A Biophysical and Structural Assessment of the Spillover Potential of Bat-Borne Sarbecoviruses

Samantha Zepeda

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

David J. Veesler, Chair

Jesse D. Bloom

Neil King

Program Authorized to Offer Degree:

Biochemistry

© Copyright 2024

Samantha Zepeda

University of Washington

Abstract

A Biophysical and Structural Assessment of the
Spillover Potential of Bat-Borne Sarbecoviruses

Samantha Zepeda

Chair of the Supervisory Committee:

David Veesler

Biochemistry

The propensity of coronaviruses to spillover from animal reservoirs into human populations has proven to have a startling impact on society in recent years. During the first two outbreaks, caused by severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002-2003 and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012-present, death tolls hovered around 800 each (World Health Organization, 2004, 2019). During the recent emergence of SARS-CoV-2 in 2019, previous death tolls were shattered as SARS-CoV-2 tore across the globe; the death toll to date stands in the millions. These viruses all began in their native reservoirs, bats, some spread to intermediate hosts palm civets (SARS-CoV-1) and dromedary camels (MERS-CoV), but all started in bats and eventually made their way into humans. They all like all other coronaviruses utilize their spike glycoproteins to mediate viral attachment and fusion to

host cells. During each of these spillover events, we were unprepared, we lacked vaccines or therapeutics, and we knew very little about what existed in animal reservoirs. Today, the situation is very much changed, vaccines and therapeutics have been developed against SARS-CoV-2, but there is still an urgent need to understand what these viruses are doing in their hosts, to identify the barriers to spillover, and characterize and stockpile vaccines and therapeutics for future spillover preparedness.

In the following dissertation, I detail the deep ancestral origins of ACE2 receptor usage in sarbecoviruses and describe the evolvability and plasticity of ACE2 binding in sarbecoviruses. I also highlight sarbecoviruses currently unable to bind and utilize human ACE2 for cellular entry but with the capacity to gain human ACE2 binding and cellular entry within 1-2 amino acid mutations. I then describe a Cambodian bat sarbecovirus, closely related to SARS-CoV-2, with the capacity to bind human ACE2 but unable to facilitate cellular entry and explore the role of spike conformational dynamics on the spillover potential of this virus. I further describe a bat-borne sarbecovirus that is much more distally related to SARS-CoV-2, PRD-0038, originating from African and explore its receptor usage, antigenicity, and immunogenicity; demonstrating that vaccination with this spike glycoprotein results in broader vaccine mismatch protection, and motivating the inclusion of clade 3 antigens in next-generation vaccines for enhanced resilience to viral evolution. I detail the discovery of ACE2 utilization in HKU5 and describe molecular basis for multiple independent evolution of ACE2 across coronaviruses. I also detail our effort towards pandemic response and zoonotic spillover preparedness; describing the discovery and characterization of several broadly neutralizing antibodies, the elicitation of broad

neutralizing antibodies with multivalent RBD-nanoparticle vaccines and describe the impact of mutation of SARS-CoV-2 Omicron variants on receptor usage, fusogenicity, and immune/therapeutic evasion. Finally, I detail the development of a ligand fitting software EMERALD, and its usage in fitting linoleic acid in spike glycoproteins and other protein datasets. Collectively, my work informs on barriers to spillover, the molecular mechanisms of coronavirus receptor usage, the evolution of sarbecoviruses in their natural reservoirs, and on the development of the next generation of vaccines and therapeutics.

,

Table of Contents

ACKNOWLEDGEMENTS 20

CHAPTER 1. INTRODUCTION 22

1.1 What is a virus? 22

1.2 Coronaviruses 23

1.3 Coronaviruses Spike Glycoproteins 23

1.4 Coronaviruses Receptor Usage and species tropism 25

1.7 Sarbecovirus Spillover Responses 26

1.8 Figures and Tables 27

CHAPTER 2. ACE2 BINDING IS AN ANCESTRAL AND EVOLVABLE TRAIT OF SARBECOVIRUSES 29

2.1 Chapter Introduction 30

2.2 Survey of ACE2 binding 31

2.3 Ancestral Origins of ACE2 Binding 34

2.4 Evolvability of ACE2 Binding 35

2.5 New Sarbecovirus lineages bind ACE2 38

2.6 Chapter Discussion 39

2.7 Methods 41

2.8 Figures & Tables 60

CHAPTER 3. RECEPTOR BINDING DOMAIN CONFORMATION ACTS AS SPILLOVER BARRIER IN BAT CAMBODIAN SARBECOVIRUS, RSHTT200 76

3.1 Chapter Introduction 76

3.2 Architecture of the RShTT200 S Trimer 78

3.3 Structural Differences in RshTT200 S result in altered antigenicity 81

3.4 Molecular Basis of RshTT200 binding of host receptor 83

3.5 Several Rhinolophus ACE2 alleles are entry receptors for RShTT200 88

3.6 Broad Tropism of RshTT200 for ACE2 orthologs Rhinolophus Species. 92

3.6 Deep Mutational Scanning Identifies mutation of concern 97

3.7 Broad Tropism of RshTT200 for ACE2 orthologs from Domesticated and Wildlife Species 99

3.8 SARS-CoV-2 Wu vaccinated polyclonal sera neutralization potency against RshTT200 Y453F 101

3.9 RBD conformational regulation modulates receptor tropism 102

3.10 Chapter Discussion 107

3.11 Methods 113
3.12 Figures & Tables 116
CHAPTER 4. BROAD RECEPTOR TROPISM AND IMMUNOGENICITY OF A CLADE 3 SARBECOVIRUS 143
4.1 Chapter Introduction 144
4.2 PRD-0038 S can utilize a broad spectrum of Rhinolophus bat ACE2 orthologs as entry receptors 146
4.3 Molecular basis of PRD-0038 RBD engagement of the *R. alcyone* ACE2 receptor 147
4.4 PRD-0038 RBD mutations enable human ACE2 utilization and expand receptor tropism to additional geographically relevant bat species 149
4.5 Architecture of the PRD-0038 S trimer 151
4.6 Antigenicity of the PRD-0038 S trimer 152
4.7 Immunogenicity of the PRD-0038 S trimer 154
4.8 Chapter Discussion 157
4.9 Methods 159
4.10 Figures & Tables 177
CHAPTER 5. MOLECULAR BASIS OF CONVERGENT EVOLUTION OF ACE2 RECEPTOR UTILIZATION AMONG HKU5 CORONAVIRUSES 198
5.1 Chapter Introduction 198
5.3 Molecular basis of HKU5 recognition of ACE2 203
5.4 Determinants of ACE2 species tropism 205
5.5 HKU5 utilization of artiodactyl and human ACE2 208
5.6 Architecture of the prefusion HKU5 S glycoprotein trimer 210
5.7 Identification of HKU5 countermeasures 212
5.8 Chapter Discussion 214
5.9 Methods 217
5.10 Figures & Tables 240
CHAPTER 6. N-TERMINAL DOMAIN ANTIGENIC MAPPING REVEALS A SITE OF VULNERABILITY FOR SARS-COV-2 267
6.1 Chapter Introduction 268
6.2 NTD-specific mAbs with potent neutralizing activity 271
6.3 NTD-specific neutralizing mAbs delineate an antigenic supersite 273
6.5 NTD mAbs are affected by the genetic diversity of circulating SARS-CoV-2 isolates and sarbecoviruses 277

6.6 SARS-CoV-2 S neutralization escape mutants reveal unconventional mechanisms of NTD mAb resistance 279
6.7 Mechanism of action of NTD-specific neutralizing mAbs 281
6.8 NTD neutralizing mAbs protect against SARS-CoV-2 challenge in hamsters 283
6.9 Discussion 284
6.10 Methods 288
6.12 Figures and Tables 314
CHAPTER 7. BROAD SARBECOVIRUS NEUTRALIZATION BY A HUMAN MONOCLONAL ANTIBODY. 329
7.1 Chapter Introduction 329
7.2 A broadly neutralizing sarbecovirus mAb 331
7.3 Structural basis of sarbecovirus neutralization 333
7.4 S2X259 has an unusually narrow escape profile 335
7.5 S2X259 protects hamsters from SARS-CoV-2 challenge 337
7.6 Discussion 339
7.7 Methods 341
7.8 Figures and Tables 364
CHAPTER 8. ANTIBODY-MEDIATED BROAD SARBECOVIRUS NEUTRALIZATION THROUGH ACE2 MOLECULAR MIMICRY. 383
8.1 Chapter Introduction 384
8.2 Identification of broadly neutralizing sarbecovirus Antibody 385
8.3 Neutralization Potency of S2K146 mAb 386
8.4 Structural Basis of S2K146 cross reactivity 387
8.5 S2K146 overcomes mutational plasticity of RBM 389
8.6 Chapter Discussion 393
8.7 Materials and Methods 394
8.9 Figures and Tables 415
CHAPTER 9. ELICITATION OF BROADLY PROTECTIVE SARBECOVIRUS IMMUNITY BY RECEPTOR-BINDING DOMAIN NANOPARTICLE VACCINES. 434
9.1 Chapter Introduction 435
9.2 Dose-sparing RBD-NP vaccination protects mice against SARS-CoV-2 challenge 438
9.3 RBD-NP vaccination induces Abs targeting diverse antigenic sites in NHPs 440

9.4 RBD-NP vaccine elicits potent neutralizing Ab responses in NHPs against a panel of SARS-CoV-2 S variants 442
9.5 RBD-NP vaccine elicits cross-reactive sarbecovirus polyclonal Abs in NHPs 446
9.6 Design, assembly, and characterization of mosaic and cocktail sarbecovirus RBD-NPs 447
9.7 Cocktail and mosaic RBD-NP vaccines elicit cross-reactive and broadly neutralizing sarbecovirus Abs 449
9.8 Mosaic RBD-NPs protect mice against heterotypic SARS-CoV challenge. 451
9.9 Chapter Discussion 453
9.10 Methods 455
9.11 Figures and Tables 480
CHAPTER 10. A POTENT PAN-SARBECOVIRUS NEUTRALIZING ANTIBODY RESILIENT TO EPITOPE DIVERSIFICATION. 496
10.1 Introduction 497
10.2 Isolation of a potent pan-sarbecovirus neutralizing mAb 499
10.3 Structural basis of VIR-7229 breadth and potency 503
10.5 VIR-7229 has an unusually high barrier to viral escape 508
10.7 Methods 517
10.8 Figures and Tables 565
CHAPTER 11. STRUCTURAL BASIS OF SARS-COV-2 OMICRON IMMUNE EVASION AND RECEPTOR ENGAGEMENT. 583
11.1 Chapter Introduction 583
11.2 Results 584
11.3 Discussion 591
11.4 Methods 592
11.5 Figures and Tables 600
CHAPTER 12. ALTERED TMPRSS2 USAGE BY SARS-COV-2 OMICRON IMPACTS INFECTIVITY AND FUSOGENICITY. 616
12.1 Chapter Introduction 616
12.2 Omicron spike-binding affinity to ACE2 617
12.3 The sensitivity of Omicron to therapeutics 618
12.4 Omicron and polyclonal antibodies 619
12.5 Omicron replication and spike cleavage 620
12.6 Omicron entry is impaired in TMPRSS2^{+} cells 622
12.7 Impaired cell–cell fusion by Omicron S 626

12.8 Discussion 628
12.9 Methods 633
12.10 Figures and Tables 655
CHAPTER 13. BROADLY NEUTRALIZING ANTIBODIES OVERCOME SARS-COV-2 OMICRON ANTIGENIC SHIFT. 680
13.1 Chapter Introduction 681
13.2 Omicron RBD binds ACE2 with increased affinity 682
13.3 Extent of Omicron escape from plasma antibodies 684
13.4 Broadly neutralizing antibodies inhibit Omicron 686
13.5 Discussion 688
13.6 Methods 689
13.7 Figures and Tables 702
CHAPTER 14. AUTOMATIC AND ACCURATE LIGAND STRUCTURE DETERMINATION GUIDED BY CRYO-ELECTRON MICROSCOPY MAPS. 725
14.1 Chapter Introduction 725
14.2 Results 727
14.3 Crystal models confirm alternate conformations for EM data 730
14.4 Docked poses reveal plausible alternate conformations 731
14.5 Low-confidence unmatched cases show pseudo-symmetry or weak density 733
14.6 Cases with worse ligand models show poor initial sampling 734
14.7 Blind modeling of linoleic acid 735
14.8 Discussion 735
14.9 Methods 736
14.9 Figures and Tables 746
CHAPTER 15. CONCLUDING REMARKS 752
REFERENCES 754

LIST OF FIGURES

Figure 1.1: Mixed open and closed conformations on SARS-CoV S glycoprotein. 27
Figure 1.2: Distinct differences in the conformation of RBD of highly pathogenic CoV compared to endemic strains. Adapted from ref 9..... 28
Figure 2.1: High-throughput survey of sarbecovirus ACE2 binding. 60
Figure 2.2: Ancestral origins of sarbecovirus ACE2 binding. 61
Figure 2.3: Evolutionary plasticity of ACE2 binding. 62
Figure 2.4: Newly sampled sarbecovirus lineages bind to ACE2. 63
Figure 3.1: Architecture of the RshTT200 S glycoprotein trimer. 116
Figure 3.2: RshTT200 binding and key molecular interactions with host R. shameli ACE2 122
Figure 3.3: RshTT200 binding and key molecular interactions with Rhinolophus ACE2 panel..... 132
Figure 3.4: RshTT200 Tropism across Domesticated and Wildlife Species ACE2..... 137
Figure 3.5: Polyclonal Sera from patients vaccinated 2-3x against SARS-CoV-2 Wuhan demonstrates equivalent or improved neutralization potency against RshT200 453F S VSV. 137
Figure 3.6: Impact of RshTT200 Y453F mutation on binding to Rhinophus and Human ACE2. 139
Figure 3.7: RshTT200 RBD in the SARS-CoV-1 S enables human ACE2 mediated cellular entry due to RBD conformation..... 140
Figure 4.1: The clade 3 PRD-0038 sarbecovirus has a broad Rhinolophus bat ACE2 tropism..... 178
Figure 4.2: PRD-0038 RBD amino acid mutations broaden receptor tropism. 179
Figure 4.3: Architecture and antigenicity of the PRD-0038 S trimer. 180
Figure 4.4: A clade 3 sarbecovirus S trimer elicits broadly reactive antibody responses. . 181
Figure 5.1: Identification of several mammalian ACE2s as functional HKU5 entry receptors. 240
Figure 5.2: Molecular basis of HKU5 recognition of the P. abramus ACE2 receptor. 242
Figure 5.3: HKU5 molecular determinants of ACE2 host species tropism 244
Figure 5.4: HKU5 utilization of artiodactyl and human ACE2 receptors..... 246
Figure 5.5: Architecture and conservation of the HKU5 S glycoprotein trimer. 247
Figure 5.6: Identification of countermeasures against HKU5 merbecoviruses..... 248
Figure 5.7: Coronaviruses have evolved ACE2 utilization at least five times independently. 249
Figure 6.1 Discovery of potent NTD-specific SARS-CoV-2 neutralizing mAbs from three convalescent individuals. 315
Figure 6.2: SARS-CoV-2 NTD neutralizing mAbs target the same antigenic supersite. ... 316
Figure 6.3: 3 The SARS-CoV-2 NTD comprises multiple antigenic sites. 317
Figure 6.4: NTD mAbs are affected by the genetic diversity of circulating SARS-CoV-2 variants and sarbecoviruses. 318
Figure 6.5: Analysis of SARS-CoV-2 S neutralization escape mutants. 319
Figure 6.6: Mechanism of action of NTD-specific neutralizing mAbs. 320

Figure 6.7: S2X333 provides robust in vivo protection against SARS-CoV-2 challenge but selects for escape mutants. 321
Figure 7.1: Identification of a broadly neutralizing sarbecovirus mAb. 364
Figure 7.2: The S2X259 broadly neutralizing sarbecovirus mAb recognizes RBD antigenic site II.| 365
Figure 7.3: S2X259 is resilient to a broad spectrum of escape mutations. 366
Figure 7.4: S2X259 protects hamsters challenged with prototypic and B.1.351 SARS-CoV-2. 367
Figure 8.1: . Identification of the S2K146 cross-reactive and broadly neutralizing sarbecovirus mAb. 415
Figure 8.2: The S2K146 broadly neutralizing mAb recognizes RBD antigenic site I. 418
Figure 8.3: S2K146 is resilient to a broad spectrum of escape mutations. 419
Figure 8.4: S2K146 blocks receptor attachment, triggers premature S refolding, and protects against SARS-CoV-2 challenge therapeutically. 422
Figure 9.1: A single immunization with RBD-NP protects BALB/c cByJ mice from SARS-CoV-2 MA10 challenge. 480
Figure 9.2: RBD-NP vaccination elicits high titers of Abs targeting diverse RBD antigenic sites in NHPs. 481
Figure 9.3: RBD-NP and HexaPro S elicit Abs with similar neutralization breadth toward SARS-CoV-2 variants. 485
Figure 9.4: In vitro assembly and accelerated stability studies of mosaic and cocktail nanoparticle immunogens. 487
Figure 9.5: Mosaic and cocktail RBD-NP vaccines elicit neutralizing Abs against multiple sarbecoviruses. 489
Figure 9.6: Mosaic and cocktail RBD-NP vaccines protect against heterotypic SARS-CoV-MA15 challenge in 15-week-old BALB/c cByJ mice. 491
Figure 10.1: VIR-7229 is a potent, pan-sarbecovirus neutralizing mAb. 565
Figure 10.2: In vivo efficacy of VIR-7229. 566
Figure 10.3: Structural basis for VIR-7229 pan-sarbecovirus neutralization. 567
Figure 10.4: VIR-7229 epitope encompasses sarbecovirus diversity. 568
Figure 10.5: VIR-7229 has high tolerance for SARS-CoV-2 epitope variation. 569
Figure 10.6: VIR-7229 exhibits a high barrier to viral escape. 570
Figure 11.1: Cryo-EM structure of the SARS-CoV-2 Omicron S trimer reveals a remodeling of the NTD antigenic supersite. 600
Figure 11.2: SARS-CoV-2 Omicron S mutations outside the NTD and RBD. 601
Figure 11.3: SARS-CoV-2 Omicron RBD mutations promote escape from a panel of clinical mAbs. 603
Figure 11.4: Molecular basis of human ACE2 recognition by the SARS-CoV-2 Omicron RBD. 604
Figure 12.1: SARS-CoV-2 Omicron and Delta Variant live virus replication in 3D tissue culture systems and 2D cell lines. 655
Figure 12.2: SARS-CoV-2 Omicron and Delta Variant live virus replication in 3D tissue culture systems and 2D cell lines. 656
Figure 12.3: SARS-CoV-2 Omicron variant spike enters cells less efficiently by TMPRSS2-dependent plasma membrane fusion. 658

Figure 12.4: SARS-CoV-2 Omicron variant spike shows impaired cell–cell fusion activity and smaller infection foci generated by live virus. 660
Figure 13.1: The Omicron RBD shows increased binding to human ACE2 and gains binding to mouse ACE2. 702
Figure 13.2: Neutralization of Omicron SARS-CoV-2 VSV pseudovirus by plasma from convalescent and vaccinated individuals. 704
Figure 13.3: Neutralization of Omicron SARS-CoV-2 VSV pseudovirus by clinical-stage monoclonal antibodies 705
Figure 13.4: Neutralization of Omicron SARS-CoV-2 VSV pseudovirus by monoclonal antibodies. 707
Figure 14.1: Overview of EMERALD docking protocol. 746
Figure 14.2: Benchmarking EMERALD against the EMDB. 746
Figure 14.3: Alternate conformations predicted with EMERALD match high- resolution crystal structures. 748
Figure 14.4: Alternate conformations found by EMERALD in cases without crystal structures 749
Figure 14.5: Examples of low-confidence docked models where there may be ambi- guity or heterogeneity in the data 750
Figure 14.6: Blind modeling of linoleic acid. 751

List of Supplementary Tables

Supplemental Figure 2.1: : Robustness of the root of the sarbecovirus ingroup. 64
Supplemental Figure 2.2: Experimental details of Sort-seq assays. 65
Supplemental Figure 2.3: Normalization and controls for biolayer interferometry binding and pseudovirus entry assays. 66
Supplemental Figure 2.4: Clade 2 RBD binding to an expanded panel of R. sinicus ACE2 alleles. 67
Supplemental Figure 2.5: Full set of RBD ancestral sequence reconstructions. 68
Supplemental Figure 2.6: Robustness to uncertainties in ancestral reconstructions. 69
Supplemental Figure 2.7: Binding of RBD single mutants to each ACE2. 71
Supplemental Figure 2.8: Existing data on sarbecoviruses in mice, and affinities of RBDs and key mutants for mouse versus human ACE2. 72
Supplemental Figure 2.9: Epistasis and turnover in mutational effects. 73
Supplemental Figure 2.10: Robustness of rooting and AncSarbecovirus phenotype in a phylogeny incorporating newly reported sequences. 74
Supplemental Figure 2.11: Uncropped images of the western blots in Fig. 2.7c. 75
***Supplemental Figure 3.1:* Stabilization of RshTT200 S ectodomain. 117**
Supplemental Figure 3.2: Data Processing Pipeline for Cryo-EM structural reconstruction of the RshTT200 S. 118
Supplemental Figure 3.3: Conservation of Linoleic Acid Binding Pocket 119
Supplemental Figure 3.4: Monoclonal antibody binding to RshTT200 and SARS-CoV-2 Spikes. 120
Supplemental Figure 3.5: Biolayer Interferometry Measuring Dimeric R. shameli ACE2 Binding. Binding. 123
Supplemental Figure 3.6: Biolayer Interferometry Measuring Dimeric Human ACE2 Binding. 124
Supplemental Figure 3.7: Normalization of RshTT200 S and SARS-CoV-2 S VSV pseudotyped particles for cell entry. 126
Supplemental Figure 3.8: Data Processing Pipeline for Cryo-EM structural reconstruction of the RshTT200 RBD in complex with R. shameli ACE2. 127
Supplemental Figure 3.9: Biolayer Interferometry Comparison of the R. shameli ACE2- Fc dimerized, and natively dimerized R. shameli ACE2. 128
Supplemental Figure 3.10: RshTT200 interactions with R. shameli ACE2. 129
Supplemental Figure 3.11: Polymorphic variation of R. affinis ACE2 impacts sarbecovirus-ACE2 binding. 130
Supplemental Figure 3.12: Polymorphic variation of R. sinicus ACE2 impacts sarbecovirus-ACE2 binding. 131
Supplemental Figure 3.13: Molecular Basis of RshTT200 RBD Interactions with Rhinolophus bat ACE2 receptors. 133
Supplemental Figure 3.14: Representative Western Blot of Spike Normalized RshTT200 Chimera S VSV 134
Supplemental Figure 3.15: Polymorphic variation of R. affinis and R. sinicus ACE2 impact on cellular entry. 135
Supplemental Figure 3.16: Biolayer Interferometry Immobilized Sarbecovirus Binding of hACE2 Fc Fusion ectodomains at 1 µM. 136

Supplemental Figure 3.17: Representative BLI Binding of immobilized RaTG13 RBD with titration of RA9479 ACE2. 136
Supplemental Figure 3.18: Polyclonal Sera Neutralization of RshTT200 453F (royal blue) and SARS-CoV-2 D614G S VSV 138
Supplemental Figure 3.19: RaTG13 RBD binds human ACE2 roughly equivalent to RshTT200 453F. 139
Supplemental Figure 3.20: D614G mutation has no impact on cell entry in RshTT200 S. 141
***Supplemental Figure 3.21:* Co-transfection of TMPRSS2 does not enable human ACE2 cellular entry by RshTT200 VSV. 141**
***Supplemental Figure 3.22:* Limited Promiscuity of *R. shameli* ACE2 across sarbecovirus clades. 142**
Supplemental Figure 4.1: CryoEM data collection and refinement of the dimeric R. alcyone ACE2-bound PRD-0038 RBD complex (related to Fig. 4.1). 183
Supplemental Figure 4.2: Structural distinctions between the PRD-0038, SARS-CoV-2 and SARS-CoV-1 RBDs near the ACE2-binding interface (RBM) (related to Fig. 4.1) 184
Supplemental Figure 4.3: Heatmaps of change in ACE2-binding avidity resulting from RBD mutations determined by DMS (related to Fig. 4.2). 185
Supplemental Figure 4.4: Binding data and Western blot analysis (related to Fig. 4.2). ... 185
Supplemental Figure 4.5: CryoEM data collection and refinement of PRD-0038 S (related to Fig. 4.3) 187
Supplemental Figure 4.6: Structural resemblance of PRD-0038 S to SARS-CoV-2 and SARS-CoV-1 (related to Fig. 4.3) 188
Supplemental Figure 4.7: Conservation analysis of epitopes targeted by monoclonal antibodies between PRD-0038 S and SARS-CoV-2 S (related to Fig. 4.3) 189
Supplemental Figure 4.8: Dose-response curves of mouse serum neutralization before and after three immunizations with SARS-CoV-2 HexaPro S or PRD-0038 HexaPro S (related to Fig. 4.4) 191
Supplemental Figure 4.9: Neutralization of SARS-CoV-1 S VSV by vaccine-elicited mouse sera using a highly diluted pseudovirus input (related to Fig. 4.4) 192
Supplemental Figure 4.10: Binding of vaccine-elicited mouse sera to yeast-displayed sarbecovirus RBDs (related to Fig. 4.4) 193
Supplemental Figure 4.11: DMS epitope mapping of serum antibodies (related to Fig. 4.4). 194
Supplemental Figure 5.1: related to Fig. 1. HKU5 utilizes several mammalian ACE2s as receptor. 251
Supplemental Figure 5.2, related to Fig. 1. Evaluation of receptor species tropism for several HKU5 strains. 253
Supplemental Figure 5.3, related to Fig. 1 and 4. HKU5 S proteolytic processing during biogenesis 254
Supplemental Figure 5.4, related to Fig. 2 and 4. Cryo-EM data processing of the P.abr ACE2-bound HKU5-19s RBD and B.tau ACE2-bound HKU5-19s RBD datasets. 257
Supplemental Figure 5.5, related to Fig. 2 and 4. Biolayer interferometry analysis of the P.abr, B.tau or N.vis ACE2 ectodomain dimers binding to immobilized HKU5-19s, HKU5-33s or HKU5-1 RBDs. 257
Supplemental Figure 5.6, , related to Fig. 5. Cryo-EM data processing of the prefusion HKU5-19s S trimer dataset. 259

Supplemental Figure 5.7, related to Fig. 6. Dose-response curves of HKU5-19s VSV pseudovirus neutralization by MERS-CoV or SARS-CoV-2 infection-elicited plasma or monoclonal antibodies..... 260
Supplemental Figure 6.1: Characteristics of SARS-CoV-2 NTD mAbs related to Fig. 1. . 323
Supplemental Figure 6.2: Cryo-EM data processing of SARS-CoV-2 S bound to S2L28, S2M28, or S2X333, related to Fig. 2. 324
Supplemental Figure 6.3: Cryo-EM data processing of SARS-CoV-2 S bound to S2X28, S2L20, S2X316, or S2M24, related to Figs. 2 and 3. 325
Supplemental Figure 6.4: NTD neutralizing mAbs inhibit S-mediated entry of the closely related RaTG13 but not of more distant sarbecoviruses, related to Fig. 4 326
Supplemental Figure 6.5: Neutralizing activity of SARS-CoV-2 NTD- and RBD-targeting mAb cocktails, related to Fig.s 1 and 6..... 327
Supplemental Figure 6.6: Sequence alignment of SARS-CoV-2 S vRNA from input virus, control group, and selected outlier animals administered with S2X333, related to Fig. 7. 328
Supplemental Figure 8.1: Overview of SARS-CoV-2 RBD cross-neutralizing antigenic sites. 423
Supplemental Figure 8.2: In vitro characterization of the S2K146 mAb 424
Supplemental Figure 8.3: Evaluation of mAb binding to the BtKY72 RBD. 425
Supplemental Figure 8.4: CryoEM data processing of the S2K146-bound SARS-CoV-2 S dataset..... 426
Supplemental Figure 8.5: S2K146 epitope conservation across sarbecoviruses. 427
Supplemental Figure 8.6: SARS-CoV-2 broadly neutralizing sarbecovirus mAbs. 428
Supplemental Figure 8.7: DMS analysis of the yeast-displayed SARS-CoV-2 RBD. 428
Supplemental Figure 8.8: S2K146 escape clone selection by plaque assays using VSV-SARS-2. 429
Supplemental Figure 8.9: Activation of FcγγRIIa and FcγγRIIIa in vitro 430
Supplemental Figure 8.10: Correlation between mAb concentration and viral loads in the lungs of 431
Supplemental Figure 9.1: Effects of mutations on binding of HCP Abs to RBD and FACS gating strategy, related to Fig. 2..... 482
Supplemental Figure 9.2: Evaluation of vaccine-elicited binding and neutralizing Ab titers against SARS-CoV-2 variants and sarbecoviruses, related to Fig. 3..... 484
Supplemental Figure 9.3: In vitro characterization and confirmation of co-display of sarbecovirus RBD-NP immunogens, related to Fig. 4. 488
Supplemental Figure 9.4: Serum Ab binding titers elicited by mosaic and cocktail RBD-NPs, related to Fig. 5. 490
Supplemental Figure 9.5: Monovalent, mosaic, and cocktail RBD-NPs protect against heterotypic SARS-CoV-MA15 challenge in 15-week-old BALB/c cByJ mice, related to Fig. 6..... 492
Supplemental Figure 10.1: Potency and breadth of comparator mAbs. Related to Fig. 1. 572
Supplemental Figure 10.2: VIR-7229 mechanisms of action. Related to Figs. 1 and 2. 573
Supplemental Figure 10.3: Cryo-EM data processing and validation of VIR-7229-bound BA.2.86 S. Related to Fig. 3. 574
Supplemental Figure 10.4: Structural analysis of VIR-7229:RBD and ACE2:RBD. Related to Figs. 3 and 5. 575

Supplemental Figure 10.5: Deep mutational scanning profiling of VIR-7229, S2V29, and S2K146 escape mutations. Related to Fig. 5. 577
Supplemental Figure 10.6: Validation and fitness of VIR-7229 escape mutations. Related to Figs. 5 and 6. 578
Supplemental Figure 11.1: CryoEM data processing and validation. 605
Supplemental Figure 11.2: Density resolving the Omicron S mutations. 606
Supplemental Figure 11.3: Comparison of Omicron and Wuhan-Hu-1 NTDs. Omicron. 607
Supplemental Figure 11.4: SARS-CoV-2 Omicron mutations promote escape from clinical mAb binding in the context of the full S ectodomain trimer..... 608
Supplemental Figure 11.5: Structural basis for S309 binding to the Omicron, Kappa and Wuhan-Hu-1 RBD..... 609
Supplemental Figure 11.6: Comparison of the region comprising residues 366 to 375 across structures, and their impact on S2X35 binding..... 610
Supplemental Figure 11.7: The SARS-CoV-2 Omicron RBD recognizes mouse ACE2 efficiently. 611
Supplemental Figure 13.1: Neutralization of WT and Omicron S pseudotyped SARS-CoV-2 virus neutralization by plasma from COVID-19 convalescent and vaccinated individuals. 719
Supplemental Figure 13.2: Neutralization of WT and Omicron S pseudotyped SARS-CoV-2 virus neutralization by plasma from COVID-19 convalescent and vaccinated individuals. 720
Supplemental Figure 13.3: Neutralization of WT and Omicron S pseudotyped SARS-CoV-2 by plasma from vaccinated dialysis patients..... 721

LIST OF TABLES

Table 5-1, related to Fig. 2, 4 and 5. Cryo-EM data collection and refinement statistics. . 261
Table 5-2, related to Fig 2.4. Conservation analysis of P.abr ACE2 residues recognized by HKU5 in hACE2. Non-conserved residues are italicized. 262
Table 5.3, related to Fig. 6. Demographic information of MERS-CoV-infected and SARS-CoV-2-infected individuals. 264

LIST OF SUPPLEMENTARY TABLES

Table S3.1: Cα RMSD between RshTT200 and SARS-CoV-2 Wuhan (PDB: 6zb5). 119
Table S3.2: Conservation/Divergence of Glycosylation Sequon. 120
Table S3.3: Kinetic Analysis of ACE2 binding to sarbecovirus RBD by Biolayer Affinity 124
***Table S3.4:* ACE2-RBD Interface Residues across Rhinolophus bat species. 130**
Table S4.1: Rhinolophus ACE2 alleles and contact residues (related to Fig. 4.1). 195
Table S 4.2: CryoEM data collection and refinement statistics (related to Fig. 4.2 and 4.3). 196
Table S 4.3: Representative KD values determined from BLI binding analysis (related to Fig. 4.2). 197
Table S 7.1: Characteristics of S2X259. VH and VL % identity refers to V gene identity compared to germline (IMGT). 379
Table S 7.2: S2X259-mediated neutralization (IC50) of VSV pseudotypes harbouring SARS-CoV-2 S representing VOC, single mutants and different sarbecovirus clades. 379
Table S 7.3: CryoEM data collection and refinement statistics. 380
Table S 7.4: X-ray data collection and refinement statistics. 381
Table S 7.5: Summary of nucleotide and amino acid mutations found in 18 neutralization-resistant VSV-SARS-CoV-2-S chimera plaques. 382
Table S 8.1: CryoEM data collection and refinement statistics. 432
Table S 8.2: Summary of nucleotide and amino acid mutations found in 36 neutralization-resistant VSV-SARS-CoV-2-S D614G chimera plaques. 433
Table S 10.1: CryoEM data collection and refinement statistics 579
Table S10.2: Crystallographic data collection and refinement statistics. 580
Table S10.3: Bioinformatic analysis of intra-individual SARS-CoV-2 genomic variability. 581
Table S 11.1 CryoEM data collection, processing and model refinement statistics. 612
Table S 11.2: Crystallographic data collection and refinement statistics. 613
Table S 11.3: Kinetic parameters and fitting details of the SPR RBD binding measurements. 614
Table S13.1: Characteristics of single point mutations present in Omicron RBD relative to Wuhan-Hu-1 RBD. 724

ACKNOWLEDGEMENTS

Through my time at the University of Washington, I have been fortunate to be surrounded by a supportive community of mentors, colleagues, collaborators, and friends. A community that has been present for me as a student, a scientist, and as an individual. Throughout personal and professional triumphs and setbacks, and a global pandemic that altered the way we all work this community has made my Ph.D. journey an enjoyable and memorable journey.

First and foremost, I would like to thank my PhD advisor, David Veesler. I have grown immensely in my years in your lab. Your patience through scientific, professional, and personal challenges throughout my graduate school journey enabled my success. Your scientific rigor, work ethic, and leadership have been a huge inspiration to me. The balance you've perfected of being a highly present and involved mentor and allowing me to explore things independently was exactly what I needed at this career stage, and I am deeply grateful to you. The training and professional development I received in your lab were unparalleled and I will forever be honored to be a Veesler lab alumni.

Next, to my colleagues, both past and present members of the Veesler lab, I have enjoyed working with you all. You have created a supportive and fun lab environment, which makes coming to lab an exciting and enjoyable part of each day. I have learned so much from each of you. To Ha and Lexi, thank you for your mentorship, your friendship and your support. I cannot express how much I respect you both. Daniel, Matt, Ollie, Max, and Amin you all have contributed to my training in different but equally impactful ways. Kaiti, thank you for being the awesome bay-mate, hood-mate, and friend that you are. Jimin, Risako and Courtney, you all have been amazing colleagues and friends, and I am grateful to each of you. I will never forget our little adventures and outings. John, Cameron, Jack, Anshu, Cecily, Emma, Albert and Ben it has been a pleasure to work with you all.

To my collaborators on various projects, thank you. It has been a privilege to work with such an amazing team of scientists. First, thank you Jesse Bloom for your time and advice and allowing me to come to your lab to pilot experiments with the assistance of your team. Allison Greaney, thank you for your assistance and insight. Tyler Starr, you are a kind and generous collaborator, and I am so grateful to you not only for our work together but also for the time spent springboarding ideas. You have been a mentor and an inspiration. To our collaborators in the King Lab, especially Cara Chao, Dan Ellis, Grace Hendricks, thank you all for your time and generosity. I have also had the pleasure of working with Andrew Muenks and the DiMaio lab. Thank you for your assistance with ligand fitting; your time and advice were invaluable! A special shout out to Matt, Young, Ale, Lexi, Jimin, and Amin. You all have been amazing collaborators, and it has been a pleasure to work alongside you. Lastly, I would like to thank our protein production team, John, Jack, and Cameron who have contributed to many of my projects and have been invaluable members of the team.

To the rest of my past and present scientific mentors, I am grateful for all your advice and guidance. Dr. Lewis, my undergraduate and masters training in your lab gave me a solid foundation to succeed in my Ph.D. Dr. Steve Whitten, your support and advice continues to assist me to this day. Dr. Ha V. Dang, I am so very grateful to have had you as my rotation mentor. Dr. Lexi Walls, you have been an amazing inspiration and your investment not only in me but in all your colleagues speaks not only to what an amazing scientist you are but also to what an amazing human you are. Dr. Park, Dr. Tortorici, Dr. Dickinson and Dr. Quispe thank you all for taking the time to teach and mentor me, I could not have grown as scientist as much as I have without your investment in my training.

Lastly, to my friends and family, I would like to thank you for your support. To my husband, Eric who has supported me throughout my Ph.D. When I moved halfway across the country, you continued to be a source of unending encouragement. I am grateful to have had you by my side during this journey. To my siblings, Jay, Kai, Saimai, Satang, Kayla, Sarah, and Sophia you have been my inspiration from the moment we first met. I am incredibly blessed to have you all in my life. To my closest friends Sarah, Chelsea, Elena, and Taylor you have been my cheerleaders, and I will always cherish you. Finally, I would like to thank my Nana Sonja, my Mee-maw Jerry who saw me start this journey but will not get a chance to see it end.

CHAPTER 1. INTRODUCTION

1.1 What is a virus?

Viruses are small intracellular parasites that hijack host machinery to carry out their life cycle and infect all cellular life forms (vertebrate and invertebrate animals, plants, fungi, bacteria and archaea) (Carter and Saunders, 2007). There are multiple levels of classification based on their genetic building blocks (RNA/DNA; single or double stranded; positive or negative sense), on whether they are enveloped, as well as well as their genomic relationships to one another, and mode of replication (Carter and Saunders, 2007, van Regenmortel, 2004; Gelderblom, 1996). A fully assembled virus is called a virion, and the simplest organization of virions consist of two components, nucleic acid and a protein coat called a capsid. The capsid is a self-assembling unit and in comprised of only one or a few structural proteins (Gelderblom, 1996). Enveloped viruses include an addition layer of coating referred to as an envelope that is comprised of a lipid bilayer that contains membrane associated or anchoring protein (Gelderblom, 1996). While envelopes are considered protective to the viral capsid and genome (Gelderblom, 1996), they are also decorated with glycosylated proteins.

For many enveloped viruses some of these glycoproteins play critical roles in attachment to host cells and fusion of the viral envelope and host membranes (Hoffmann, 2004; Walls, 2020; Li, 2023, Wang, 2022). However, these glycoproteins can also act a weakness for these viruses by acting as immunogenic targets (Luo, 2012; Ellis, 2022; Dang, 2019; Dang, 2021; Walls, 2020; Piccoli et al. 2020; Bowen et al. 2022; Greaney et al. 2021).

1.2 Coronaviruses

One family of viruses of particular interest is coronaviruses. There have been three major outbreaks of severe and often fatal pneumonia due to zoonotic coronaviruses (CoV). During the previous two outbreaks, caused by severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002-2003 and Middle-East respiratory syndrome coronavirus (MERS-CoV) in 2012-present, death tolls hovered around 800 each (WHO, 2004; WHO, 2019). During the recent emergence of SARS-CoV-2, the world was brought to a halt, global death tolls exceeded 500,000 and the total infected reached over 10 million in 6 months' time (WHO, 2020). This highlighted a dire need to understand the factors that allow zoonotic CoVs to spillover in humans and to screen for the emergence of zoonotic CoVs with the potential to cross the species barrier. Understanding of the molecular determinants of CoV spillover and screening wildlife can allow for the development of early warning systems. As two of the three major spillover events, originated from the sarbecovirus branch, including SARS-CoV-1 and SARS-CoV-2 sarbecoviruses are of particular interest.

1.3 Coronaviruses Spike Glycoproteins

To prepare for future spillover, we must understand the steps that enable successful infection. One of the earliest, and fortunately targetable steps of infection is cellular entry which is one of the key factors for spillover of viruses into humans. Mediating this early step of infection is the spike glycoprotein (S). The S initiates coronavirus infection by promoting binding to host receptors and fusion of the viral and host membranes ((Walls et al. 2016; Walls et al. 2017; Tortorici and Veesler 2019; Walls et al. 2019;

Walls et al. 2020) Furthermore, S is the predominant target of neutralizing antibodies, due to its essential role in cellular entry and its crown-like organization at the surface of coronavirus particles (Piccoli et al. 2020; Bowen et al. 2022; Greaney et al. 2021). As a result, coronavirus S participates in determining host and tissue tropism as well as pathogenicity, and is a key factor governing the ability of coronaviruses to cross the species barrier and trigger outbreaks (Menachery et al. 2015). Studies of sarbecovirus S glycoproteins enable the identification of viruses with a high likelihood of spillover to humans and inform the deployment of monoclonal antibody therapies and the design of broadly protective vaccines (Xiong et al. 2022).

The cleavage of the CoV S protein into S1 and S2 subunits has been demonstrated to be an important step in infectivity (Hoffmann, 2020) and pathogenicity (Lau, 2020). In SARS-CoV-2 the presence of a novel polybasic furin cleavage site at the spike S1/S2 boundary, which is unique within the sarbecovirus subgenus, leads to questions about how this site may have impacted the spillover of this virus into humans. The presence of such polybasic cleavage sites has been a signature feature of other highly pathogenic virsuses (Nao, 2017). Additionally, the presence of insertions at this site may be more prevalent than previously realized as a novel bat CoV, RmYN02, has been identified that contains an insertion of several amino acids at the S1/S2 boundary (Xiao, 2020). In SARS-CoV-2, it has been determined that the insertion of the polybasic furin cleavage site is important for cellular entry into Calu-3 cells and for pathogenesis in a rodent model of infection (Hoffmann, 2020).

All three highly pathogenic CoVs (MERS-CoV, SARS-CoV, and SARS CoV-2) were found to exist in multiple conformational states, including partially open states

(**Fig. 1.1**) (Gui, 2017' Kirchdoerfer, 2018; Pallsen, 2017; Song, 2018; Walls, 2019; Yuan, 2017). This is distinct from endemic strains of CoVs that cause colds in humans (**Fig. 1.2**) (Tortorici and Veesler, 2019; Walls, 2016; Kirchdoefer, 2016; Wong, 2017; Wu, 2009). This leads to the hypothesis that presence of spontaneously open conformations is important in pathogenicity.

1.4 Coronaviruses Receptor Usage and species tropism

Successful viral entry into human cells of two highly pathogenic CoVs (SARS-CoV and SARS-CoV-2) relies on hACE2 as the entry receptor (Li, 2003; Walls, 2004). Furthermore, the binding affinity for hACE2 of various SARS-CoV isolates from the 2002-2003 epidemics was shown to be a key factor of transmissibility and viral replication in different species (Li, 2003; Li, 2005a; Guan, 2003; Wan, 2020). However, bats were identified as the reservoir host in the transmission of SARS-CoV, whereas palm civets are believed to have served as intermediate hosts (Graham, 2013; Ge, 2013; Yang, 2015). The SARS-CoV-2 reservoir host is most likely to also be bats, and pangolins have been investigated as a possible intermediate host (Lam, 2020; Zhou, 2020; Xiao, 2020; Zhou, 2020), more recently however identification of viruses most closely related to SARS-CoV-2 have been identified in bats (Temmam et al. 2022) and an intermediate host has not been identified.

It is also important to understand the receptor usage of these viruses in their native reservoirs. Furthermore, learning how these viruses spread and evolve in their native reservoirs can help us prepare for the next pandemic. It has been shown that co-infections with two or more sarbecoviruses can occur in wild bats (Wang et al. 2022). Moreover, co-roosting of some of these bat species may favor viral transmission,

evolution and recombination. Additionally, clustering of lineages of CoV into geographical regions has been demonstrated (Yu, 2019) further supporting the idea that viruses spread and evolve between species with range overlap.

1.7 Sarbecovirus Spillover Responses

Early on in the COVID pandemic, responses to spillover were largely reactive; treatments had limited efficacy, and no vaccines existed. Even though there were warnings that several sarbecoviruses circulating in nature, also had the capacity to utilize human ACE2 and the alarm was raised that future spillover from sarbecoviruses were likely following the SARS-CoV-1 outbreak (Menachery, 2015; Menachery, 2016; Hu, 2017; Ge, 2013). While vaccines and therapeutics have been developed against SARS-CoV-2 and its variants; a proactive approach is necessary to look forward and prepare for future spillover.

The work I will present in the subsequent chapters begins with studies on bat-borne sarbecoviruses and explores their spike glycoprotein structures, receptor usage, barriers to spill over, antigenicity, and immunogenicity. The following chapters are focused on pandemic response, exploring SARS-CoV-2 variants evolution and the susceptibility of SARS-CoV-2 variants and bat-borne sarbecoviruses to vaccines and therapeutic treatments that have been designed and/or isolated throughout the course of the pandemic. The final chapter describes the development of a ligand fitting software, EMERALD, that was used to fit linoleic acid in our sarbecovirus structure, a molecule that is correlated to tightly closed or “locked” conformations of sarbecovirus spike.

1.8 Figures and Tables

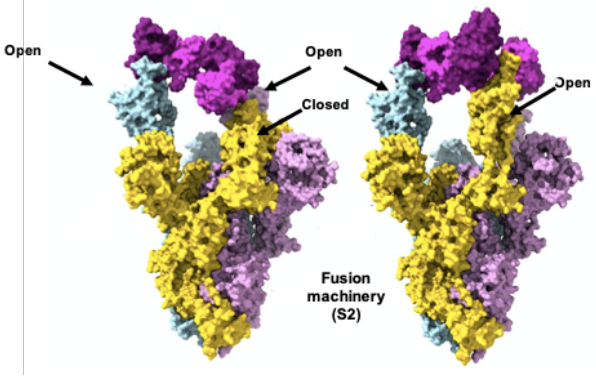


Figure 1.1: Mixed open and closed conformations on SARS-CoV S glycoprotein.

Cryo-EM structures of the SARS-CoV glycoprotein in complex with the S230 neutralizing antibody (magenta). Molecular surface representation of complex with one open, one partially open, and one closed receptor binding domain, PDB:6NB6 (left) and with three open domains (right) PDB:6NB7. Each protomer of the S glycoprotein is rendered with different colors (light blue, gold, and plum).

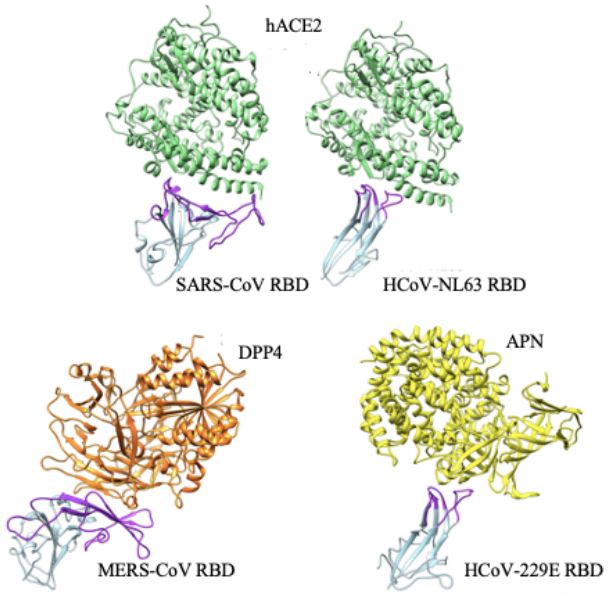


Figure 1.2: Distinct differences in the conformation of RBD of highly pathogenic CoV compared to endemic strains. Adapted from ref 9.

Ribbon diagrams of the RBD of two highly pathogenic strains of CoV, SARS-CoV (PDB: 2AJF) and MERS CoV (PDB: 4L72) (left) and two endemic strains of CoV, HCoV-NL63 (PDB: 3KBH) and HCoV-229E (PDB: 6ATK) (right) are shown in complex with their receptors. The receptors hACE2 (green), DPP4 (peach), and APN (yellow) are shown in complex with the RBD (light blue). The receptor binding motifs are rendered in purple.

CHAPTER 2. ACE2 BINDING IS AN ANCESTRAL AND EVOLVABLE TRAIT OF SARBECOVIRUSES

The first story included was initiated early in the COVID-19 pandemic when little was known about sarbecoviruses that were circulating in nature. The emergence of SARS-CoV-2 was the third major outbreak of severe and often fatal pneumonia due to zoonotic CoV. During the previous two outbreaks, caused by severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002-2003 and Middle-East respiratory syndrome coronavirus (MERS-CoV) in 2012-present, death tolls hovered around 800 each. At the start of this project, the world has been brought to a halt, global death tolls exceeded 500,000 and the total infected were over 10 million in 6 months' time.

In the subsequent sections, I detail our efforts to understand receptor usage of bat-borne sarbecoviruses in and outside of Asia, describing the evolutionary history of ACE2 binding across a diverse range of sarbecoviruses and ACE2 orthologs. I also detail our efforts to understand the evolvability of ACE2 usage amongst sarbecoviruses and describe how the effects of individual mutations can differ considerably between viruses. Our results point to the deep ancestral origin and evolutionary plasticity of ACE2 binding, broadening the range of sarbecoviruses that should be considered to have spillover potential.

Adapted from: Starr, T.N., Zepeda, S.K., Walls, A.C. *et al.* ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. *Nature* **603**, 913–918 (2022). https://doi.org/10.1038/s41586-022-04464-z

2.1 Chapter Introduction

Both SARS-CoV-2 and SARS-CoV-1 use human ACE2 as their receptor (Zhou, 2020; Li, 2003; Walls, 2020; Wrapp, 2020; Hoffmann, 2020). Sampling of bats has identified multiple lineages of sarbecoviruses with receptor-binding domains (RBDs) exhibiting different ACE2-binding properties (Letko, 2020; Hu, 2005; Li, 2005; Lau, 2005; Ge, 2013; Starr, 2020; Tao, 2019; Drexler, 2019; Zhou, 2021; Liu, 2021; Ren, 2008; Becker 2008; Wells, 2021) that are exchanged through recombination (Hu, 2005; Wells, 2021; Boni, 2020). Before the emergence of SARS-CoV-2, all bat sarbecoviruses with a demonstrated ability to bind to any ACE2 orthologue contained RBDs related to SARS-CoV-1 and were sampled from *Rhinolophus sinicus* and *Rhinolophus affinis* bats in Yunnan province in southwest China (Letko, 2020; Hu, 2017; Ge, 2013; He, 2014;Yang, 2015). More recently, sarbecoviruses related to SARS-CoV-2 that bind to ACE2 have been found more widely across Asia and from a broader diversity of *Rhinolophus* species (Zhou, 2020; Liu, 2008; Lam, 2020; Murakami, 2020; Delaune, 2021). However, ACE2 binding has not been observed within a prevalent group of sarbecovirus RBDs sampled in southeast Asia (RBD clade 2) (Letko, 2020; Hu, 2017, Ren, 2008), nor has it been observed in distantly related sarbecoviruses from Africa and Europe (RBD clade 3) (Letko, 2020; Wells, 2021) (**Fig. 2.1a**). It is therefore unclear whether ACE2 binding is an ancestral trait of sarbecovirus RBDs that has been lost in some RBD lineages, or a trait that was acquired more recently in a subset of Asian sarbecovirus RBDs (Wells, 2021; Boni, 2020). As ACE2 is also variable among *Rhinolophus* bats, particularly in the surface recognized by sarbecoviruses (Guo, 2020; Demogines, 2012; Mou, 2021), it is important to understand how

sarbecoviruses acquire the ability to bind to new ACE2 orthologues, including that of humans, through amino acid mutations.

2.2 Survey of ACE2 binding

To trace the evolutionary history of sarbecovirus binding to ACE2, we assembled a gene library encoding 45 sarbecovirus RBDs spanning all four known RBD phylogenetic clades (**Fig. 2.1a,b and Fig. S2.1**). We cloned the RBD library into a yeast-surface display platform that enables high-throughput measurement of ACE2-binding avidities using titration assays combining fluorescence-activated cell sorting (FACS) and deep sequencing (Starr, 2020) **(Fig. S2.2a-d).** We also assembled a panel of recombinant, dimeric ACE2 proteins from human, civet, pangolin and mouse, as well as two alleles each from *R. affinis* and *R. sinicus* bats (Guo, 2020) (**Fig. 2.1c**). The *R. affinis* alleles encode the two distinct RBD-interface sequences found among 23 *R. affinis* bats from Yunnan and Hubei, China. The *R. sinicus* alleles encode two out of the eight distinct RBD-interface sequences found among 25 *R. sinicus* bats from Yunnan, Hubei, Guangdong and Guangxi provinces, and Hong Kong (Guo, 2020), including one allele (3364) that is closest to consensus among the 8 RBD-interface sequences, and another (1434) that does not support entry by some clade 1a sarbecoviruses (Guo, 2020). We measured the apparent dissociation constant ($K_{D,app}$) of each RBD for each of the eight ACE2 orthologues (**Fig. 2.1b,d and Fig. S2.2**). We performed all of the experiments in duplicate using independently constructed libraries, and the measurements were highly correlated between replicates ($R^2 > 0.99$; **Fig. S2.2g**).

Consistent with a previous survey of human ACE2-mediated cellular infectivity(Letko, 2020), human ACE2 binding is restricted to RBDs within the SARS-CoV-1 and SARS-CoV-2 clades (**Fig. 2.1b**), although binding affinities vary among RBDs within these clades. Specifically, the RBDs from SARS-CoV-2 and related viruses from pangolins bind to human ACE2 with high affinity, whereas the RBD from the bat virus RaTG13 exhibits much lower affinity (Liu, 2021). The RBDs of SARS-CoV-1 isolates from the 2002–2003 epidemic bind to human ACE2 strongly, whereas RBDs from civet and sporadic 2004 human isolates (GD03T0013, GZ0402) show weaker binding, consistent with their civet origin and limited transmission (Li, 2005; Song, 2005). SARS-CoV-1-related bat virus RBDs bind to human ACE2, in some cases with higher affinity than SARS-CoV-1 itself.

Binding to civet ACE2 was detected only within the SARS-CoV-1 clade, whereas pangolin ACE2 binding is more widespread within the SARS-CoV-2 clade, consistent with viruses isolated from civet or pangolin partitioning specifically within each of these clades. Mice are not a natural host of sarbecoviruses, and RBDs from the SARS-CoV-1 and SARS-CoV-2 clades bind to mouse ACE2 only sporadically, typically with modest to weak affinity relative to other ACE2 orthologues. The highest binding affinity for mouse ACE2 is found in the cluster of RBDs related to RsSHC014, which can mediate infection and pathogenesis in mice (Menachery, 2015).

Binding to ACE2 of *R. affinis* and particularly *R. sinicus* bats varies considerably among strains in the SARS-CoV-1 and SARS-CoV-2 clades, consistent with an evolutionary arms race driving ACE2 variation in *Rhinolophus* bats (Guo, 2020; Demogines, 2012). The two *R. sinicus* bat ACE2 proteins tested interacted only with

SARS-CoV-1 isolates and the bat RsSHC014-cluster RBDs, which are notable for their broad ACE2-binding specificity in our assay. By contrast, we detected strong binding to both *R. affinis* ACE2 proteins among many RBDs in the SARS-CoV-1 and SARS-CoV-2 clades. However, the RBDs of the two viruses sampled from *R. affinis* in our panel bound only modestly (LYRa11) or very weakly (RaTG13) to the *R. affinis* ACE2s that we tested.

Strikingly, we detected binding to *R. affinis* ACE2 proteins by the RBD of the BtKY72 virus from Kenya (Tao, 2019) (**Fig. 2.1b,d),** the first described binding to any ACE2 orthologue for a sarbecovirus outside of Asia (Letko, 2020; Wells, 2021). To validate this finding, we purified the BtKY72 RBD and *R. affinis* ACE2–Fc fusion proteins recombinantly expressed in human cells and characterized their interaction using biolayer interferometry (BLI). In agreement with the yeast-display results, the BtKY72 RBD bound to the *R. affinis* 9479 ACE2 and more weakly to the *R. affinis* 787 allele (**Fig. 2.1e and Fig. S2.3a**). Furthermore, HEK293T cells transfected with the *R. affinis* 9479 or 787 *ACE2* alleles supported the entry of vesicular stomatitis virus (VSV) particles pseudotyped with the BtKY72 spike, thereby demonstrating that ACE2 is a bona fide entry receptor for this virus (**Fig. 2.1f and Fig. S2.3c,d**). The geographical range of *R. affinis* does not extend outside of Asia (Zhou, 2021), but this result indicates that BtKY72 may bind to ACE2 orthologues of bats found in Africa, although the full range of non-Asian bat species that harbour sarbecoviruses and their ACE2 sequences are underexplored (Tao, 2019; Drexler, 2010; Wells, 2021; Alkhovsky, 2022).

We did not detect ACE2 binding by any of the clade 2 RBDs. In our panel, 9 out of the 23 clade 2 RBDs were sampled from *R. sinicus*, in some cases from the same

caves and even found co-infecting the same *R. sinicus* bats (Hu, 2017) as ACE2-utilizing SARS-CoV-1-related RBDs. We tested binding by two clade 2 RBDs isolated from *R. sinicus* (YN2013 from Yunnan and HKU3-1 from Hong Kong Special Administrative Region) to an expanded ACE2 panel comprising all RBD-interface sequences observed in *R. sinicus* bats(Guo, 2020), including those sampled in Yunnan and Hong Kong Special Administrative Region. In contrast to SARS-CoV-1 Urbani and RsSHC014 (a clade 1a RBD isolated from *R. sinicus* in Yunnan (Ge, 2013),YN2013 and HKU3-1 RBDs did not bind to any of the eight *R. sinicus* ACE2 proteins (**Fig. S2.4**). Previous experiments with clade 2 RBDs have also demonstrated a lack of binding to *R. pearsoni* (Ren, 2008) and human(Letko, 2020; Hu, 2017; Starr, 2020; Ren, 2008) ACE2. Clade 2 RBDs have two large deletions within the receptor-binding motif (Letko, 2020; Hu, 2017; Wells, 2021), which has led to the hypothesis that this clade uses an unidentified alternative receptor, which could be bound by either the RBD or the spike N-terminal domain (Reguera, 2014; Tortorici, 2019; Tortorici, 2019; Bosch, 2019). Our results are consistent with this hypothesis, although we cannot rule out that clade 2 RBDs bind to other ACE2 orthologues that have not yet been tested.

2.3 Ancestral Origins of ACE2 Binding

Our finding that the BtKY72 RBD binds to ACE2 suggests that ACE2 binding was present in the ancestor of all sarbecoviruses before the split of Asian and non-Asian RBD clades (**Fig. 2.2a**). To test this hypothesis, we used ancestral sequence reconstruction (Thornton, 2004) to infer plausible sequences representing ancestral

nodes on the sarbecovirus RBD phylogeny (**Fig. 2.2a and Fig. S2.5a**). We evaluated ACE2 binding for the most probable reconstructed ancestral sequences (**Fig. 2.2b and Fig. S2.5b**) and in alternative reconstructions that incorporate statistical or phylogenetic ambiguities inherent to ancestral reconstruction (**Fig. S2.6**). Consistent with the distribution of ACE2 binding among extant sarbecoviruses, the reconstructed ancestor of all sarbecovirus RBDs (AncSarbecovirus) bound to the *R. affinis* 9479 ACE2 (**Fig. 2.2b**)**.** Broader ACE2 binding (including to human ACE2) was acquired on the branch connecting AncSarbecovirus to the ancestor of the three Asian sarbecoviruses RBD clades (AncAsia). ACE2 binding was then lost along the branch to the clade 2 ancestor (AncClade2), due to the combination of 48 amino-acid substitutions and 2 deletions within the ACE2-binding region that occurred along this branch **(Fig. 2.2c).**

This evolutionary history of ACE2 binding is robust to some but not all analyses of uncertainty in our phylogenetic reconstructions (Eick, 2017; Hanson-Smith, 2010). The key phenotypes represented in **Fig. 2.2b** are robust to uncertainties in the topology of the RBD phylogeny (**Fig. S2.6a,b**) or possible recombination within the RBD impacting the cluster of RBDs related to RsSHC014 (**Fig. S2.6c-f).** However, statistical uncertainty in the identity of some ACE2-contact positions affects our inferences, with some reasonably plausible 'second-best' reconstructed states altering ancestral phenotypes **(Fig. S2.6b).** Nonetheless, our hypothesis of an ancestral origin of sarbecovirus ACE2 binding is supported by the most plausible ancestral reconstructions as well as the distribution of ACE2 binding among the directly sampled sarbecovirus RBDs in **Fig. 2.1a,b.**

2.4 Evolvability of ACE2 Binding

To examine how easily RBDs can acquire ACE2 binding through single amino-acid mutations, we constructed mutant libraries in 14 RBD backgrounds spanning the RBD phylogeny. In each background, we introduced all single amino acid mutations at six RBD positions previously implicated in the evolution of receptor binding in SARS-CoV-2 and SARS-CoV-1 (Starr, 2020 ;Wan, 2020) (SARS-CoV-2 residues Leu455, Phe486, Gln493, Ser494, Gln498 and Asn501; **Fig. 2.3a**; we use SARS-CoV-2 numbering for mutations in all of the homologues below). We recovered nearly all 1,596 of the intended mutations, and measured the binding of each mutant RBD to each ACE2 orthologue using high-throughput titrations as described above.

The results show that ACE2 binding is a remarkably evolvable trait (**Fig. 2.3b,c and Fig. S2.7**). In almost all cases in which a parental RBD binds to a particular ACE2, there are single amino acid mutations that improve binding by greater than fivefold. Thus, ACE2 binding can easily be enhanced by mutation, which may facilitate the frequent host jumps seen among sarbecoviruses(Latinne, 2020). Notably, our data on mouse ACE2 binding could inform the development of mouse-adapted sarbecovirus strains for in vivo studies(Menachery, 2015; Dinnon, 2020; Roberts, 2007; Menachery, 2016), including potentially safer strains that bind to mouse but not human ACE2 (**Fig. S2.8**).

In the majority of cases in which an RBD does not bind to a particular ACE2 orthologue, single mutations can confer low to moderate binding affinity (**Fig. 2.3b,c**). The only exceptions are BM48-31 and AncClade2, for which none of the tested mutations enabled binding to any of the ACE2 variants. We found that the mutation K493Y in AncSarbecovirus enables binding to human ACE2 (**Fig. 2.3b and Fig. S2.7**),

although this particular mutation did not occur on the branch to AncAsia where we inferred that human ACE2 binding was historically acquired, illustrating the existence of multiple evolutionary paths to acquiring human ACE2 binding. We identified single mutations at positions 493, 498 and 501 that enable the BtKY72 RBD to bind to human ACE2 (**Fig. 2.3b and Fig. S2.7**), suggesting that human ACE2 binding is evolutionarily accessible in this lineage.

We validated that the mutations K493Y and T498W enable the RBD of the African sarbecovirus BtKY72 to interact with human ACE2 using purified recombinant proteins. Binding to human ACE2–Fc is not detectable with the parental BtKY72 RBD using BLI but is conferred by T498W and enhanced for the K493Y/T498W double mutant (**Fig. 2.3d and Fig. S2.3b**). To evaluate whether the observed binding translated into cell entry, we generated VSV particles pseudotyped with the wild-type or mutant BtKY72 spikes and tested entry in HEK293T cells expressing human ACE2. We detected robust spike-mediated entry for the K493Y/T498W double mutant but not the T498W single mutant (**Fig. 2.3e and Fig. S2.3c,e**), reflecting their apparent avidities (**Fig. 2.3d**) and confirming the evolvability of human ACE2 binding in this African sarbecovirus lineage.

Finally, we examined how the mutations that enhance ACE2 binding differ among sarbecovirus backgrounds, reflecting epistatic turnover in mutation effects (Starr, 2020; Starr, 2016). For example, the N501Y mutation increases human ACE2-binding affinity for SARS-CoV-2 where it has arisen in variants of concern(Tao, 2021), but the homologous mutation in the SARS-CoV-1 RBD (position 487) is highly deleterious for human ACE2 binding (**Fig. 2.3f**). More broadly, variation in mutant effects increases as

RBD sequences diverge (**Fig. 2.3g and Fig. S2.9**). However, the rate of this epistatic turnover varies across positions—for example, the effects on human ACE2 binding for mutations at positions 486 and 494 remain relatively constant across sequence backgrounds, whereas variability in the effects of mutations at positions 498 and 501 increases substantially as RBDs diverge.

2.5 New Sarbecovirus lineages bind ACE2

Given that ACE2 binding is an ancestral sarbecovirus trait with plastic evolutionary potential, unsampled sarbecoviruses lineages probably have the ability to bind to ACE2 and evolve to bind to human ACE2 unless these traits have been specifically lost as occurred in clade 2. To test this idea, we investigated sarbecoviruses reported after the initiation of our study, including viruses from Africa(Wells, 2021) and Europe (Alkhovsky, 2022; Crook, 2021) and a new RBD lineage represented by RsYN04 from a *Rhinolophus stheno* bat in Yunnan, China (Zhou, 2021), which branches separately from the four RBD clades previously described (**Fig. 2.4a**).

We determined the ACE2-binding abilities of these RBDs using our yeast-display platform. We found that two newly described sarbecoviruses from the Caucasus region of Russia (Alkhovsky, 2022) bind to ACE2 (**Fig. 2.4b**): the Khosta-1 RBD binds to *R. affinis* ACE2s with avidity that is improved by the T498W mutation and, strikingly, the Khosta-2 RBD binds to human ACE2 even in the absence of mutations. The Khosta-2 RBD was also recently shown to enable cell entry through human ACE2 (Seifert, 2021; Roelle, 2021). This finding indicates that the evolvability of human ACE2 binding that we describe for other African and European sarbecoviruses has been realized in naturally

circulating viruses that are geographically and phylogenetically separated from the southeast Asian clades from which spillover has been described to date. Our results also reinforce our observation of ACE2 binding in African sarbecoviruses (**Fig. 2.4c**)—similar to BtKY72, RBDs of the newly described African sarbecoviruses PDF-2380 and PRD-0038 (Wells, 2021) bind to *R. affinis*ACE2s, and the K493Y/T498W double mutant confers human ACE2 binding to the PRD-0038 RBD as it does for BtKY72. Finally, the uniquely branching RsYN04 RBD binds to *R. affinis* 787 ACE2 (**Fig. 2.4d**), as was recently shown for the closely related RaTG15 spike(Menachery, 2017). The RsYN04 RBD can also acquire binding to human ACE2 through the single T498W mutation. Incorporation of newly described sarbecovirus sequences into an updated phylogenetic reconstruction of the AncSarbecovirus RBD sequence reaffirms the conclusion that the ancestral sarbecovirus binds to bat ACE2 and can evolve human ACE2 binding through single amino-acid mutation (**Fig. S2.10**). These results illustrate that the ancestral traits of ACE2 binding and ability to evolve human ACE2 binding are maintained in geographically and phylogenetically diverse sarbecoviruses, including lineages that are just beginning to be described(Tao, 2019; Zhou, 2021; Well, 2021; Alkhovsky, 2022; Guo, 2021).

2.6 Chapter Discussion

Our experiments reveal that binding to bat ACE2 is an ancestral trait of sarbecoviruses that is also present in viruses from outside of Asia (Tao, 2019; Wells, 2021; Alkhovsky, 2020). Binding to human ACE2 arose in the common ancestor of SARS-CoV-1- and SARS-CoV-2-related RBDs before their divergence, and human ACE2 binding is evolvable in other phylogenetic clades. Binding to the ACE2

orthologues that we tested was then lost on the branch leading to the clade 2 RBDs, which either bind to an alternative receptor or ACE2 orthologues that were not evaluated here. These results imply that unsampled RBD lineages in the phylogenetic interval between BtKY72 and SARS-CoV-1/SARS-CoV-2 probably use ACE2 as an entry receptor and have the ability to evolve affinity for human ACE2. Indeed, the Khosta-2 virus from Russia provides an example of a RBD for which this evolutionary potential for human ACE2 binding has been realized.

Our research also shows that ACE2 binding is a highly evolvable trait of sarbecovirus RBDs. For every ACE2-binding RBD that we studied, there were single amino acid mutations that enhanced affinity for ACE2 orthologues that a RBD could already bind to or that conferred binding to new ACE2 orthologues from different species. Host jumps are common among the wide diversity of bats that are naturally infected with these viruses (Hu, 2017; Zhou, 2021; Latinne, 2020). In addition to frequent exchange of RBDs among viral backbones through recombination (Hu, 2017; Wells, 2021; Boni, 2020) the evolutionary plasticity of RBD binding to ACE2 is probably a key contributor to the ecological dynamics of sarbecoviruses, and perhaps other coronaviruses that frequently transmit across species (Menachery, 2017). As the effects of RBD mutations on ACE2 binding can differ across sarbecovirus backgrounds, it is not trivial to predict the ACE2-binding properties of a given RBD solely from its sequence. Thus, high-throughput approaches such as the one we have used here, which enables rapid and comprehensive measurement of ACE2-binding affinities of RBD variants in a non-viral context, can aid efforts to understand the evolutionary diversity and dynamics of sarbecoviruses and develop broadly protective therapeutics.

Sarbecoviruses are of particular concern, as two different strains have caused human outbreaks. Although human infectivity depends on many factors, the ability to bind to human receptors is certainly a key factor. Our results show that the ability of sarbecoviruses to bind to human ACE2 is evolvable and has arisen independently in regions outside of southeast Asia. Our high-throughput yeast-display platform enables the study of possible host tropism of sarbecoviruses without requiring work with replication-competent viruses that can pose biosafety concerns. The geographical breadth of ACE2 binding that we describe suggests that care should be taken in the sampling and study of replication-competent sarbecoviruses even outside regions such as southeast Asia in which spillover potential is considered greatest, and that efforts to develop vaccines and antibody therapeutics for pandemic preparedness should consider sarbecoviruses circulating worldwide.

2.7 Methods

Phylogenetics and ancestral sequence reconstruction

All steps of the bioinformatic analysis, including specific programmatic commands, alignments, raw data and output files are provided at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/tree/master/RBD_ASR). A panel of unique sarbecovirus RBD sequences was assembled incorporating the RBD sequences curated in ref. Lekto, 2020, all unique RBD sequences among SARS-CoV-1 human and civet strains reported in ref. Song, 2005, and recently reported sarbecoviruses BtKY72 (ref. Tao, 2019), RaTG13 (ref. Zhou, 2020) GD-Pangolin-CoV (consensus RBD sequence reported in **Fig. 2.3a** of ref. Lam, 2020) and GX-Pangolin-

CoV (Lam, 2020) (P2V, ambiguous nucleotide in codon 515 (SARS-CoV-2 numbering) was resolved to retain amino acid Phe515, which is conserved across all other sarbecoviruses). We also incorporated newly described sarbecovirus sequences RsYN04 (ref. Zhou, 2021), PDF-2370 and PRD-0038 (ref. Wells, 2021), Khosta-1 and Khosta-2 (ref. Alkovsky, 2020), RhGB01 (ref. Crook, 2021), RshSTT182 (ref. Delaune, 2021) and Rc-o319 (ref. Murakami, 2020) into updated phylogenies and functional work after the initiation of our study (**Fig. 2.4 and Fig. S2.10**). The Hibecovirus sequence Hp-BetaCoV/Zhejiang2013 (GenBank: KF636752) was used to root the sarbecovirus phylogeny. For Figs. S1 and S10a-d, additional betacoronavirus outgroups were included in rooting. All virus names, species and location of sampling, and sequence accessions or citations are provided at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/RBD_ASR/RBD_accessions.csv).

Amino acid sequences were aligned by mafft (v.7.471)(Katoh, 2013) with a gap opening penalty of 4.5. RBD sequences were subsetted from spike alignments according to our domain boundary defined for SARS-CoV-2 (Wuhan-Hu-1 GenBank: MN908947, residues Asn331–Thr531). Nucleotide alignments were constructed from amino acid alignments using PAL2NAL (v.14)(Suyama, 2006). Phylogenies were inferred with RAxML (v.8.2.12)(Stamatakis, 2014) using the LG+Γ substitution model for amino acid sequence alignments or GTR+Γ with separate data partitions applied to the first, second and third codon positions for nucleotide sequence alignments. Constraint files specifying specific clade relationships (but free topologies within clades) were used to fix particular topologies in **Fig. S2.6a** (alternative relationships between RBD clades 1a, 1b and 2) and **Fig. 2.4a** (monophyletic Europe

and Africa RBD clade; **Fig. S2.10a-d**). RBD gene segments were used as our primary boundary for phylogenetic inference and ancestral sequence reconstruction due to the presence of frequent recombination within broader spike alignments (Wells, 2021; Boni, 2020).

Marginal likelihood ancestral sequence reconstruction was performed with FastML (v.3.11)(Ashkenazy, 2012) using the amino acid sequence alignment, the maximum likelihood nucleotide tree topology from RAxML, the LG+Γ substitution matrix, re-optimization of branch lengths and FastML's likelihood-based indel reconstruction model. The maximum a posteriori ancestral sequences at nodes of interest were determined from the marginal reconstructions as the string of amino acids at each alignment site with the highest posterior probability, censored by deletions as inferred from the indel reconstruction. To test the robustness of ancestral phenotypes to statistical uncertainty in reconstructed ancestral states, we also constructed 'alt' ancestors in which all second-most-probable states with posterior probability > 0.2 were introduced simultaneously (Eick, 2017).

To identify potential recombination breakpoints within the RBD alignment, we used GARD (v.0.2)(Kosakovsky, 2006), which identified a possible recombination breakpoint (**Fig. S2.6c**) that produces two alignment segments exhibiting phylogenetic incongruence with a gain in overall likelihood sufficient to justify the duplication of phylogenetic parameters (ΔAIC = −85). To determine the impact of this possible recombination on ancestral sequence reconstructions, the alignment was split into separate segments at the proposed breakpoint. Phylogenies were inferred and ancestral sequences reconstructed on separate segments as described above, and

reconstructed ancestral sequences at matched nodes for each segment were concatenated, as shown in **Fig. S2.6e.**

RBD library construction

Genes encoding all 73 unique extant and ancestral RBD amino acid sequences were ordered from Twist Bioscience, Genscript, and IDT. Gene sequences are provided at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/RBD_ASR/parsed_sequences/RBD_sequence_set_annotated.csv). Genes were cloned in bulk into the pETcon yeast surface-display vector (plasmid 2649) as described previously (Starr, 2020). As described in this previous publication, randomized N16 barcodes were appended by PCR downstream from RBD coding sequences. RBD sequences were pooled and barcoded in two independently processed replicates. The pooled, barcoded parental RBD libraries were electroporated into *Escherichia coli* and plated at an estimated bottleneck of ~22,000 colony-forming units, yielding an estimated ~300 barcodes per parental RBD within each library replicate.

In parallel, we cloned site saturation mutagenesis libraries of six positions in select RBD backgrounds. The positions targeted correspond to SARS-CoV-2 positions 455, 486, 493, 494, 498 and 501. The RBD-indexed position targeted in each background is provided at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/RBD_ASR/parsed_sequences/RBD_sequence_set_annotated.csv). Precise site saturation mutagenesis pools were produced by Genscript,

provided as plasmid libraries. Failed positions in the Genscript mutagenesis libraries (all six positions in SARS-CoV-1 Urbani, position 494 in SARS-CoV-2, and position 455 in RaTG13 and GD-Pangolin) or backgrounds chosen for mutagenesis subsequent to initial library design (BtKY72) were produced in-house by PCR-based mutagenesis using NNS degenerate mutagenic primers followed by Gibson Assembly of the mutagenized fragments. In duplicate, mutant libraries were pooled and N16 barcodes were appended downstream from the RBD coding sequence. The pooled, barcoded mutant libraries were electroporated into *E. coli* and plated at a target bottleneck corresponding to an average of 20 barcodes per mutant within each library replicate.

Colonies from bottlenecked transformation plates were scraped and plasmids were purified. Parental RBD and mutant pools were combined at ratios corresponding to expected barcode diversity, yielding the two separately barcoded library replicates used in high-throughput experiments. Plasmid libraries were transformed into yeast (AWY101 strain(Wentz, 2007)) according to a previously described protocol(Gietz, 2007), transforming 10 µg of plasmid at 10× scale.

PacBio sequencing and analysis

As described previously(Starr, 2020), PacBio sequencing was used to acquire long sequence reads spanning the N16 barcode and RBD coding sequence. PacBio sequencing constructs were prepared from library plasmid pools by NotI digestion and gel purification, followed by SMRTbell ligation. Each library was sequenced across three SMRT Cells on a PacBio Sequel using 20 h video collection times. PacBio circular consensus sequences (CCSs) were generated from subreads using the ccs program

(v.5.0.0), requiring 99.9% accuracy and a minimum of 3 passes. The resulting CCSs are available on the NCBI Sequence Read Archive (SRA), BioSample SAMN18316101.

CCSs were processed using alignparse (v.0.1.6)(Crawford, 2019) to identify the RBD target sequence, call any mutations and determine the associated N16 barcode sequence, requiring no more than 18 nucleotide mutations from the intended target sequence, an expected 16-nucleotide-length barcode sequence and no more than 3 mismatches across the sequenced portions of the vector backbone.

We next used processed CCSs to link each barcode to the associated RBD sequence. We first filtered sequences with ccs-determined accuracies of <99.99% or indels. The empirical sequencing accuracy estimated by comparing RBD variants associated with barcode sequences sampled across multiple CCSs (https://jbloomlab.github.io/alignparse/alignparse.consensus.html#alignparse.consensus.empirical_accuracy) was 99.0% and 98.4% in libraries 1 and 2, respectively. For barcodes sampled across multiple CCSs, we derived consensus RBD variant sequences, discarding barcodes of which CCSs with identical barcodes exhibited >1 point mutation or >2 indels, or of which >10% or >25% of CCSs with an identical barcode contained a secondary non-consensus mutation or indel, respectively. The CCS processing pipeline is available at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/summary/process_ccs.md). The final barcode-variant lookup table, which links each N16 barcode with its associated RBD sequence, is available at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/variants/nucleotide_variant_table.csv).

ACE2 proteins for yeast-display assays

Recombinant dimeric ACE2 proteins for yeast-display binding assays were purchased or produced from commercial sources. Recombinant human ACE2 (UniProt: Q9BYF1-1) was purchased from ACROBiosystems (AC2-H82E6), consisting of residues 18–740 spanning an intrinsic dimerization domain, followed by a His tag and biotinylated Avitag used for downstream detection. Civet (*Paguma larvata*) ACE2 (UniProt: Q56NL1-1) was purchased from ACROBiosystems (AC2-P5248), consisting of residues 18–740 spanning an intrinsic dimerization domain, with an N-terminal His tag used for downstream detection. Mouse (*Mus musculus*) ACE2 (UniProt: Q8R0I0-1) was purchased from Sino Biological (50249-M03H), consisting of residues 18–740 spanning an intrinsic dimerization domain, followed by a His tag and human IgG1 Fc domain used for downstream detection.

The remaining ACE2s for yeast-display binding assays (with the exception of **Fig. S4**) were produced by Genscript. Specifically, pangolin (*Manis javanica*, GenBank: XP_017505746.1), *R. affinis* 787 (GenBank: QMQ39222), *R. affinis* 9479 (GenBank: QMQ39227), *R. sinicus* 3364 (GenBank: QMQ39219) and *R. sinicus* 1434 (GenBank: QMQ39216) ACE2 residues 19–615 were cloned with a C-terminal human IgG1 Fc domain for dimerization and downstream detection. pcDNA3.4 expression plasmids were transfected into HD 293F cells for protein expression. ACE2–Fc fusion

Library measurements of RBD expression and RBD$^+$ enrichment

Transformed yeast library aliquots were grown overnight in a shaker at 30 °C in SD-CAA medium (6.7 g l^{-1} yeast nitrogen base, 5.0 g l^{-1} casamino acids, 2.13 g l^{-1} MES and 2% (w/v) dextrose, pH 5.3). To induce RBD expression, yeast was washed and resuspended in SG-CAA + 0.1% D medium (6.7 g l^{-1} yeast nitrogen base, 5.0 g l^{-1} casamino acids, 2.13 g l^{-1} MES, 2% (w/v) galactose and 0.1% (w/v) dextrose, pH 5.3) at an initial optical density at 600 nm (OD_{600}) of 0.67, and incubated at room temperature for 16–18 h with mild agitation.

For each library, 45 OD_{600} of induced culture was washed twice with PBS-BSA (0.2 mg ml^{-1}), and RBD surface expression was labelled by a C-terminal c-Myc tag with 1:100 diluted FITC-conjugated chicken anti-c-Myc antibodies (Immunology Consultants Lab, CMYC-45F) in 3 ml PBS-BSA. Labelled cells were washed twice in PBS-BSA, and resuspended in PBS for FACS analysis.

Yeast library sorting experiments were conducted on the BD FACSAria II system with FACSDiva software (v.8.0.2). For high-throughput measurements of RBD expression levels, cells were gated for single cells (**Fig. S2b**) and partitioned into 4 bins of FITC fluorescence (**Fig. S2.2c**), where bin 1 captures 99% of unstained cells, and bins 2–4 split the remaining library population into tertiles. Cells were sorted into 5 ml tubes pre-wet with 1 ml of SD-CAA with 1% BSA. We recovered ~8 million cells per library across the 4 bins. Sorted cells were resuspended to 2×10^6 cells per ml in fresh SD-CAA with 1:100 penicillin–streptomycin, and grown overnight at 30 °C. Plasmids were purified from post-sort yeast samples of $<4 \times 10^7$ cells per miniprep column using

the Zymo Yeast Miniprep II kit (D2004) according to the manufacturer's instructions, with the addition of an extended (>2 h) Zymolyase treatment and a −80 °C freeze–thaw cycle before cell lysis. N16 barcodes were PCR amplified from each plasmid aliquot as described previously(Starr, 2020) and submitted for Illumina HiSeq 50 bp single-end sequencing.

To enrich properly expressing RBD variants for downstream titration experiments, we also sorted around 2×10^7 cells per library using the RBD^+ ($FITC^+$) bin (**Fig. S2.2b**). RBD^+-enriched populations were resuspended to 1×10^6 cells per ml for overnight outgrowth, and frozen at −80 °C in 9 OD_{600} aliquots for subsequent titration experiments. A pool of mutants that were added after the first set of experiments (mutations at position 455 in RaTG13 and GD-Pangolin, and mutations at all six positions in BtKY72) were not RBD^+enriched and were not part of the bulk expression Sort-seq measurement, but were pooled with the RBD^+-enriched population of the primary libraries for subsequent titration assays.

Library measurements of ACE2-binding affinities

For high-throughput measurements of ACE2-binding affinities, yeast libraries were induced for RBD expression as described above. Induced cultures were aliquoted at 8 OD_{600} per titration sample and washed twice with PBS-BSA. Cells were resuspended across a range of ACE2 concentrations from 1×10^{-6} M to 1×10^{-13} M in 1 M intervals, plus a 0 M ACE2 concentration. The samples were incubated overnight at room temperature with mild agitation. The samples were washed twice in ice-cold PBS-BSA, and resuspended in 1 ml secondary label (1:100 Myc-FITC and 1:200 PE-

conjugated streptavidin (Thermo Fisher Scientific, S866) for human ACE2, 1:200 iFluor647-conjugated mouse anti-His (Genscript, A01802) for civet ACE2 and 1:200 PE-conjugated goat anti-human IgG (Jackson ImmunoResearch Labs 109-115-098) for all other Fc-tagged ACE2 ligands), and incubated for 1 h on ice. Cells were washed twice with PBS-BSA and resuspended in PBS for FACS analysis.

Titration samples were binned for single RBD-expressing cells (**Fig. S2.2b**), which were then partitioned into four bins on the basis of ACE2 binding (**Fig. S2.2d**). At each concentration, a minimum of 5×10^6 cells were collected across the 4 bins. Sorted cells were resuspended in 1 ml SD-CAA with 1:100 penicillin–streptomycin, and grown overnight at 30 °C in deep-well plates. Plasmid aliquots from each population were purified using the Zymo Yeast 96-Well Miniprep kit (D2005) according to the manufacturer's instructions, with the addition of an extended (>2 h) Zymolyase treatment and a −80 °C freeze–thaw cycle before cell lysis. N16 barcodes were PCR amplified from each plasmid aliquot as described previously(Starr, 2020) and submitted for Illumina HiSeq 50 bp single-end sequencing.

For the pool of mutants that were added after the first set of experiments (mutations at position 455 in RaTG13 and GD-Pangolin, and mutations at all six positions in BtKY72), duplicate titrations were already conducted with the primary pool for human ACE2 and *R. affinis* 787 ACE2. Titrations with this smaller library sub-pool with these ACE2 ligands were conducted as described above, but scaled to 1.6 OD_{600} per sample, collecting >1 million cells per concentration.

Illumina barcode sequencing analysis

Demultiplexed sequence reads (available on the NCBI SRA, BioSample SAMN20174027) were aligned to library barcodes as determined from PacBio sequencing using dms_variants (v.0.8.5), yielding a count of the number of times each barcode was sequenced within each FACS bin. Read counts within each FACS bin were downweighted by the ratio of total reads from a bin compared to the number of cells that were actually sorted into that bin. The table giving downweighted counts of each barcode in each FACS bin is available at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/counts/variant_counts.csv).

We estimated the RBD expression level of each barcoded variant on the basis of its distribution of counts across FACS bins and the known log-transformed fluorescence boundaries of each sort bin using a maximum likelihood approach (Starr, 2020, Peterman, 2016), implemented with the fitdistrplus package (v.1.0.14)(Delignette, 2015) in R. Expression measurements were retained for barcodes for which greater than 20 counts were observed across the four FACS bins. The full pipeline for computing per-barcode expression values is described at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/summary/compute_expression_meanF.md).

We estimated the level of ACE2 binding of each barcoded variant at each titration concentration on the basis of its distribution of counts across FACS bins calculated as a simple mean(Peterman, 2016), as described previously (Starr, 2020). We determined the apparent binding constant $K_{D,app}$ describing the affinity of each barcoded variant for each ACE2 along with free parameters *a* (titration response range)

and *b* (titration curve baseline) with nonlinear least-squares regression using the standard non-cooperative Hill equation relating the mean bin response variable to the ACE2 labelling concentration:

The measured mean bin value at a given ACE2 concentration was excluded from a variant's curve fit if fewer than 10 counts were observed across the four FACS bins at that concentration. Individual concentration points were also excluded from the curve fit if they demonstrated evidence of bimodality (>40% of counts of a barcode were found in each of two non-consecutive bins 1 + 3 or 2 + 4, or >20% of counts of a barcode were found in each of the boundary bins 1 + 4). To avoid errant fits, we constrained the fit baseline parameter *b* to be between 1 and 1.5, the response parameter *a* to be between 2 and 3, and the $K_{D,app}$ parameter to be between 1×10^{-15} and 1×10^{-5}. The fit for a barcoded variant was discarded if the average count across all sample concentrations was below 10, or if >20% of sample concentrations were missing due to counts below 10. We also discarded curve fits in cases in which the normalized mean square residual (residuals normalized from 0 to 1 relative to the fit response parameter *a*) is >10× the median normalized mean square residual across all titrations with all ACE2s. $K_{D,app}$ binding constants were expressed as $-\log_{10}(K_{D,app})$, where higher values indicate higher-affinity binding. The full pipeline for computing per-barcode binding affinities is described at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/summary/compute_binding_Kd.md).

To derive our final measurements we collapsed measurements across internally replicated barcodes representing each RBD genotype. For each RBD genotype, we discarded the top and bottom 5% (expression measurements) or 2.5% (titration

affinities) of per-barcode measurements, and computed the mean value across the remaining barcodes within each library. The correlations in these barcode-averaged measurements between independently barcoded and assayed library replicates are shown in **Fig. S2.2g**. Final measurements were determined as the mean of the barcode-collapsed mean measurements from each replicate. The total number of barcodes collapsed into these final measurements from both replicates is shown in the histograms in **Fig. S2.2f**. Final measurements for an RBD genotype were discarded if the RBD genotype was not sampled with at least one non-filtered barcode in each replicate, or sampled with at least five non-filtered barcodes in a single replicate. The full pipeline for barcode collapsing is described at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/summary/barcode_to_genotype_phenotypes.md). The final processed measurements of expression and ACE2 binding for parental and mutant RBDs can be found at GitHub (https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/final_variant_scores/wt_variant_scores.csv and https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/master/results/final_variant_scores/mut_variant_scores.csv.

Isogenic ACE2-binding assays

For RBDs assayed subsequent to library experiments (**Fig. 2.4 and Figs. S2.4, S6f, and S10e**), RBDs were cloned as isogenic stocks into the 2649 plasmid, sequence verified and transformed individually into yeast using the LiAc/ssDNA transformation method (Gietz, 2007). Cultures were induced for RBD expression and labelled across ACE2 concentration series as described above in V-bottom 96-well plates with 0.067

OD_{600} yeast per well. ACE2 labelling of RBD^+ cells was measured using the BD LSRFortessa X50 flow cytometer and data were processed using FlowJo (v.10). Binding curves of PE (ACE2) mean fluorescence intensity versus ACE2 labelling concentration were fit as above, with the inclusion of a Hill coefficient slope parameter *n*.

Transient expression of R. affinis and R. sinicus ACE2–Fc

The *R. affinis* 787 (GenBank: QMQ39222.1), *R. affinis* 9479 (GenBank: QMQ39227.1), *R. sinicus*1446 (GenBank: QMQ39213.1), *R. sinicus* WJ1 (GenBank: QMQ39206.1), *R. sinicus* GQ262791 (GenBank: ACT66275.1), *R. sinicus* 3364 (GenBank: QMQ39219.1), *R. sinicus* WJ4 (GenBank: QMQ39200.1), *R. sinicus* 1438 (GenBank: QMQ39203.1), *R. sinicus* 1434 (GenBank: QMQ39216.1) and *R. sinicus* 3358 (GenBank: QMQ39212.1) ACE2 ectodomains constructs were synthesized by GenScript and placed into a pCMV plasmid. The domain boundaries for the ectodomain are residues 19–615. The native signal tag was identified using SignalP-5.0 (residues 1–18) and replaced with an N-terminal mu-phosphatase signal peptide. These constructs were then fused to a sequence encoding a thrombin cleavage site and a human Fc fragment at the C-terminus. All ACE2–Fc constructs were produced in Expi293F cells (Thermo Fisher Scientific, A14527) in Gibco Expi293 Expression Medium at 37 °C in a humidified 8% CO_2incubator rotating at 130 rpm. The cultures were transfected using PEI-25K (Polyscience) with cells grown to a density of 3 million cells per ml and cultivated for 4–5 days. Proteins were purified from clarified supernatants using a 1 ml HiTrap Protein A HP affinity column (Cytiva), concentrated and flash-frozen in 1× PBS, pH 7.4 (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl,

137 mM NaCl). Cell lines were not authenticated or tested for mycoplasma contamination.

Transient expression of BtKY72 parental and mutant RBDs

BtKY72 RBD construct (BtKY72 S residues 318–520) was synthesized by GenScript into a CMVR plasmid with an N-terminal mu-phosphatase signal peptide and a C-terminal hexa-histidine tag (-HHHHHHHH) joined by a short linker (-GGSS) to an Avi tag (-GLNDIFEAQKIEWHE). BtKY72 mutant constructs T498W (BtKY72 S residue 487) and K493Y/T498W (BtKY72 S residue 482/487) were subcloned by GenScript from the BtKY72 RBD construct. BtKY72 and BtKY72 mutant RBD constructs were produced in Expi293F cells in Gibco Expi293 Expression Medium at 37 °C in a humidified 8% CO_2 incubator rotating at 130 rpm. The cultures were transfected using PEI-25K with cells grown to a density of 3 million cells per ml and cultivated for 3–5 days. Proteins were purified from clarified supernatants using a 1 ml HisTrap HP affinity column (Cytiva), concentrated and then biotinylated using a commercial BirA kit (Avidity). Proteins were then purified from the BirA enzyme by affinity purification using a 1 ml HisTrap HP affinity column (Cytiva), concentrated and flash-frozen in 1× PBS, pH 7.4. Cell lines were not authenticated or tested for mycoplasma contamination.

BLI analysis

Assays were performed on an Octet Red (Forte Bio) instrument at 30 °C with shaking at 1,000 rpm. Streptavidin biosensors were hydrated in water for 10 min before incubation for 60 s in 10× kinetics buffer (undiluted). Biotinylated RBDs were loaded at 5–10 μg ml^{-1} in 10× kinetics buffer for 100–600 s before baseline equilibration for 120 s

in 10× kinetics buffer. Association of ACE2–Fc (dimeric) was performed at 1 μM in 10× kinetics buffer. These data were baseline-subtracted. The experiments were performed with three separate purification batches of BtKY72 RBDs. All RBDs were immobilized to identical levels, that is, 1 nm shift. The data were plotted in GraphPad Prism and a representative plot is shown.

Generation of VSV pseudovirus

The BtKY72 S construct was synthesized by GenScript and cloned into an HDM plasmid with a C-terminal 3× Flag tag. The BtKY72 mutant S constructs T498W (BtKY72 S residue 487) and K493Y/T498W (BtKY72 S residue 482/487) were subcloned by GenScript from the BtKY72 S construct. Pseudotyped VSV particles were prepared using HEK293T (ATCC CRL-11268) cells seeded into 10 cm dishes. HEK293T cells were transfected using Lipofectamine 2000 (Life Technologies) with a *S*-encoding plasmid in Opti-MEM transfection medium and incubated for 5 h at 37 °C with 8% CO_2 supplemented with DMEM containing 10% FBS. One day after transfection, cells were infected with VSV (G*ΔG-luciferase) and, after 2 h, infected cells were washed five times with DMEM before adding medium supplemented with anti-VSV G antibodies (I1-mouse hybridoma supernatant diluted 1:40, ATCC CRL-2700). Pseudotyped particles were collected 18–24 h after inoculation, clarified from cellular debris by centrifugation at 3,000*g* for 10 min, concentrated 100× using a 100 MWCO membrane for 10 min at 3,000 rpm and frozen at −80 °C. Mock pseudotyped VSV pseudovirus was generated as above but in the absence of *S*. Cell lines were not authenticated or tested for mycoplasma contamination.

VSV pseudovirus entry assays

HEK293T cells (ATCC CRL-11268) and HEK293T cells with stable transfection of human ACE2 (ref. Crawford, 2020) were cultured in 10% FBS, 1% penicillin–streptomycin DMEM at 37 °C in a humidified 8% CO_2 incubator. Cells were plated into poly-lysine-coated 96-well plates. For *R. affinis* ACE2 entry, transient transfection of *R. affinis* ACE2 in HEK293T cells was performed 36–48 h before infection using Lipofectamine 2000 (Life Technologies) and an HDM plasmid containing full length *R. affinis* ACE2 (synthesized by GenScript) in Opti-MEM. After 5 h incubation at 37 °C in a humidified 8% CO_2 incubator, DMEM with 10% FBS was added and cells were incubated at 37 °C in a humidified 8% CO_2 incubator for 36–48 h. Cell lines were not authenticated or tested for mycoplasma contamination.

Immediately before infection, HEK293T cells with stable expression of human *ACE2*, transient expression of *R. affinis ACE2* or not transduced to express *ACE2* were washed once with DMEM, then plated with normalized pseudovirus in DMEM. Infection in DMEM was performed with cells between 60–80% confluence (human ACE2-293T) or between 80–90% confluence (*R. affinis* ACE2-293T) for 2.5 h before adding FBS and penicillin–streptomycin to final concentrations of 10% and 1%, respectively. After 24 h of infection, One-Glo-EX (Promega) was added to the cells and incubated in the dark for 5 min before reading on a Synergy H1 Hybrid Multi-Mode plate reader (Biotek). Normalized cell entry levels of pseudovirus generated on different days

(biological replicates) were plotted in GraphPad Prism as individual points, and average cell entry across biological replicates was calculated as the geometric mean.

BtKY72 S parental and mutant pseudoviral particle inputs for the above cell entry assays were normalized to spike incorporation quantified using western blotting. Detection of S was performed using mouse monoclonal anti-Flag M2 antibodies (Sigma-Aldrich, F3165) and Alexa Fluor 680 AffiniPure Goat Anti-Mouse IgG, light chain specific (Jackson ImmunoResearch Labs, 115-625-174). Detection of the VSV backbone was performed using anti-VSV-M [23H12] antibodies (Kerafast, EB0011) and Alexa Fluor 680 AffiniPure Goat Anti-Mouse IgG, light chain specific (Jackson ImmunoResearch Labs 115-625-174). A representative blot is shown in **Fig. S3c**. Expression of the *R. affinis ACE2* alleles was not quantified or normalized.

Biosafety considerations

We characterized the human ACE2 binding of sarbecovirus RBDs and identified point mutants that increase the affinity of some RBDs. This work includes identifying sarbecovirus RBDs from outside southeast Asia that can naturally bind to human ACE2 (Khosta-2 RBD from Russia) or adapt to bind to human ACE2 with just a few mutations (BtKY72 RBD from Kenya). We verified this latter finding using non-replicative spike-pseudotyped VSV particles. None of our experiments pose a biosafety risk, as they involve only RBD protein (purified or expressed in yeast) or non-replicative pseudotyped VSV viral particles, and not live virus. However, it is possible that another researcher could perform experiments on actual sarbecoviruses with RBDs such as the ones we described, and such experiments could pose a risk. Against that possible information

misuse, we weigh the following benefits of the information conveyed by our study: (1) as stated in the concluding paragraph of the Discussion, we used safe methods to highlight the need for care when sampling sarbecoviruses including those from outside southeast Asia; (2) we identified a broader swath of spike proteins that should be included in biochemical studies to engineer countermeasures (such as broad antibodies (Starr, 2021; Tortorici, 2021b) or stabilized spike immunogens); (3) we characterized mutations that could enable safer mouse-adapted laboratory strains with reduced human ACE2 affinity (**Fig. S2.8c**); (4) we provide data that can improve sequence-based phenotypic predictions. We emphasize that our research indicates that live-virus experiments with any new sarbecovirus should involve careful consideration of risks, as human ACE2 binding may be widespread. The actual ability of a sarbecovirus to infect humans will depend not only on its ACE2 affinity, but also other properties including proteolytic activation of the spike protein (Johnson, 2021), innate immunity and other poorly understood factors.

2.8 Figures & Tables

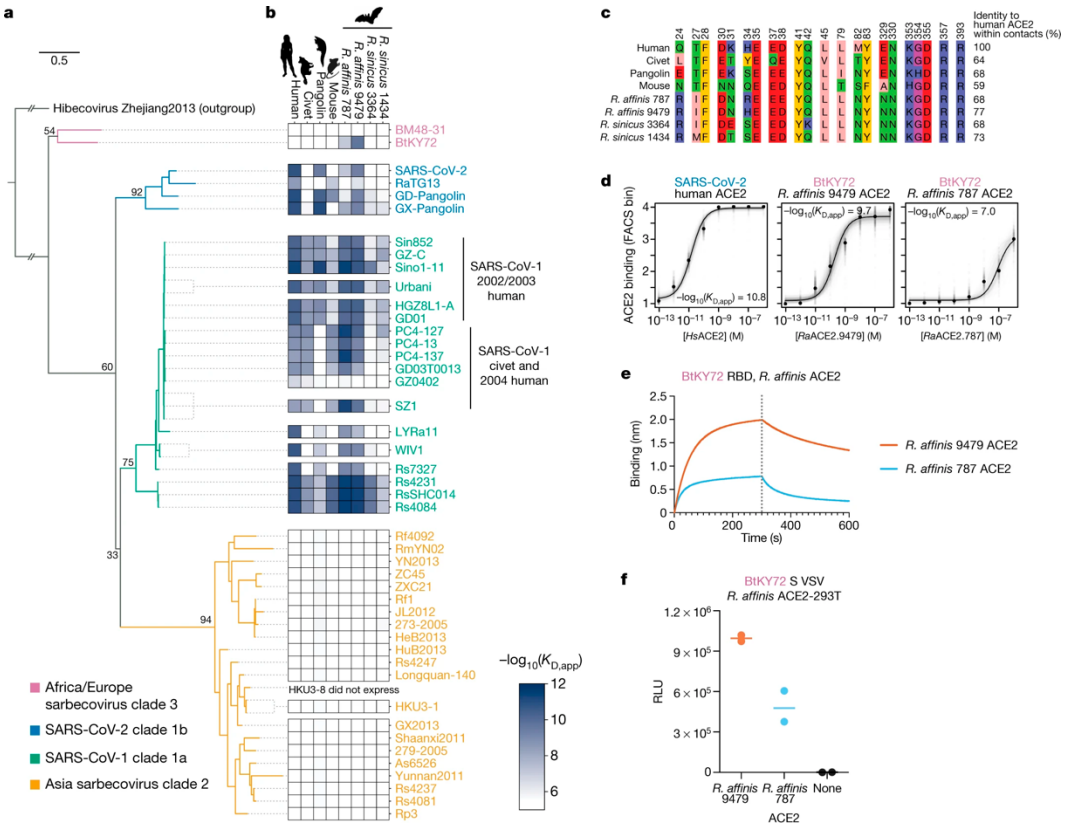


Figure 2.1: High-throughput survey of sarbecovirus ACE2 binding.

(**a)** Maximum likelihood phylogeny of sarbecovirus RBDs constructed from RBD nucleotide sequences. The node labels indicate bootstrap support values. Details on rooting are shown in **Fig. S2.1**. Scale bar, 0.5 nucleotide substitutions per site. (**b**) Binding avidities of sarbecovirus RBDs for eight ACE2 orthologues determined using high-throughput yeast-displayed RBD titration assays (**Fig. S2.2**). (**c**) Alignment of tested ACE2 orthologues within RBD-contact positions (4 Å cut-off in Protein Data Bank (PDB) 6M0J or 2AJF). (**d**) Representative ACE2-binding curves from high-throughput titrations. Underlying titration curves for individual replicate-barcoded representatives of a genotype are shown in light grey, and the average binding across all barcodes is indicated in black. (**e**) BLI binding analysis of 1 μM *R. affinis* ACE2–Fc binding to biotinylated BtKY72 RBD immobilized at the surface of streptavidin biosensors (see **Fig. 2.7a** for analysis of the robustness of the result to ACE2–Fc concentration). Data are representative of three assays using independent preparations of RBD (biological triplicate). (**f**) Entry of VSV particles pseudotyped with the BtKY72 spike into HEK293T

cells transiently expressing *R. affinis ACE2* alleles 9479 or 787. Each point represents the mean of technical triplicates for assays performed with independent preparation of pseudoviral particles (biological duplicate). The geometric mean is shown by the horizontal line. The normalized pseudovirus western blot, and mock (VSV prepared without spike plasmid) pseudovirus entry in *R. affinis* ACE2 HEK293T cells are shown in **Fig. S2.3c, d**.

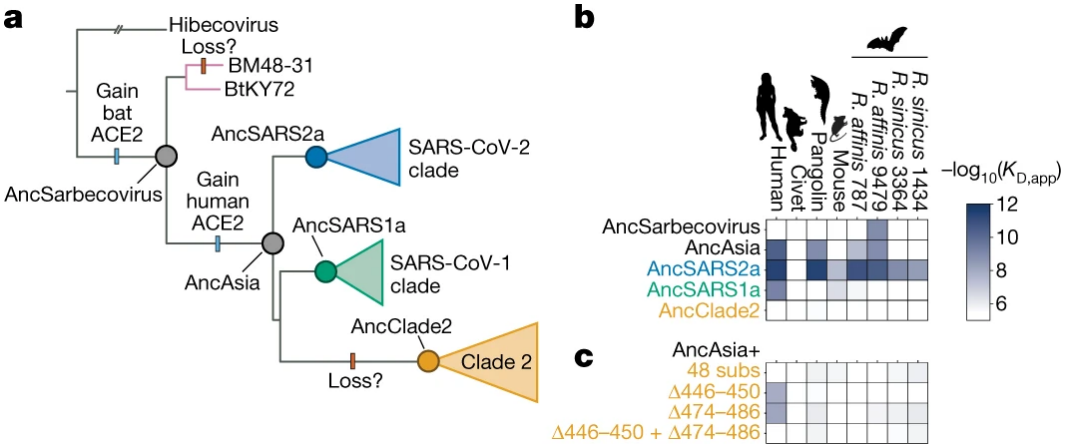


Figure 2.2: Ancestral origins of sarbecovirus ACE2 binding.

(**a**) Clade-collapsed RBD phylogeny. The circles represent nodes at which ancestral sequences were inferred. The bars indicate putative gains and losses in ACE2 binding. (**b**) ACE2 binding of ancestrally reconstructed, yeast-displayed RBDs (**Figs. S2.5 and S2.6**). (**c**) ACE2 binding of AncAsia RBD plus introduction of the 48 substitutions or 2 sequence deletions that occurred on the phylogenetic branch leading to AncClade2 RBD.

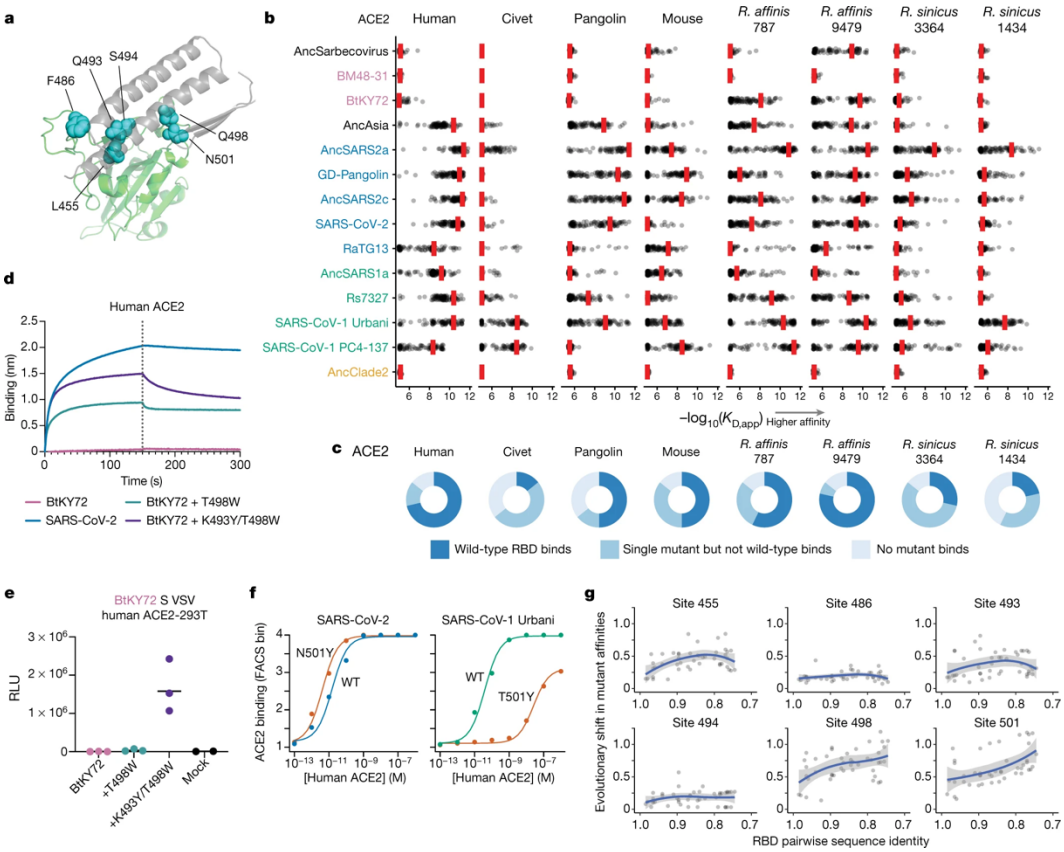


Figure 2.3: Evolutionary plasticity of ACE2 binding.

(**a**) The structural context of positions targeted for mutagenesis. Green cartoon, RBD; grey cartoon, ACE2 interaction motifs; blue spheres, residues targeted through mutagenesis (SARS-CoV-2 identities). (**b**) Mutational scanning measurements. The red bars mark the binding avidity of the parental RBD, and the points mark mutant avidities (see **Fig. S2.7** for mutation-level measurements). (**c**) The fraction of the 14 RBD backgrounds for which the parental RBD binds to the indicated ACE2 orthologue ($-\log_{10}(K_{D,app}) > 7$), a single mutant binds but the parental RBD does not, or no tested mutants bind. (**d**) Binding of 1 μM human ACE2–Fc to biotinylated RBDs immobilized at the surface of streptavidin biosensors (see **Fig. S2.3b** for an analysis of the robustness of the result to ACE2–Fc concentration). Data are representative of three assays using independent preparations of RBD (biological triplicate). (**e**) Entry of BtKY72 spike-pseudotyped VSV in HEK293T cells stably expressing human ACE2. Each point represents the mean of technical triplicates in assays performed with independent preparation of pseudoviral particles (biological triplicate). The horizontal line shows the geometric mean. Mock, VSV particles produced in cells in which no spike gene was transfected. A western blot of pseudotyped particles is shown in **Fig. S2.3c**, and entry into HEK293T cells lacking ACE2 is shown in **Fig. 2.3e**. (**f**) Titration curves illustrating the effect of mutation to tyrosine 501 (SARS-CoV-2 numbering) in the SARS-CoV-2 and SARS-CoV-1 Urbani RBD backgrounds. (**g**) Epistatic turnover in mutation effects. Each point represents, for a pair of RBDs, the mean absolute error (residual) in their correlated mutant avidities for human ACE2 (**Fig. S2.9a**) versus their pairwise amino acid sequence identity. Correlations were computed only for pairs in which the parental

RBDs bind with $-\log_{10}(K_{D,app}) > 7$. Data are LOESS mean (blue line) ± 95% confidence intervals trendline (grey shading) (see **Fig. S2.9b** for an analysis across all ACE2 orthologues).

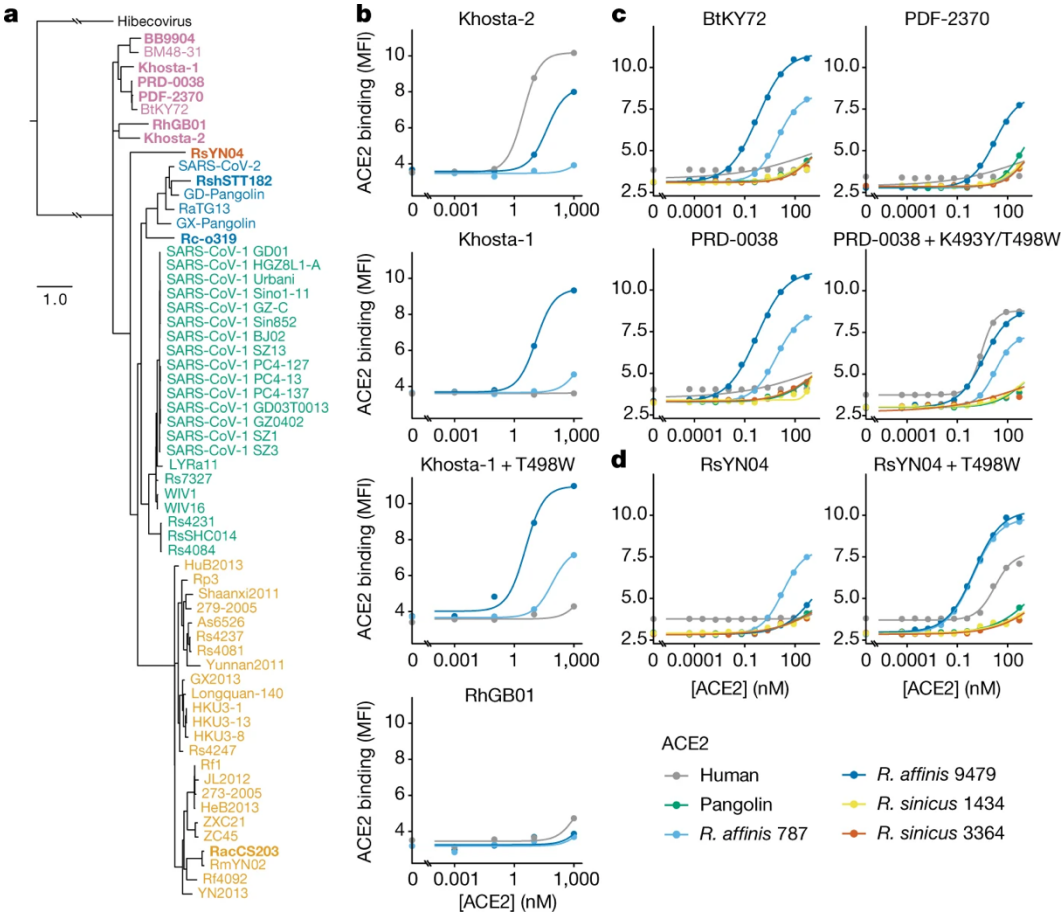
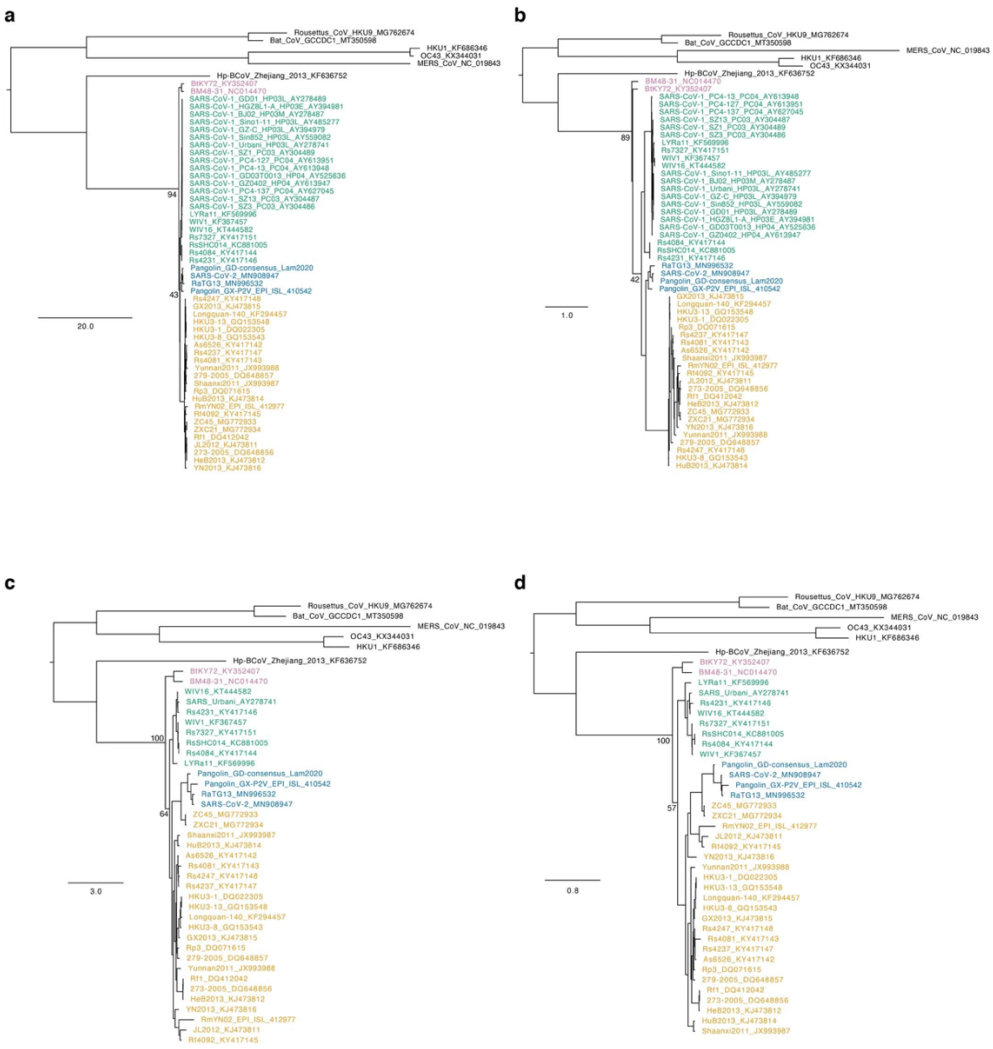


Figure 2.4: Newly sampled sarbecovirus lineages bind to ACE2.

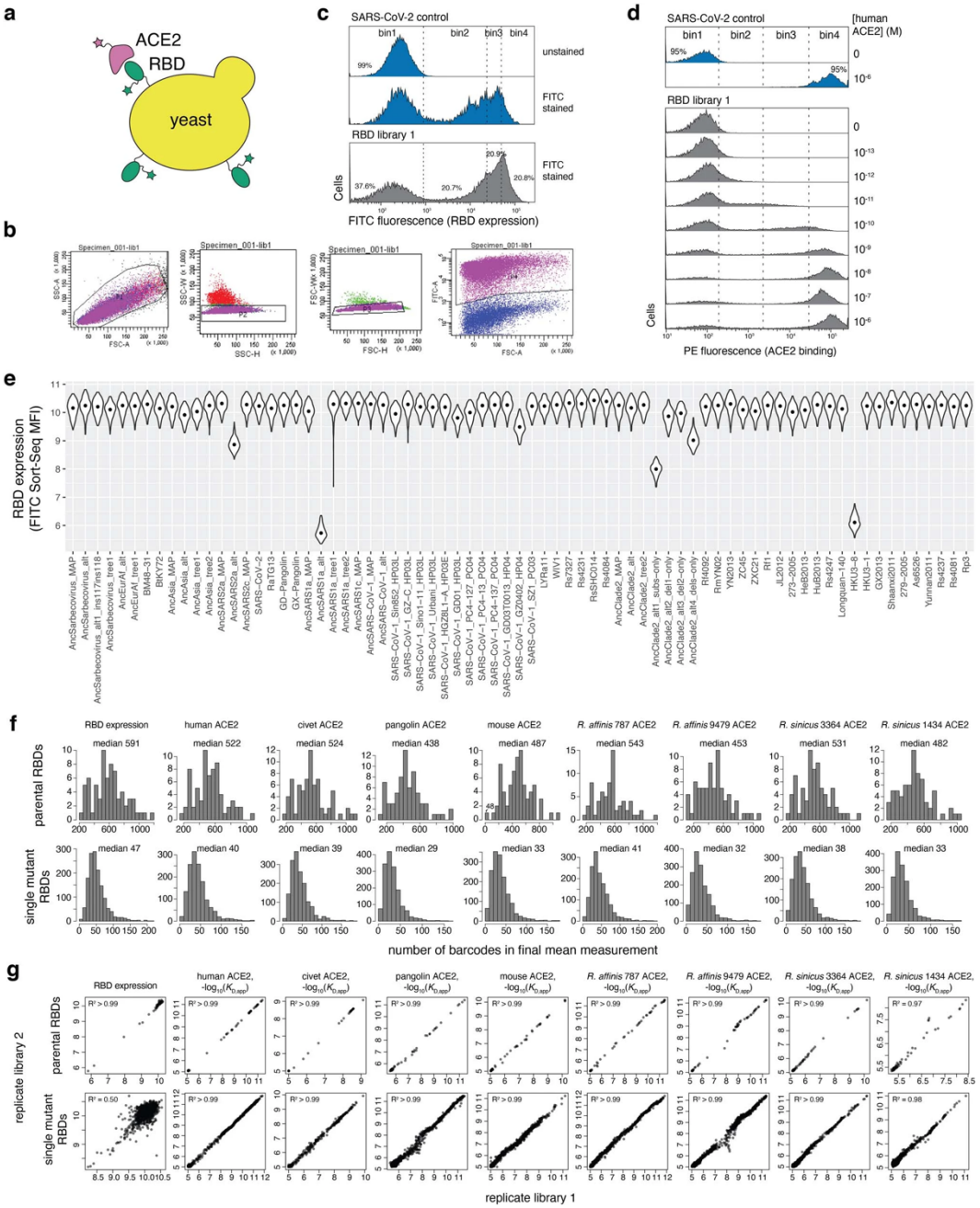
(**a**) Phylogenetic placement of the newly described sarbecovirus RBDs. The new sequences are shown in bold font. RBDs are coloured according to the key in **Fig. 2.1a** (**Fig. 2.14**). Scale bar, expected nucleotide substitutions per site. (**b–d**), Binding curves for newly described sarbecovirus RBDs from Europe (**b**), Africa (**c**) and Asia (**d**), and candidate mutations that confer human ACE2 binding. Measurements were performed with yeast-displayed RBDs and purified dimeric ACE2 proteins, measured using flow cytometry. Data are from a single experimental replicate.



Supplemental Figure 2.1: : Robustness of the root of the sarbecovirus ingroup.

To establish robustness of our conclusion that the first sarbecovirus divergence is between sarbecoviruses from Africa and Europe and those from Asia, we inferred phylogenies based on alignments of RBD (SARS-CoV-2 spike residues N331-T531) (**a**,**b**) or the full spike gene (**c**,**d**) and nucleotide (**a**, **c**) or amino-acid (**b**, **d**) alignments and substitution models. In all four cases, the first sarbecovirus bipartition is placed between sarbecoviruses in Africa/Europe and those in Asia. The placement of the overall tree root is arbitrary with respect to the relationship among non-sarbecovirus outgroups, but this arbitrary placement does not impact the sarbecovirus ingroup rooting. The primary variations among trees includes a potential paraphyletic separation of BtKY72 and BM48–31 from Europe and Africa such that they do not form a monophyletic clade (**b**; also seen in **Fig. 2.14a-c**), and variation in the relationships among the three Asia sarbecovirus clades (whose relationship is also inferred with a very low bootstrap support value in our primary phylogeny in **Fig. 2.1a**). Known recombination of RBDs with respect to other spike segments among viruses creates incongruencies between spike and RBD trees among Asian sarbecovirus lineages (e.g.

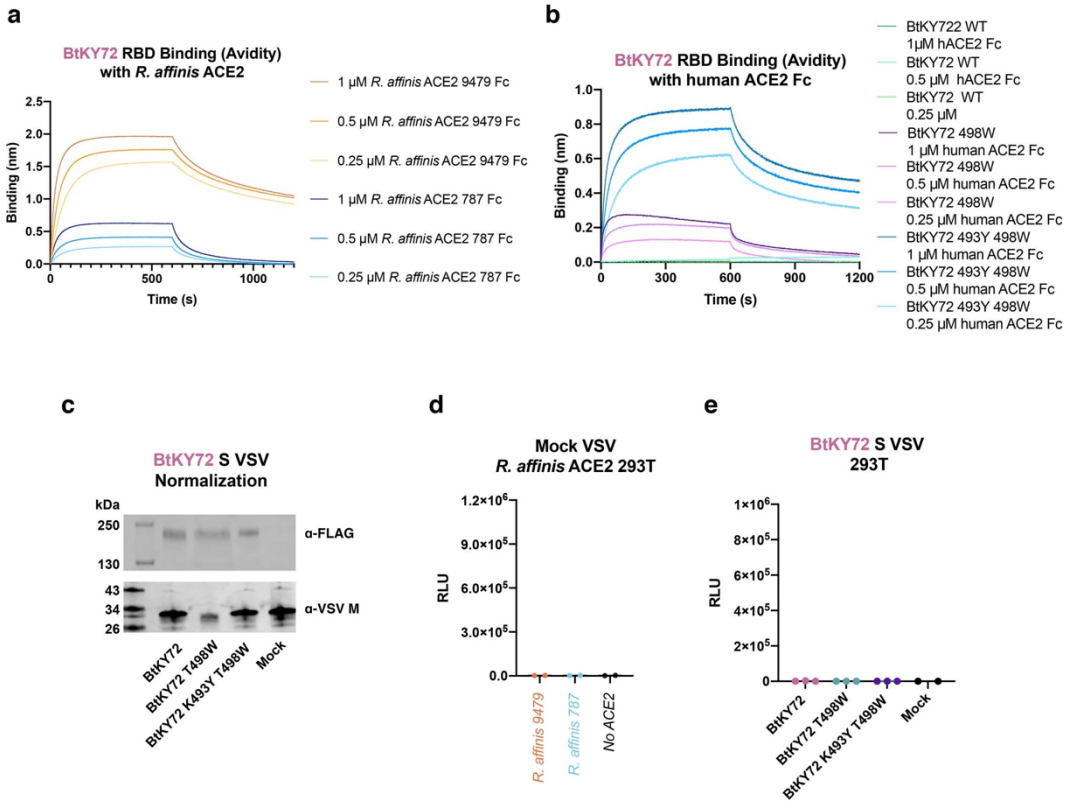
ZC45 and ZXC21), though recombination has not been reported among the Africa and Europe spikes and those in Asia. Scale bar, expected number nucleotide or amino-acid substitutions per site. Node labels illustrate bootstrap support values for sarbecovirus and Asia sarbecovirus monophyly. Sequences colored by their RBD clade as in **Fig. 2.1a**.



Supplemental Figure 2.2: Experimental details of Sort-seq assays.

(a) RBD yeast-surface display enables detection of folded RBD expression and ACE2 binding. (**b**) Representative gating for single (SSC-A versus FSC-A, SSC-W versus SSC-H, and FSC-W versus FSC-A), RBD+ (FITC versus FSC-A) cells. (**c)**

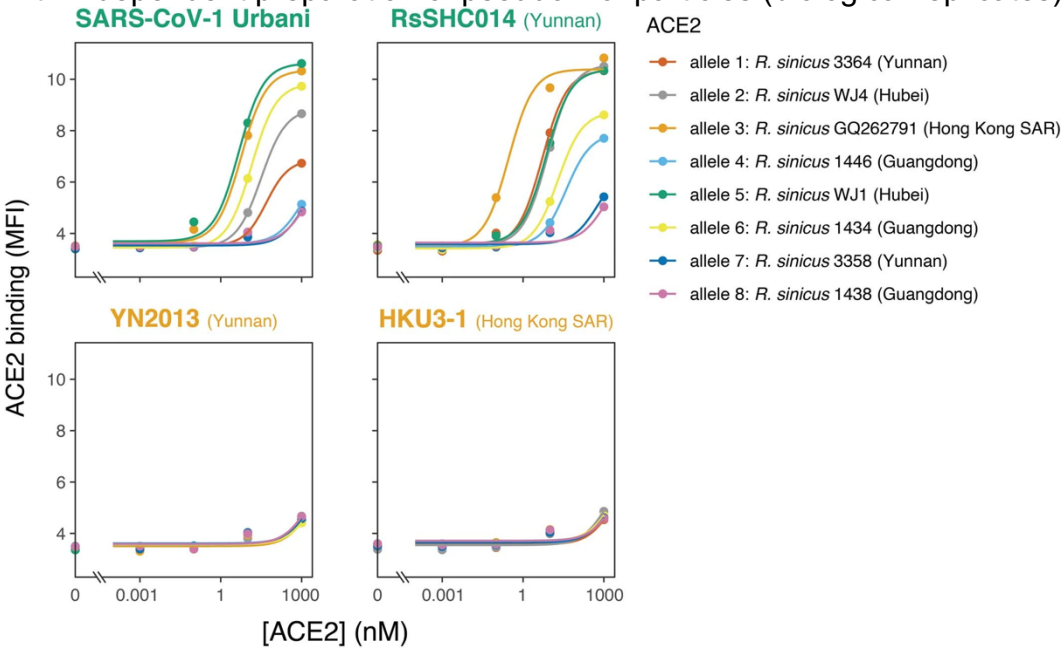
Representative bins drawn on single cells for expression Sort-seq measurements. (**d**) Representative bins drawn on single, RBD+ cells for ACE2 Tite-seq)(Starr, 2020; Adams, 2016) measurements. (**e)** Per-variant expression, shown as violin plots across replicate barcodes representing each variant within the gene libraries. (**f)** Number of distinct barcodes for each parental (top) or mutant (bottom) RBD genotype used in the determination of final pooled measurements across libraries. (**g)** Correlation in measured phenotypes between independently assembled and barcoded gene library duplicates for parental (top) or mutant (bottom) RBD genotypes.



Supplemental Figure 2.3: Normalization and controls for biolayer interferometry binding and pseudovirus entry assays.

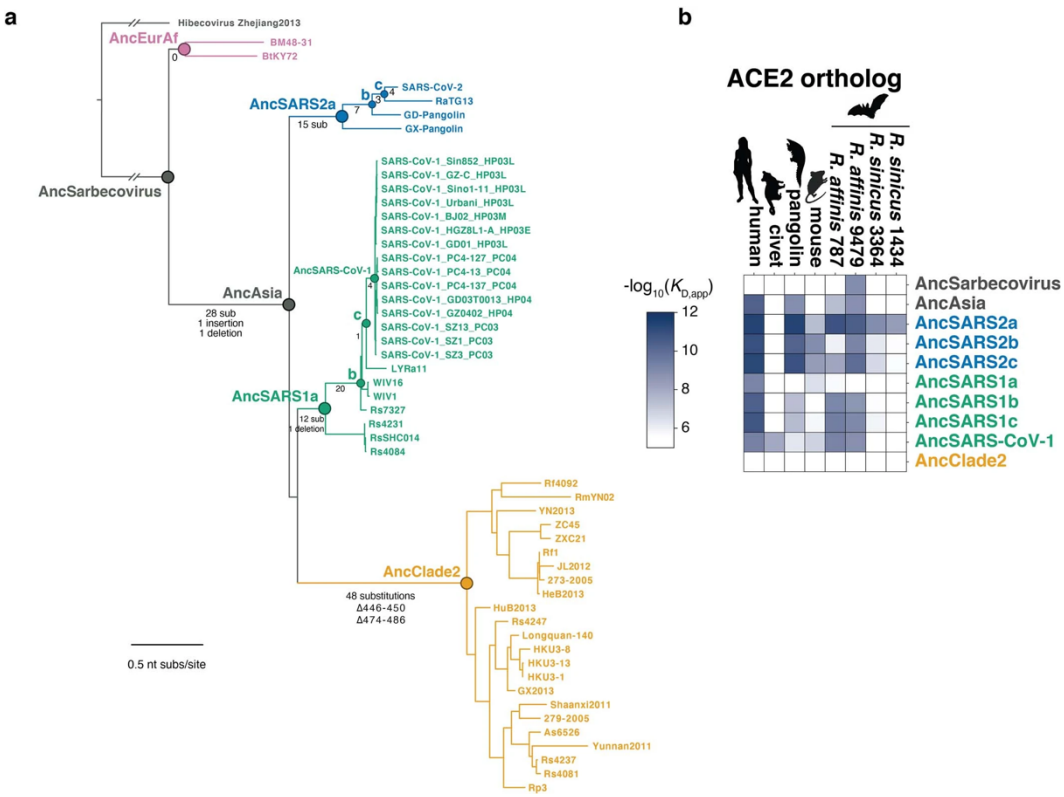
(a-b) Biolayer interferometry binding analysis of a range of *R. affinis* ACE2-Fc (**a**) or human ACE2-Fc (**b**) concentrations to biotinylated BtKY72 RBD (parental or mutant) immobilized at the surface of streptavidin biosensors. (**c)** Representative Western blots for quantification of spike incorporation into pseudoviral particles. Anti-FLAG (Sigma F3165) identifies incorporation of 3xFLAG-tagged spike, and anti-VSV-M (Kerafast EB0011) identifies level of VSV backbone. Viral inputs into cell entry assays were normalized across pseudoviral particles by S incorporation as determined in the anti-FLAG Western blot. Blot representative of biological duplicate generations of each pseudovirus. For gel source data, see **Fig. 2.15**)(**d)** Entry into *R. affinis* ACE2-expressing HEK293T cells by mock VSV particles produced in cells in which no spike gene was transfected. Each point represents the mean of technical triplicates for assays performed with independent preparation of pseudoviral particles (biological replicates). (**e**) Entry of pseudoviral particles into HEK293T cells not transfected with

any ACE2. Each point represents the mean of technical triplicates for assays performed with independent preparation of pseudoviral particles (biological replicates).



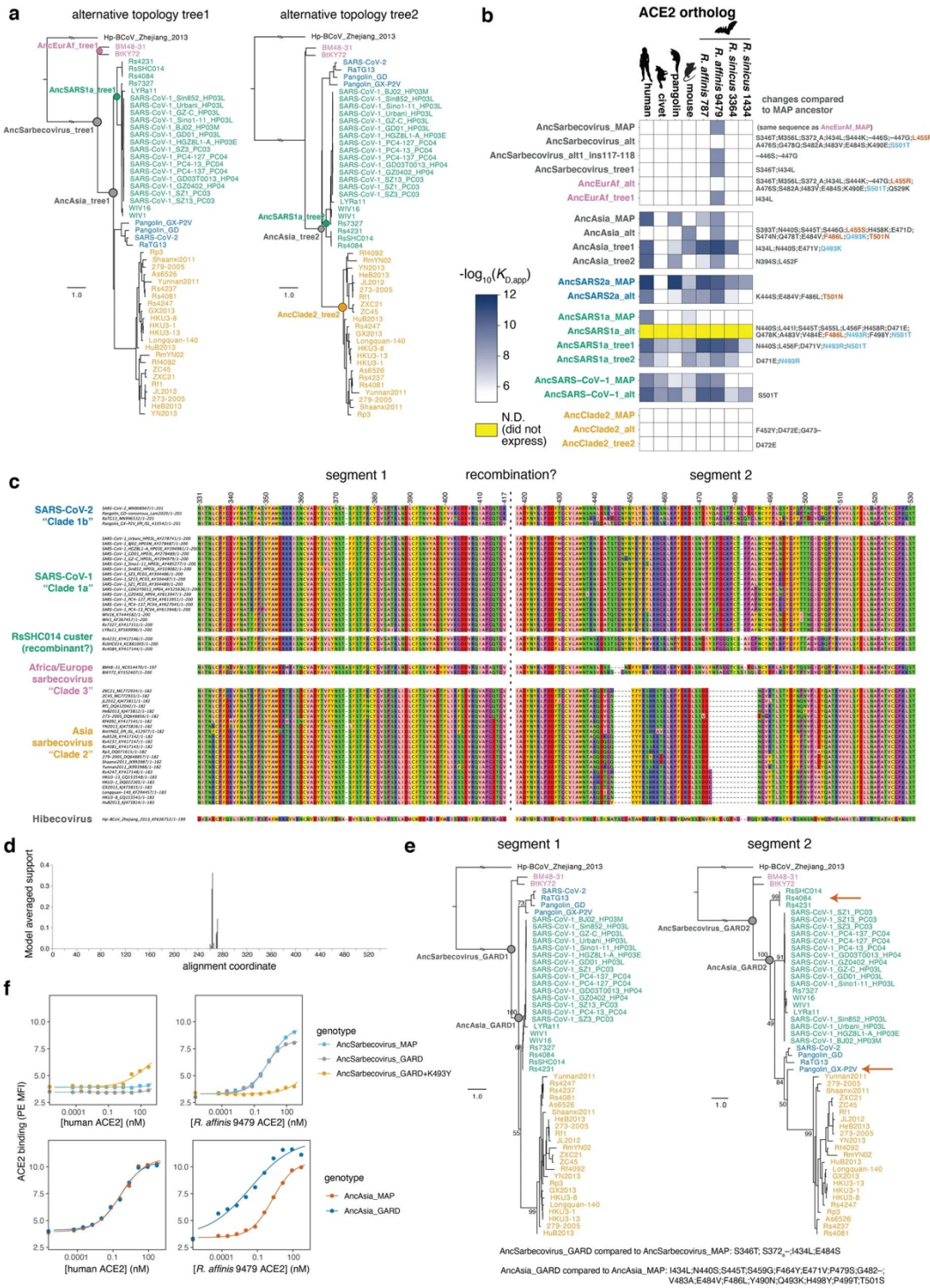
Supplemental Figure 2.4: Clade 2 RBD binding to an expanded panel of R. sinicus ACE2 alleles.

Binding curves for Clade 1a (SARS-CoV-1 Urbani and RsSHC014) and Clade 2 (YN2013 and HKU3-1) sarbecovirus RBDs for 8 *R. sinicus* ACE2 alleles. Measurements performed with yeast-displayed RBDs and purified dimeric ACE2 proteins, measured by flow cytometry. Data from a single experimental replicate. Region of sampling for bat sarbecovirusess and *R. sinicus* ACE2 alleles are provided. RsSHC014, YN2013, and HKU3-1 were all sampled from *R. sinicus* bats.



Supplemental Figure 2.5: Full set of RBD ancestral sequence reconstructions.

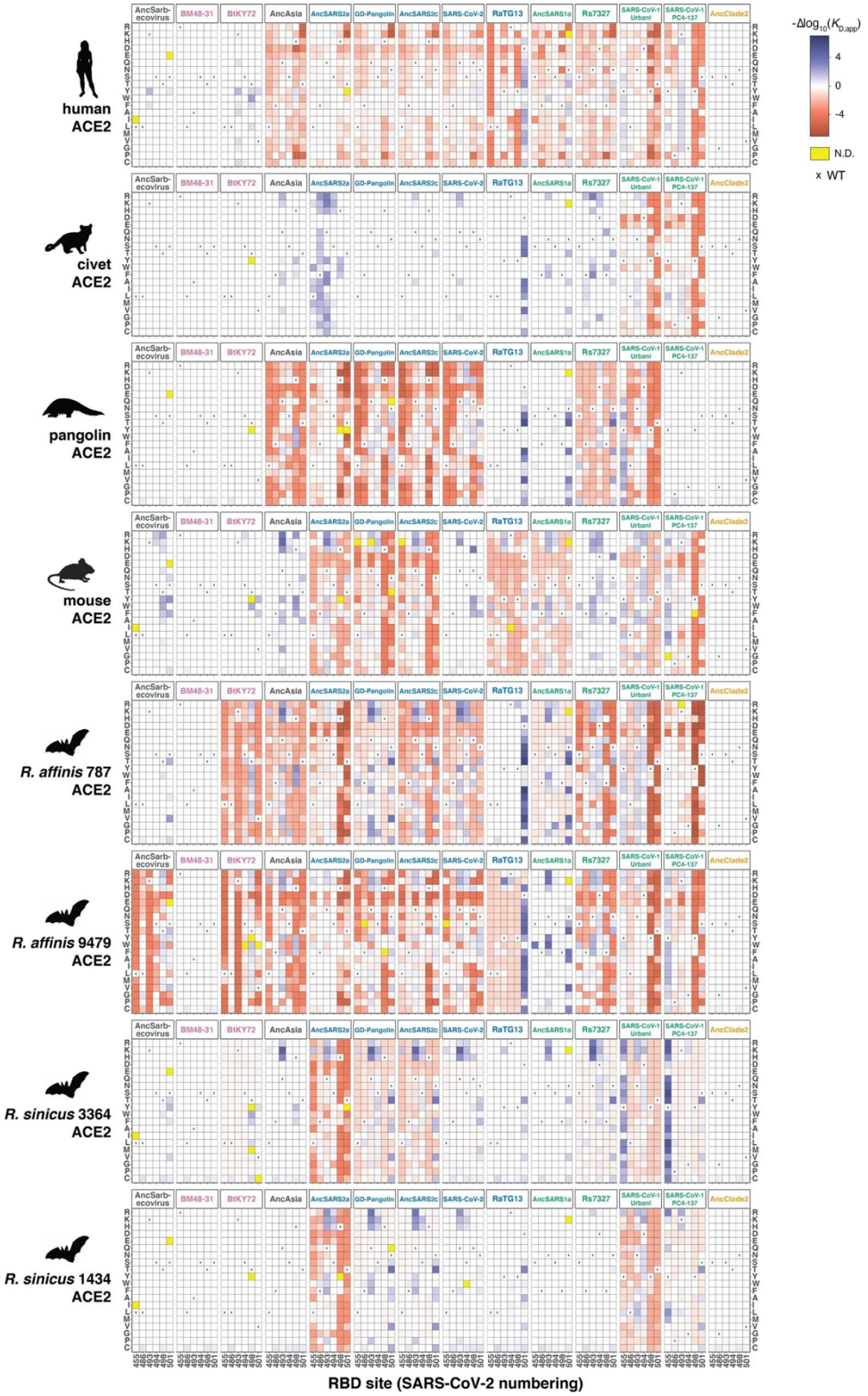
(**a**) Phylogeny with labelled nodes representing all ancestors tested, including nodes within the SARS-CoV-1 and SARS-CoV-2 clades leading to the human viruses. Branches are annotated with the number of amino-acid substitutions and indels that are inferred to have occurred along each branch. (**b**) Phenotypes of all most plausible ancestral sequences (including repetition of the data represented in **Fig. 2.2b**).



Supplemental Figure 2.6: Robustness to uncertainties in ancestral reconstructions.

(a) We performed ancestral sequence reconstructions on phylogenies constraining sister relationships between SARS-CoV-2 clade and clade 2 (tree1) or SARS-CoV-1 and SARS-CoV-2 clades (tree2) due to ambiguity in these relationships (**Fig. 2.1a and Fig. 2.5).** (**b)** ACE2 binding of alternative reconstructions. "Alt" ancestors incorporate all secondary reconstructed states with posterior probability > 0.2; (Eick, 2017) "tree1" and

“tree2” ancestors are inferred on the constrained trees in (**a**); and “ins117-118” tests the ambiguity of an indel separate from the remaining substitutions in AncSarbecovirus_alt. Sequence differences are listed at right relative to the maximum a posteriori (MAP) ancestors from **Fig. 2.2b** and **Fig. 2.9b**. Mutations are colored red if they were sufficient to abolish the ancestral phenotype and blue if they reinforced it (**Fig. 2.11**). Dramatic changes to inferred ancestral phenotypes are mostly observed in the alt ancestors which are the most probabilistically distant, while the tree1 and tree2 alternatives generally recapitulate the MAP phenotypes. The exception is AncSARS1a, where the tree1 and tree2 alternatives better match what would be expected based on the descendent RBD phenotypes (**Fig. 2.1b**). (**c)** RBD amino acid alignment, indicating a potential recombination breakpoint identified by GARD56 (from underlying nucleotide sequence). (**d)** Relative support values for possible recombination breakpoints. (**e)** Phylogenies inferred for the putative non-recombinant RBD segments. Arrows point to key changes in the segment 2 sub-tree. Each change is supported by weak bootstrap support values, and this hypothesis introduces a non-parsimonious history with respect to an indel at position 482. We reconstructed AncSarbecovirus_GARD and AncAsia_GARD as concatenated segment 1 and 2 reconstructions. Mutations that distinguish the GARD and MAP ancestor are listed at bottom. (**f**) Binding of GARD ancestors to human and *R. affinis* 9479 ACE2 was determined in isogenic yeast-display titrations.



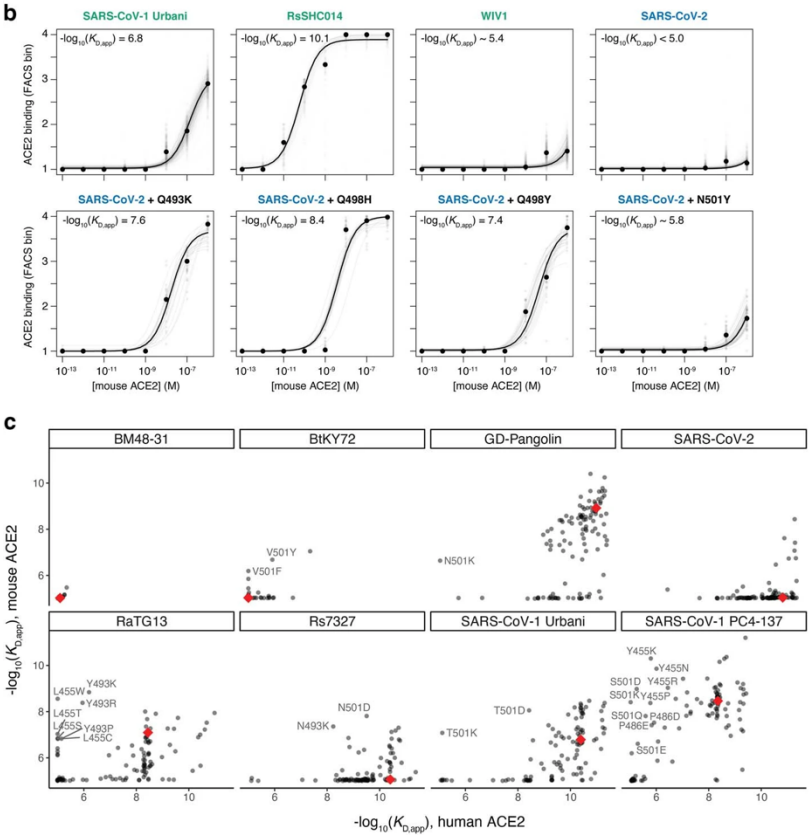
Supplemental Figure 2.7: Binding of RBD single mutants to each ACE2.

Each heatmap square illustrates the change in binding caused by the indicated mutation at the indicated position (SARS-CoV-2 numbering), according to the color key shown on

the upper-right. Yellow, mutations that were absent from the library or not sampled with sufficient depth in a particular experiment. x markers indicate the wildtype state at each position in each RBD background.

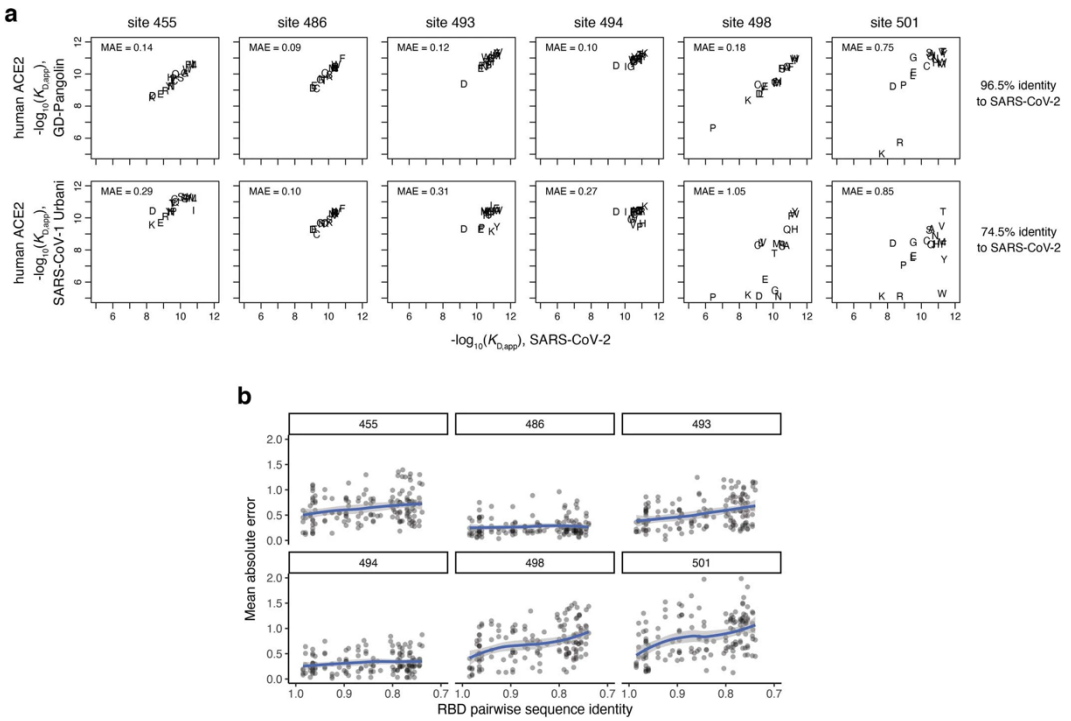
a

virus	replication in lung of young BALB/c mouse	pathogenesis in young BALB/c mouse	notes	reference
SARS-CoV-1 Urbani infectious clone (i.c.)	efficient	none	N/A	Roberts et al. *PLOS Path* 2007
SARS-CoV-1 Urbani MA15	enhanced compared to SARS-CoV-1 i.c.	weight loss, lethality	six changes selected in serial passage; S Y436H key to pathogenesis	Roberts et al. *PLOS Path* 2007; Frieman et al. *JVI* 2012
RsSHC014 i.c.	efficient, but lower than SARS-CoV-1 i.c.	none	N/A	Menachery et al. *Nat Med* 2015
RsSHC014 S/ MA15 backbone	similar efficiency as SARS-CoV-1 MA15	weight loss, no lethality	MA15 backbone minus S Y436H (RsSHC014 S)	Menachery et al. *Nat Med* 2015
WIV1 i.c.	inefficient	none	N/A	Menachery et al. *PNAS* 2016
WIV1 S/ MA15 backbone	inefficient	none	MA15 backbone minus S Y436H (WIV1 S)	Menachery et al. *PNAS* 2016
SARS-CoV-2 WA1 i.c.	undetectable	none	N/A	Dinnon et al. *Nature* 2020
SARS-CoV-2 WA1 MA	efficient	no weight loss or lethality; signs of impaired lung function	rationally designed, Q498Y+P499T	Dinnon et al. *Nature* 2020
SARS-CoV-2 MASCp6	efficient	no weight loss or lethality; signs of impaired lung function	five changes selected in serial passage; includes S N501Y	Gu et al. *Science* 2020
SARS-CoV-2 WBP-1	efficient	lethality	1 deletion and 5 nonsyn mutations selected in serial passage; includes S Q493K and Q498H	Huang et al. *EBioMedicine* 2021
SARS-CoV-2 B.1.1.7	inefficient	none	alpha VOC; includes S N501Y	Montagutelli et al. *bioRxiv* 2021
SARS-CoV-2 B.1.351	efficient	none	beta VOC; includes S K417N, E484K, N501Y	Montagutelli et al. *bioRxiv* 2021
SARS-CoV-2 P.1	efficient	none	gamma VOC; includes S K417T, E484K, N501Y	Montagutelli et al. *bioRxiv* 2021



Supplemental Figure 2.8: Existing data on sarbecoviruses in mice, and affinities of RBDs and key mutants for mouse versus human ACE2.

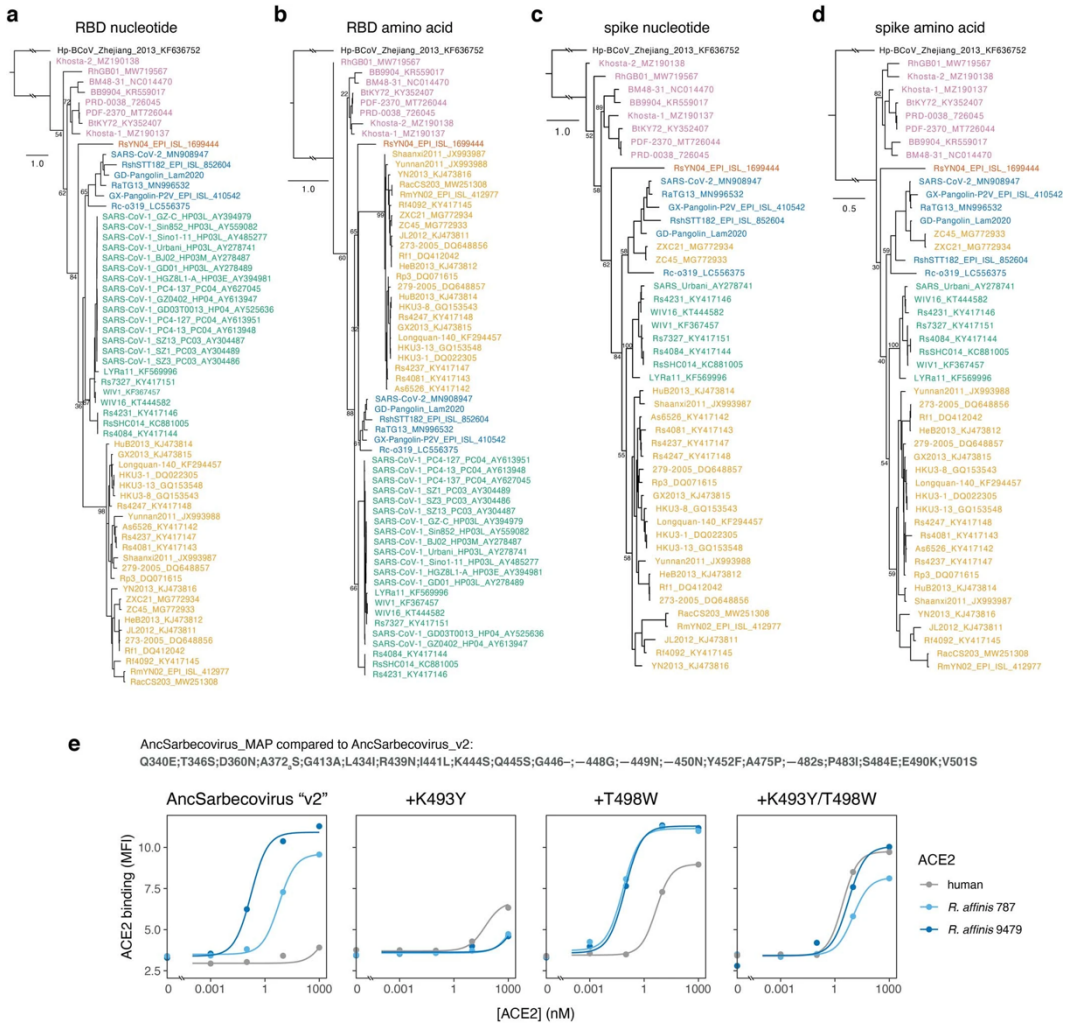
A, Summary of infectivity and pathogenesis of natural sarbecovirus and mouse-adapted strains from prior studies (Menachery, 2015; Dinnon, 2020; Roberts, 2007; Menachery, 2016; Frieman, 2012; Gu, 2020; Huang, 2021; Montagutelli, 2021). **b**, High-throughput titration curves for relevant genotypes from (**a**). Details as in **Fig. 2.1d**. Strength of binding to mouse ACE2 explains the infectivity and pathogenesis of SARS-CoV-1 Urbani and RsSHC014 (Menachery, 2015; Roberts, 2007), relative to the weak or absent replication of WIV1 (Menachery, 2016) and SARS-CoV-2 (Dinnon, 2020) in mice. Mutagenesis data explain the inefficient mouse infectivity of the SARS-CoV-2 B.1.1.7 isolate (Montagutelli, 2021) which incorporates the N501Y RBD mutation, relative to the efficient replication of the mouse-adapted SARS-CoV-2 isolate containing Q498Y (Dinnon,2020) or the pathogenic WBP-1 strain containing Q493K and Q498H (Huang, 2021). (**c)** An ideal mouse-adapted laboratory sarbecovirus strain would bind mouse ACE2 but not human ACE2 due to biosafety considerations. The large red points indicate the affinity of the parental RBD for human and mouse ACE2. The smaller black points indicate mutations, and key mutations that enhance binding to mouse versus human ACE2 are labelled (using SARS-CoV-2 numbering). Further mouse ACE2 specificity may be enabled via mutations at other positions not surveyed in our set of six positions.



Supplemental Figure 2.9: Epistasis and turnover in mutational effects.

(a) Example correlations in binding affinities for mutants in distinct RBD backgrounds at each site for human ACE2. Plots illustrate mutant avidities for human ACE2 and mean absolute error (residual) in the correlation for mutation measurements in GD-Pangolin (top) and SARS-CoV-1 Urbani (bottom) versus SARS-CoV-2. Plotting symbols indicate amino acid for each measurement. (**b)** Epistatic turnover in mutational effects across RBD backgrounds. Details as in **Fig. 2.3g**, but incorporating mutation effects among

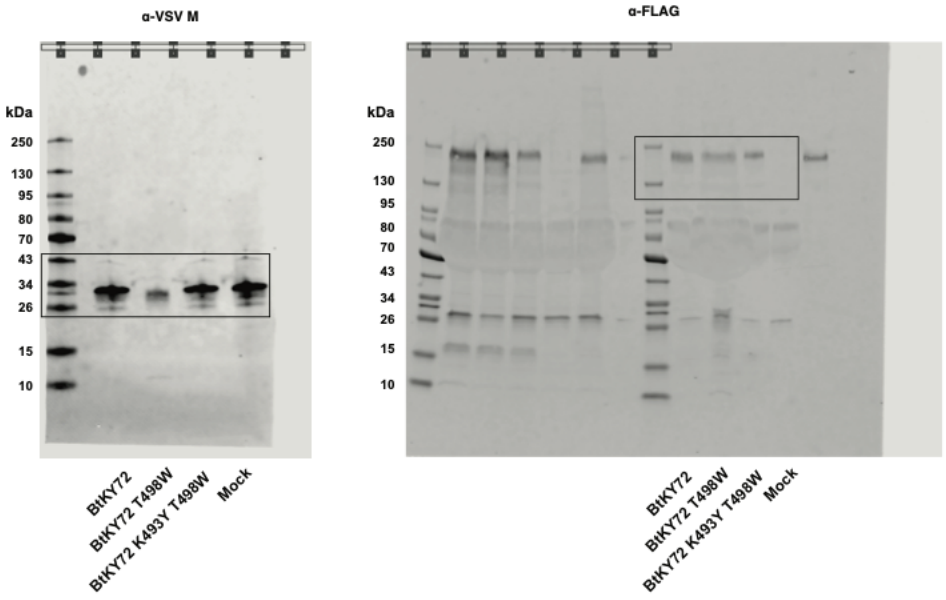
RBD pairs across all tested ACE2s. Blue line and shaded grey, LOESS mean and 95% CI trendline. See **Fig. 2.13b** for analysis across all ACE2 orthologues.



Supplemental Figure 2.10: Robustness of rooting and AncSarbecovirus phenotype in a phylogeny incorporating newly reported sequences.

a-d, Phylogenetic inference with inclusion of newly reported sarbecovirus sequences (**Fig. 2.4a**). As in **Fig. 2.5** , we infer phylogenies with RBD (**a**,**b**) and full spike alignments (**c**,**d**), both on nucleotide sequences (**a**,**c**) and translated amino acid (**b**,**d**) sequence alignments. The full set of outgroup betacoronavirus sequences shown in **Fig. 2.5** were also included in this tree inference but truncated from the display for visual clarity. The phylogeny in **Fig. 2.4a** is a constrained version of the RBD nucleotide tree from (**a**) where we constrained a monophyletic relationship among Africa/Europe sarbecoviruses due to uncertainty in the exact placement of the root within or relative ot the Africa/Europe sarbecovirus clade. (**e)** ACE2 binding by parental RBD and candidate mutants in an updated AncSarbecovirus sequence ("v2") inferred from the phylogeny in **Fig. 2.4a** that incorporates many newly described sarbecovirus RBDs, including some in important new phylogenetic locations. The unconstrained tree in (**a**) leads to inference of an AncSarbecovirus sequence that is identical to Khosta-2 (which also

binds ACE2). Sequence differences between the original MAP AncSarbecovirus and the “v2” reconstruction are listed at top. Measurements performed with yeast-displayed RBDs and purified dimeric ACE2 proteins, measured by flow cytometry. Data from a single experimental replicate.



Supplemental Figure 2.11: Uncropped images of the western blots in **Fig. 2.7c**.

CHAPTER 3. RECEPTOR BINDING DOMAIN CONFORMATION ACTS AS SPILLOVER BARRIER IN BAT CAMBODIAN SARBECOVIRUS, RSHTT200

For this next story, in an effort to understand the spillover of SARS-CoV-2 and the spillover potential of closely related viruses we did an in-depth study of a Cambodia bat-borne sarbecovirus, RshTT200. In the subsequent sections I describe the RshTT200 S structure and identify regions of divergence with important implications for receptor usage and antigenicity. I describe the structure of the RshTT200 RBD in complex with its host ACE2 receptor and detail the molecular interaction involved in its receptor association. I then detail its spillover potential across Rhinolophus bat species and other domesticated and wildlife animal species. I characterize it's RBD binding with human ACE2 but the absence of cellular entry with human ACE2. We identify a single nucleotide mutation that enable RshTT200 to enter human ACE2 but also demonstrate that RBD conformation also acts as a barrier of spillover. We identify the range of binding in which RBD conformational dynamics can act as a barrier to spillover, which is critical as all highly pathogenic CoV to date demonstrate RBD open and closed conformations, but bat-borne/ animal vector sarbecoviruses S structures solved to date have only demonstrated closed conformations.

3.1 Chapter Introduction

In the last two decades there have been three major outbreaks of severe and often fatal pneumonia due to the emergence of zoonotic coronaviruses (CoVs). Two of these events have resulted from sarbecovirus spillovers for which the native reservoir is Rhinolophus bats. Both severe acute respiratory syndrome coronavirus (SARS-CoV-1) and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) utilize the human ACE2 receptor to gain cellular entry during infection ((Walls et al. 2020; Hoffmann et al.

2020; Zhou et al. 2020; Li et al. 2003; Li et al. 2005; McCallum et al. 2022). Although the SARS-CoV-1 outbreak was contained after ~8,000 infections causing ~800 deaths, millions of deaths were recorded due to SARS-CoV-2. Sarbecoviruses found in wildlife, including in bat reservoirs, are continuously evolving and are likely to continue crossing the species barrier in the future, as exemplified by serological studies describing additional spillovers (Wang et al. 2018; Li et al. 2020; Evans et al. 2023) . Several coronaviruses closely related to SARS-CoV-2 (clade 1b (Starr et al. 2022)) have been identified in bats and pangolins and the study of these viruses may provide insights into both the evolution of sarbecoviruses as well as inform on future spillover potential for pandemic preparedness (Lam et al. 2020; Zhou et al. 2020; Zhou et al. 2021; Temmam et al. 2022).

Two such viruses, RshTT182 and RshTT200 were isolated from *Rhinolophus shameli* bats in 2010, with their sequences becoming publicly available in 2021 (Delaune et al. 2021). The identification of RshTT182/200 as well as other clade 1b viruses in Japan and Thailand demonstrates that SARS-CoV-2-like coronaviruses were circulating outside of China at least a decade prior to the emergence of SARS-CoV-2 (Murakami et al. 2020; Wacharapluesadee et al. 2021). The more recent identification of several bat sarbecoviruses very similar to SARS-CoV-2 in Laos further supports the idea that these viruses are widely circulating in bat populations across Asia (Temmam et al. 2022). It is therefore key to study these viruses to understand the origin and evolution of SARS-CoV-2 and allow identification of members with spillover potential.

The spike glycoprotein (S) initiates coronavirus infection by promoting binding to host receptors and fusion of the viral and host membranes ((Walls et al. 2016; Walls et

al. 2017; Tortorici and Veesler 2019; Walls et al. 2019; Walls et al. 2020) Furthermore, S is the predominant target of neutralizing antibodies, due to its essential role in cellular entry and its crown-like organization at the surface of coronavirus particles (Piccoli et al. 2020; Bowen et al. 2022; Greaney et al. 2021). As a result, coronavirus S participates in determining host and tissue tropism as well as pathogenicity, and is a key factor governing the ability of coronaviruses to cross the species barrier and trigger outbreaks (Menachery et al. 2015). Studies of sarbecovirus S glycoproteins enable the identification of viruses with a high likelihood of spillover to humans and inform the deployment of monoclonal antibody therapies and the design of broadly protective vaccines (Xiong et al. 2022). Here, we evaluate the RshTT200 S receptor species tropism , explore the balance between receptor affinity and accessibility for human ACE2-dependent entry of RshTT200 S pseudovirus into cells, and identify countermeasures that could be implemented for pandemic preparedness.

3.2 Architecture of the RShTT200 S Trimer

To unveil the organization of the RshTT200 infection machinery, we characterized a trimeric ectodomain S construct comprising residues 14-1189 with the HexaPro prefusion-stabilizing mutations (Hsieh et al. 2020) using single particle cryo-electron microscopy (cryoEM). 3D classification of the cryoEM dataset revealed a single conformation of the S trimer, with all three receptor binding domains (RBDs) adopting a closed state, for which we determined a structure at 2.8Å overall resolution (**Fig. 3.1, Fig. S3.1- S3.2**). In addition, a S_1 subunit class was observed, wherein all three RBDs were also closed (**Fig. S3.2A**). Stabilization of the RshTT200 prefusion S was

unsuccessful with the more conservative 2P mutation that has been widely used for other coronaviruses (SARS-CoV-1, SARS-CoV-2, and MERS-CoV) (**Fig. S3.1**).

The RshTT200 S atomic model comprises residues 15-1120, with internal breaks corresponding to flexible regions. Although the RshTT200 S trimer structure resembles the closely related prefusion SARS-CoV-2 S structure (Walls et al. 2020), with which it shares ~84.5% amino acid sequence identity, there are several architectural differences in the N-terminal domain (NTD) and in the RBD, that are the main targets of neutralizing antibodies in SARS-CoV-2 infected and in vaccinated individuals (Piccoli et al. 2020; Stamatatos et al. 2021; Greaney et al. 2021; McCallum et al. 2021; Bowen et al. 2022). The RshTT200 and SARS-CoV-2 S (PDB: 6VXX) structures can be superimposed with an Cα R.M.S.D. of 4.0Å (**Fig. S3.3**). The Cα R.M.S.D. drops to 0.813Å for S_2 (**Fig. S3.3, Table S3.1**), concurring with the high sequence conservation of the fusion machinery. RshTT200 S lacks a polybasic cleavage site at the S_1/S_2 junction (**Fig. 3.1F**), like other closely related sarbecoviruses discovered to date except SARS-CoV-2. Some important structural features are virtually identical to SARS-CoV-2 S including the S_2‘ cleavage site, the fusion peptide region (**Fig. 3.1F**), and the stem helix (**Fig. 3.1G**). While others like the NTD, the RBD, the C domain and the D domain, demonstrate higher levels of diversity (**Fig. 3.1H-J**).

The RshTT200 NTD is the region of highest sequence and structural divergence from SARS-CoV-2 (**Fig. 3.1H-I**) (Delaune et al. 2021). Specifically, the topological equivalent of the SARS-CoV-2 NTD antigenic supersite (McCallum et al. 2021; Suryadevara et al. 2021; Cerutti et al. 2021) is completely remodeled in the RshTT200 NTD due to residue deletions in the supersite β-hairpin and the supersite loop, the latter

region being largely disordered in the RshTT200 cryoEM map. Moreover, the RshTT200 NTD lacks several N-linked glycans relative to SARS-CoV-2, such as the equivalent of the SARS-CoV-2 N17 and N149 oligosaccharides that are flanking the antigenic supersite (**Fig. 3.1I, Table S3.2**). Conversely, the RshTT200 NTD possesses an N-linked glycosylation at position N111 which is absent in SARS-CoV-2 S (N112) but shared with SARS-CoV-1 S (N109) (**Fig. 3.1I**). These major structural changes are reminiscent of the structural plasticity of the SARS-CoV-2 variant NTDs, and explain the lack of cross-reactivity of most SARS-CoV-2 NTD-targeted antibodies with distinct variants or other sarbecoviruses (McCallum et al. 2022; McCallum et al. 2021; McCallum et al. 2021; McCallum et al. 2021; Bowen et al. 2022).

The RshTT200 and SARS-CoV-2 RBDs share 81% amino acid sequence identity and their backbones differ most notably in the helix comprising residues 351-360 (SARS-CoV-2 S residues 364-373) based on apo SARS-CoV-2 S structures (Walls et al. 2020; Wrapp et al. 2020) . This RshTT200 helix, however, has an identical conformation to the one identified in the linoleic acid-bound prefusion SARS-CoV-2 S structures (Toelzer et al. 2020; Wrobel et al. 2020) and our cryoEM map resolves a linoleic acid bound to the RshTT200 RBD through strictly conserved interactions (**Fig. S3.3**). This finding likely explains the observed closed S conformation due to linoleic acid-mediated bridging of two neighboring RBDs in the closed state (**Fig S3.3A**). The RshTT200 HexaPro S trimer resembles SARS-CoV-2 “locked” S, wherein RBDs are tightly packed (Xiong et al. 2020), and differs from SARS-CoV-2 HexaPro S for which structures with open RBD conformations have been solved (Hsieh et al. 2020; Li et al. 2024). The prevalence of closed RBD conformations is also reminiscent of other bat

and pangolin sarbecovirus S structures that have been described to date (Wrobel et al. 2020; Zhang et al. 2021; Wrobel et al. 2021; Lee et al. 2023), several of which also contained linoleic acid bound to their RBDs through similar interactions.

The RshTT200 RBD possesses an N-linked glycan at position N357, which is absent in SARS-CoV-2 S (N370) but present in the RBD of all other sarbecoviruses identified to date. **(Fig. 3.1J, Table S3.2**), and can be attributed to T359 (RshTT200)/A372 (SARS-CoV-2) at position +2 of the glycosylation sequon. This alteration of glycosylation may impact antigenicity compared to SARS-CoV-2, as previously observed by deep-mutational scanning (Lee et al. 2023). Sarbecovirus ACE2 tropism has been found to be correlated with the presence of indels and deletions along two sites (indel site i and ii) within the receptor binding motif (RBM), wherein long RBM exhibit the broadest tropism, and those with deletions in either site i or ii exhibiting narrower ACE2 tropism (Si et al. 2024). The RshTT200 RBD harbors a markedly shorter 430-434 loop (indel site i) than the equivalent SARS-CoV-2 RBD loop (residues 443-451) as well as a single amino acid deletion corresponding to SARS-CoV-2 residue G482 (indel site ii) (**Fig. 3.1J**). In addition to the potential impact on ACE2 tropism, both deletions map to opposite sides of the SARS-CoV-2 RBD antigenic site I that is targeted by several monoclonal antibodies that have been used clinically during the COVID-19 pandemic (Corti et al. 2021; Piccoli et al. 2020; Barnes et al. 2020; Westendorf et al. 2022; Park et al. 2022; Zost et al. 2020; Dong et al. 2021; Hansen et al. 2020; Baum et al. 2020).

3.3 Structural Differences in RshTT200 S result in altered antigenicity

To probe the impact of these notable structural differences between RshTT200 and SARS-CoV-2 on antigenicity of the RshTT200 S, we selected a panel of NTD, RBD, and

S2 targeting antibodies and performed ELISA assays to compare the EC50s of their binding to the S glycoproteins of SARS-CoV-2 and RshTT200. None of the NTD antibodies targeting the the antigenic supersite that flanks N17 and N149 in SARS-CoV-2, SL28, S2M11, or S2X333 bound the RshTT200 S (**Fig. S3.4A-C**)**.** Due to the high degree of variance in the N-terminal domain including deletions in the neighboring contact residues for S2X333, S2L28, and S2M11 of RshTT200 equivalent to SARS-CoV-2 residues 152-157, and 247-252 (**Fig. 3.1I**). These result in significant changes along the antigenic supersite, most noticeably along residues 143-157 (**Fig. 3.1I**). However, residues corresponding to 245-252 are not resolved in the RshTT200 structure, similar to the antibody unbound SARS-CoV-2 structures (PDB: **6VXX**).

Most neutralizing antibodies however target several sites in the RBD of SARS-CoV-2. Therefore, we next assessed whether broadly neutralizing RBD targeting antibodies would bind the RshTT200 S. This revealed site ii targeting antibodies, S2X35 and S2X259 bound the RshTT200 S with comparable EC50s to SARS-CoV-2 S (**Fig. 3.1K; Fig S3.4D-E,J-K**). Furthermore, site v (S2H97) and iv (S309) targeting antibodies, S2H97 and S309 showed similar trends of binding with the RshTT200 S and SARS-CoV-2 S (**Fig. 3.1K; Fig. S3.4F-G, J-K**). S2E12, a site i binding antibody failed to bind the RshTT00 S, whereas has been previously shown to bind the closely related sarbecoviruses RaTG13, Pangolin GD and Pangolin GX (Starr et al. 2021). The absence of binding observed is most likely due to the P468 in RshTT200; in SARS-CoV-2, the equivalent mutation F486P has been demonstrated to be an escape mutation from S2E12 (Park et al. 2022).

While S2 targeting antibodies are identified more rarely, the high level of conservation of the S2 subunit makes them a promising candidate for broadly neutralizing antibodies. The stem helix binding monoclonal antibody P6 (Pinto et al. 2021), targets a highly conserved epitope on the RshTT200 S (**Fig. 3.1G**) and is bound the RshTT200 S and SARS-CoV-2 S equivalently (**Fig. 3.1K**). The RshTT200 fusion peptide, which is identical between RshTT200 and SARS-CoV-2 (**Fig 3.1F**) is targeted by RAY53 (Silva et al. 2023), which recognizes the fusion machinery apex and and 76E1 (Sun et al. 2022); both of which recognized RshTT200 equivalently to SARS-CoV-2 (**Fig. S3.4K**). Collectively, these antibodies highlight targetable regions of conservation between RshTT200 and identify monoclonal antibodies that could be stockpiled for use in the event of spillover.

3.4 Molecular Basis of RshTT200 binding of host receptor

To understand the potential for RshTT200 spiller, an understanding of receptor usage in its host species is necessary. Previous studies indicated that *R. shameli* ACE2 could promote cellular entry of SARS-CoV-2 S and RshTT200 S pseudoviruses, concurring with the fact that RshTT200 was identified in this bat species (Delaune et al. 2021)). To compare the binding of RshTT200 and SARS-CoV-2 RBD to their cognate receptors, we utilized natively dimerized ACE2 ectodomain and assessed binding by Bio-layer Interferometry (BLI). SARS-CoV-2 binding with dimeric human ACE2 yielded an apparent affinity (KD_{app}) (due to avidity) of 0.75 nM compared to RshTT200 binding with dimeric *R. shameli* ACE2 with a KD_{app} of 77 nM (**Fig. 3.2A-B, S3.5, S3.6, Table S3.2, S3.3**). Comparatively, SARS-CoV-2 had a much weaker association with the *R.*

shameli ACE2 of ~30μM (**Fig. 23.C, Table S3.2**), which is nevertheless a functional entry receptor (possibly due to overexpression in our cell entry assay). (**Fig. 3.2D**). To compare these interactions to those that have been observed between other bat-borne sarbecoviruses and their hosts, we measured the binding avidity between the RaTG13 RBD and its host *R. affinis* ACE2 receptor. *R. affinis* is among the Rhinolophus species that are commonly infected with sarbecoviruses and show polymorphic variation of their ACE2 receptors. Due to the overall low binding affinity of RaTG13 and limitations of BLI, we were unable to surround the KD in our BLI assays so an approximate KD_{app} of 3 μM is reported with *R. affinis* ACE2 allele 9479. Binding kinetics with *R. affinis 787* with its weaker association were not determined. While previous reports have explored of RatG13 binding to its host's, *R. affinis,* ACE2 (33μM by SPR) such studies have not taken into account the impact of polymorphism in the ACE2 genes of *R. affinis,* nor have they explored the enhanced binding enabled by ACE2 dimerization and the resultant avidity, as it exists on the cell surface (Liu et al. 2021). However, even accounting for this in our assay points to micromolar affinity for its host receptor, a marked difference from RshTT200 and its nanomolar affinity with *R. shameli* ACE2. While RshTT200 binding is weaker than that of SARS-CoV-2_{Wuhan} for human ACE2, it is considerably stronger than that of RaTG13 for *R. affinis* ACE2.

To characterize the molecular basis of RshTT200 S-mediated recognition of the reservoir host *R. shameli* ACE2 receptor, we characterized a complex between the RshTT200 RBD and the natively dimeric *R. shameli* ACE2 (residues 19-740) using single particle cryoEM (**Fig. S3.8**, **S3.9**). Symmetry-expansion and local refinement yielded a structure at 3.1 Å resolution of the ACE2 peptidase domain bound to the RBD revealing

the interactions at the complex interface. Although *R. shameli* ACE2 shares 80.7% overall amino acid sequence identity with human ACE2 (hACE2), the conservation decreases to 60.7% when considering only the 28 hACE2 residues that face the SARS-CoV-2 RBD in published structures (**Table S3.4, S3.8**). Nevertheless, similar surface areas are buried at the *R. shameli* ACE2/RshTT200 RBD interface (837.5 $Å^2$) as compared to the hACE2-bound to the SARS-CoV-2 RBD complex pdb 6m0j (~840 $Å^2$) (Lan et al. 2020). Our *R. shameli* ACE2 structure reveals differences in N-linked glycosylation along the sarbecovirus interacting surface most notably with absence of glycosylation at $N90_{hACE2}$ and with the addition of $N329_{hACE2}$ however N90, N322, and N546 are conserved in our structure (**Fig. S3.9**).

Like SARS-CoV-2- hACE2 structures, the RshTT200- *R. shameli* ACE2 interaction relies on a network of hydrogen bonds, electrostatic interactions, and van der Waals interactions. The RshTT200 RBD interface is largely conserved with five of six positions previously implicated in the evolution of human ACE2 receptor binding for SARS-CoV-2 and SARS-CoV being conserved in the RshTT200 RBD: L438 (SARS-CoV-2 L455), Q475 (SARS-CoV-2 Q493), S476 (SARS-CoV-2 S494), Q480 (SARS-CoV-2 Q498), N483 (SARS-CoV-2 N501) whereas P468 corresponds to a non-conservative mutation relative to SARS-CoV-2 F486 (Starr et al. 2020; Wan et al. 2020). However, this comes at the consequence of disrupting interactions that in SARS-CoV-2 are facilitated by F486, specifically π-stacking with $Y83_{hACE2}$, and Van Der Waals interactions with $M82_{hACE2}$ (**Fig. 3.2M**). However this interaction interface would also be altered by differences on the hosts ACE2 ($F83_{RshACE2}$ and $N82_{RshACE2}$)(**Fig 3.2L**). Additionally, tolerance of proline at this site is evolvable in different backgrounds and is not deleterious

in XBB.1.5. Despite the tolerance of mutation without loss of ACE2 binding in some SARS-CoV-2 variant backgrounds, deep mutational scanning at this position shows mutation at $P468_{RshTT200}$ is favored, most strongly with P468L, but P468F as seen in SARS-CoV-2 is also weakly favored. (**Fig. 3.2N**).

Like SARS-CoV-2- hACE2 structures, the RshTT200- *R. shameli* ACE2 interaction relies on a network of hydrogen bonds, electrostatic interactions, and van der Waals interactions. The RshTT200 RBD interface is largely conserved with five of six positions previously implicated in the evolution of human ACE2 receptor binding for SARS-CoV-2 and SARS-CoV being conserved in the RshTT200 RBD: L438 (SARS-CoV-2 L455), Q475 (SARS-CoV-2 Q493), S476 (SARS-CoV-2 S494), Q480 (SARS-CoV-2 Q498), N483 (SARS-CoV-2 N501) whereas P468 corresponds to a non-conservative mutation relative to SARS-CoV-2 F486 (Starr et al. 2020; Wan et al. 2020). As previously mentioned $P468_{RshTT200}$ in place of $F486_{SARS\text{-}CoV\text{-}2}$ is observed in correlation with a deletion $G482_{SARS\text{-}CoV\text{-}2}$ in RshTT200 to presumably maintain the overall structure of the x-x loop. However, this comes at the consequence of disrupting interactions that in SARS-CoV-2 are facilitated by F486, specifically π-stacking with $Y83_{hACE2}$, and Van Der Waals interactions with $M82_{hACE2}$ (**Fig. 3.2M**). However, this interaction interface would also be altered by differences on the hosts ACE2 ($F83_{RshACE2}$ and $N82_{RshACE2}$)(**Fig. 3.2L**). Additionally, tolerance of proline at this site is evolvable in different backgrounds and is not deleterious in XBB.1.5. Despite the tolerance of mutation without loss of ACE2 binding in some SARS-CoV-2 variant backgrounds, deep mutational scanning at this position shows mutation at $468_{RshTT200}$ is

favored, most strongly with P468L, but P468F as seen in SARS-CoV-2 is also weakly favored. (**Fig. 3.2N**).

Other notable alterations across the ACE2: RBD binding interface were observed in our structure. Specifically, while D469 side chains were unable to be built into our density map, they are expected to fall in range for the formation of a novel but weak salt bridge between $R24_{RshACE2}$ - $D469_{RshTT200,}$ (**Fig. S3.10B**). Supporting the role of this interaction, deep mutational scanning revealed any mutation other than R24 is detrimental to *R. shameli* binding. This differs from the SARS-CoV-2 for which the equivalent site $N487_{SARS\text{-}CoV\text{-}2}$ participates in hydrogen bonding with $Q24_{hACE2}$ (**Fig. S3.10C**). Another difference in the host ACE2 compared to human ACE2, is that $H41_{RshACE2}$ facilitates hydrogen bonding with $D355_{RshACE2}$ (an interaction that is not observed in human ACE2 due to the presence of $G355_{hACE2}$). This interaction is clearly supported by our density map (**Fig. 3.2F-G**). An alternative rotamer conformation of $H41_{RshACE2}$ could enable hydrogen bonding with $N483_{RshTT200}$ (similar to the interaction between $N501_{SARS\text{-}CoV\text{-}2}$ and $Y41_{hACE2}$) (**Fig. S3.10**); however, our density map more strongly supports interactions with $D355_{RshACE2}$.

The presence of $Y478_{RshTT200}$ enables van der waals interactions with $E38_{RshACE2}$, an interaction that is absent between the equivalent $496G_{SARS\text{-}CoV\text{-}2}$ and $D38_{hACE2}$. Interactions between $Q498_{SARS\text{-}CoV\text{-}2}$ and $Q42_{hACE2}$/$K353_{hACE2}$, are also conserved in our structure between $Q480_{RshTT200}$ and $Q42_{RshACE2}$/$K353_{RshACE2}$(**Fig. 3.2F-G**). In addition, hydrogen bonding between $K353_{hACE2}$ and the backbone carbonyl of $G496_{SARS\text{-}CoV\text{-}2}$ is conserved in our structure between $Y478_{RshTT200}$ and $K353_{RshACE2}$(**Fig.**

3.2F-G). Other structural differences along the ACE2 binding interface have implications for ACE2 binding. The markedly shorter 430-434 loop (SARS-CoV-2 residues 443-451) results in the loss of key interactions observed between $Y449_{SARS\text{-}CoV\text{-}2}$ and $Q42/D38_{hACE2}$ (**Fig. 3.2H-I**). The loss of these interactions likely greatly contributes to the overall reduction of the RshTT200 RBD binding with its host receptor *R. shameli* compared with SARS-CoV-2 RBD binding with human ACE2 (**Fig. 3.2A-B**). However, the introduction of $Y478_{RshTT200}$ presumably acts to compensate at least marginally for the loss of these interactions, through Van der Waals interactions with E38 (**Fig. 3.2H**). Together these differences explain the markedly different binding of RshTT200 and SARS-CoV-2 with their cognate receptors.

3.5 Several Rhinolophus ACE2 alleles are entry receptors for RShTT200

Several sarbecoviruses have been identified in Rhinolophus bats in Southeast Asia (Hu et al. 2017; Wang et al. 2022; Ge et al. 2013; Latinne et al. 2020) and it has been shown that co-infections with two or more sarbecoviruses can occur in wild bats (Wang et al. 2022). Moreover, co-roosting of some of these bat species may favor viral transmission, evolution and recombination. Several SARS-CoV-1-related (clade 1a) and SARS-CoV-2-related (clade 1b) sarbecoviruses were identified from *R. sinicus* (e.g. WIV1 (Ge et al. 2013) and WIV-16 (Yang et al. 2015)) and *R. affinis* bats (e.g. RaTG13 (Zhou et al. 2020) and LyRa11 (He et al. 2014), Ra22QT10 (Hassanin et al. 2023)), which share geographical overlap with *R. shameli*. Sequencing data revealed marked bat ACE2 polymorphism which led to the description of eight and two distinct ACE2 alleles for *R. sinicus* and *R. affinis*, respectively, defined based on sarbecovirus RBD contact residues, (**Table S3.4**) (Starr et al. 2022; Guo et al. 2020). Notably *R. affinis*

ACE2 receptors have been shown to be highly promiscuous receptors for sarbecoviruses ((Starr et al. 2022; Guo et al. 2020; Lee et al. 2023)).

To assess the promiscuity of RshTT200 host receptor usage compared to other bat-borne clade 1a and clade 1b sarbecoviruses, we evaluated binding of a panel of Rhinolophus ACE2 orthologs harboring a human Fc-fusion to immobilized RBDs using BLI (**Fig. S3.10-S11**). Our ACE2 panel featured 8 distinct *R. sinicus* ACE2 alleles and 2 *R. affinis* ACE2 alleles selected to represent the diversity of polymorphic variants at ACE2:RBD contact sites. RshTT200 bound three ACE2 alleles: *R. affinis* 9479 (RA9479), *R. affinis* 787 (RA787) and *R. sinicus* WJ1 (RSWJ1) (**Fig. 3.3**, **Fig. S3.10-S3.11**). RaTG13 demonstrated narrower binding breadth relative to RshTT200, interacting only with *R. affinis* ACE2 allele 9479 (**Fig. S3.10-S3.11, S3.13**). Comparatively, SARS-CoV-1 and SARS-CoV-2 RBDs showed broad ACE2 promiscuity bound seven and five ACE2 alleles respectively. Another bat borne sarbecovirus WIV1, showed similar promiscuity to SARS-CoV-1, reflective of its close evolutionary relationship with SARS-CoV-1 (clade 1a), though the identities of the ACE2 alleles it bound were slightly altered (**Fig. S3.10-S3.11**). Overall, these findings are in line with the observation that several clade 1a sarbecoviruses have been isolated from *R. sinicus* bats and therefore evolved to interact with *R. sinicus* bat ACE2. The polymorphic variation in *R. sinicus* bats has been suggested to result from an evolutionary arms race between this species and the sarbecoviruses they host and thus clade 1a sarbecovirus (like SARS-CoV-1 and WIV1), which have co-evolved within this species demonstrate broader tropism across R. sinicus ACE2 allele variants than clade 1b (like SARS-CoV-2, RshTT200, and RaTG13) that have undergone other selective pressures. To date no

clade 1b sarbecoviruses (such as SARS-CoV-2, RatG13, and RshTT200) have been observed in *R. sinicus* bats and have only been isolated from other Rhinolophus species.

Having established which Rhinolophus ACE2 receptors could interact with the RshTT200 RBD, we next evaluated the kinetics and avidity of binding of the natively dimerized RA9479, RA787, and RSWJ1 ACE2 ectodomains to the immobilized RshTT200 RBD. We determined K_{Dapp} of 206 nM, 55.1 nM, 82.5 nM, respectively (**Fig. 3.3A-C**). Interestingly, unlike other sarbecoviruses in our panel, the RshTT200 RBD demonstrated enhanced binding to the RA787 allele, relative to RA9479 (**Fig. S3.10**, **Fig. 3.3A-C**). These two alleles differ only by two mutations at the binding interface: $H34_{RA9479}/R34_{RA787}$ and $D38_{RA9479}/E38_{RA787}$ (**Fig. 3.3A-B,D**).

To understand the molecular basis of these differences we used Alphafold 3 to predict the structure of these ACE2 alleles in complex with the RshTT200 RBD (**Fig. 3.3F-G**, **Fig. S3.12D-I**). Like $E38_{RShACE2}$ in our RshTT200 RBD: *R. shameli* ACE2 structure (**Fig. S3.12A**), $E38_{RA787ACE2}$ is modeled in an orientation that would enable van der Waals interactions with Y478 (**Fig. S3.12G**). This differs from $D38_{RA9479ACE2}$ with its shorter side chain which is predicted to have less optimal positioning for forming such interactions (**Fig. S3.12D**). In our RshTT200 RBD *R. shameli* ACE2 structure the $D466_{RshTT200}$ side chain position is not fully resolved but appears to be in range for the formation of a weak salt bridge with $K31_{RShACE2}$(**Fig. S3.12B**). This interaction would not be possible with either *R. affinis* ACE2 alleles as a result of the $K31_{RSh}/N31_{RA9479ACE2/RA787ACE2}$ residue substitution (**Fig. S3.12E,H**). $K31_{RshTT200}$ is also positioned to form hydrogen bonds with the peptide backbone of $Y472_{RshTT200}$. N31

in the model of RA9479 but not RA787 forms hydrogen bonds with $Q475_{RshTT200}$ likely due in part to the presence of R34 in RA787 which sterically restricts the position of the $N31_{RA787}$ side chain (**Fig. SXH**). This is a factor not at play in RA9479 due to the presence of H34 at this position (**Fig. SXf).** However, $R34_{RA787}$ appears to form hydrogen bonds with peptide backbone in $Y472_{RshTT200}$ and to be in range for the formation of a weak salt bridge with $D466_{RshTT200}$ (**Fig. 3.3F-G**). The positioning of $H34_{RA9479ACE2}$ would be unable to form this interaction in our model (**Fig. 3.3F-G**). These differences contribute to the differential binding observed between the *R. affinis* alleles in our binding assays.

The lone *R. sinicus* ACE2 allele that demonstrated binding with RshTT200, demonstrated a high level of conservation with both *R. affinis* alleles. However, it differs notably at $K31_{RSWJ1}$, and $S34_{RSWJ1}$. $K31_{RSWJ1}$ like in our *R. shameli* structure is expected to participate in a salt bridge with $D466_{RshTT200}$ but with more optimal positioning in our models. Additionally, S31 is expected to form hydrogen bonds with $Q475_{RshTT200}$, similar to those expected between $H34_{RA9479}$/$R34_{RA787}$ but not observed with $P34_{Rsh}$. It is anticipated that these differences act in concert to drive differential RshTT200 binding across these ACE2 alleles.

We subsequently evaluated the ability of each of the corresponding full-length (membrane-anchored) ACE2 constructs of *R. affinis*, *R. sinicus* as well as ACE2 orthologs from other Rhinolophus species *R. alcyone*, *R. landeri* and *R. pusillus*, which also harbor sarbecoviruses, to promote entry into HEK293T cells of non-replicating vesicular stomatitis virus (VSV) pseudotyped with RShTT200 S, SARS-CoV-1 S, or SARS-CoV-2 S. Our data show that RShTT200 S VSV can utilize both of the *R. affinis*

ACE2 contact alleles tested as well as the *R. sinicus* ACE2 WJ1 allele as entry receptors (**Fig. 3.3I-L, Fig. S3.14)**, in line with our binding data. None of the other ACE2 orthologs tested supported RShTT200 S VSV cellular entry. As expected, SARS-CoV-1 and SARS-CoV-2 S VSV pseudoviruses exhibited a broader ACE2 tropism within our panel and entered cells transfected with a broader set of *R. sinicus* ACE2s as well as with *R. alcyone* and *R. landeri* ACE2s (**Fig. 2I-L**).

These results show that RshTT200 can utilize multiple bat ACE2s as entry receptors and underscore the ability of *R. affinis,* and *R. alcyone* ACE2 alleles to promote S-mediated entry into cells of several sarbecoviruses (Starr et al. 2022; Roelle et al. 2021). However, it also highlights the relatively limited promiscuity of RshTT200 across our Rhinolophus ACE2, panel compared to SARS-CoV-1 and SARS-CoV-2. Both of which demonstrate particularly wide tropism; a feature likely contributing to their spillover. However, when comparing this tropism to what is observed in other bat borne sarbecoviruses, specifically PRD-0038 (Lee et al. 2023) against an identical ACE2 panel the limited tropism of RshTT200 begins to stand out for its limited tropism.

A likely explanation is the markedly shorter 430-434 loop (SARS-CoV-2 residues 443-451) as deletions in this loop (indel site i) and/or the $455\text{-}491_{\text{SARS-CoV-2}}$ loop (indel site 2) have been correlated with limited tropism and a more specialist interaction with its host ACE2 (Si et al. 2024). While the limited tropism may reduce the risk of spillover for this sarbecovirus the ability of it to utilize one ACE2 allele of *R. sinicus*, and two alleles of *R. affinis* bats, both species with ranges overlapping that of R. shameli suggestive that this virus could spread and evolve potentially increasing spillover risk.

3.6 Broad Tropism of RshTT200 for ACE2 orthologs Rhinolophus Species

Several sarbecoviruses have been identified in Rhinolophus bats in Southeast Asia (Hu et al. 2017; Wang et al. 2022; Ge et al. 2013; Latinne et al. 2020) and it has been shown that co-infections with two or more sarbecoviruses can occur in wild bats (Wang et al. 2022). Moreover, co-roosting of some of these bat species may favor viral transmission, evolution and recombination. Several SARS-CoV-1-related (clade 1a) and SARS-CoV-2-related (clade 1b) sarbecoviruses were identified from *R. sinicus* (e.g. WIV1 (Ge et al. 2013) and WIV-16 (Yang et al. 2015)) and *R. affinis* bats (e.g. RaTG13 (Zhou et al. 2020) and LyRa11 (He et al. 2014), Ra22QT10 (Hassanin et al. 2023)), which share geographical overlap with *R. shameli*. Sequencing data revealed marked bat ACE2 polymorphism which led to the description of eight and two distinct ACE2 alleles for *R. sinicus* and *R. affinis*, respectively, defined based on sarbecovirus RBD contact residues, (Table S4) (Starr et al. 2022; Guo et al. 2020). Notably *R. affinis* ACE2 receptors have been shown to be highly promiscuous receptors for sarbecoviruses ((Starr et al. 2022; Guo et al. 2020; Lee et al. 2023)).

To assess the promiscuity of RshTT200 host receptor usage compared to other bat-borne clade 1a and clade 1b sarbecoviruses, we evaluated binding of a panel of Rhinolophus ACE2 orthologs harboring a human Fc-fusion to immobilized RBDs using BLI (**Fig. S3.11, S3.12**). Our ACE2 panel featured 8 distinct *R. sinicus* ACE2 alleles and 2 *R. affinis* ACE2 alleles selected to represent the diversity of polymorphic variants at ACE2:RBD contact sites. RshTT200 bound three ACE2 alleles: *R. affinis* 9479 (RA9479), *R. affinis* 787 (RA787) and *R. sinicus* WJ1 (RSWJ1) (Fig. 3 **Fig. S3.11**, **Fig. S3.12**). RaTG13 demonstrated narrower binding breadth relative to RshTT200, interacting only with *R. affinis* ACE2 allele 9479 (**Fig. S3.11**, **Fig. S3.12**).

Comparatively, SARS-CoV-1 and SARS-CoV-2 RBDs showed broad ACE2 promiscuity bound seven and five ACE2 alleles respectively. Another bat borne sarbecovirus WIV1, showed similar promiscuity to SARS-CoV-1, reflective of its close evolutionary relationship with SARS-CoV-1 (clade 1a), though the identities of the ACE2 alleles it bound were slightly altered (**Fig. S3.11**, **Fig. S3.12**). Overall, these findings are in line with the observation that several clade 1a sarbecoviruses have been isolated from *R. sinicus* bats and therefore evolved to interact with *R. sinicus* bat ACE2. The polymorphic variation in *R. sinicus* bats has been suggested to result from an evolutionary arms race between this species and the sarbecoviruses they host and thus clade 1a sarbecovirus (like SARS-CoV-1 and WIV1), which have co-evolved within this species demonstrate broader tropism across R. sinicus ACE2 allele variants than clade 1b (like SARS-CoV-2, RshTT200, and RaTG13) that have undergone other selective pressures. To date no clade 1b sarbecoviruses (such as SARS-CoV-2, RatG13, and RshTT200) have been observed in *R. sinicus* bats and have only been isolated from other Rhinolophus species.

Having established which Rhinolophus ACE2 receptors could interact with the RshTT200 RBD, we next evaluated the kinetics and avidity of binding of the natively dimerized RA9479, RA787, and RSWJ1 ACE2 ectodomains to the immobilized RshTT200 RBD. We determined K_{Dapp} of 206 nM, 55.1 nM, 82.5 nM, respectively (**Fig. 3.3A-C**). Interestingly, unlike other sarbecoviruses in our panel, the RshTT200 RBD demonstrated enhanced binding to the RA787 allele, relative to RA9479 (**Fig. S3.11**, **Fig. 3.3A-C**). These two alleles differ only by two mutations at the binding interface: $H34_{RA9479}/R34_{RA787}$ and $D38_{RA9479}/E38_{RA787}$ (**Fig. 3.3A-B,D**).

To understand the molecular basis of these differences we used Alphafold 3 to predict the structure of these ACE2 alleles in complex with the RshTT200 RBD (**Fig. 3.3F-G**, **Fig. S3.13D-I**). Like $E38_{RShACE2}$ in our RshTT200 RBD: *R. shameli* ACE2 structure (**Fig. S3.13A**), $E38_{RA787ACE2}$ is modeled in an orientation that would enable van der Waals interactions with Y478 (**Fig. S3.13G**). This differs from $D38_{RA9479ACE2}$ with its shorter side chain which is predicted to have less optimal positioning for forming such interactions (**Fig. S3.13D**). In our RshTT200 RBD *R. shameli* ACE2 structure the $D466_{RshTT200}$ side chain position is not fully resolved but appears to be in range for the formation of a weak salt bridge with $K31_{RShACE2}$(**Fig. S3.13B**). This interaction would not be possible with either *R. affinis* ACE2 alleles as a result of the $K31_{RSh}$/$N31_{RA9479ACE2/RA787ACE2}$ residue substitution (**Fig. S3.13E,3.H**). $K31_{RshTT200}$ is also positioned to form hydrogen bonds with the peptide backbone of $Y472_{RshTT200}$. N31 in the model of RA9479 but not RA787 forms hydrogen bonds with $Q475_{RshTT200}$ likely due in part to the presence of R34 in RA787 which sterically restricts the position of the $N31_{RA787}$ side chain (Fig. SXH). This is a factor not at play in RA9479 due to the presence of H34 at this position (Fig. SX. However, $R34_{RA787}$ appears to form hydrogen bonds with peptide backbone in $Y472_{RshTT200}$ and to be in range for the formation of a weak salt bridge with $D466_{RshTT200}$ (**Fig. 3.3F-G**). The positioning of $H34_{RA9479ACE2}$ would be unable to form this interaction in our model (**Fig. 3.3F-G**). These differences contribute to the differential binding observed between the *R. affinis* alleles in our binding assays.

The lone *R. sinicus* ACE2 allele that demonstrated binding with RshTT200, demonstrated a high level of conservation with both *R. affinis* alleles. However, it differs

notably at $K31_{RSWJ1}$, and $S34_{RSWJ1}$. $K31_{RSWJ1}$ like in our *R. shameli* structure is expected to participate in a salt bridge with $D466_{RshTT200}$ but with more optimal positioning in our models. Additionally, S31 is expected to form hydrogen bonds with $Q475_{RshTT200}$, similar to those expected between $H34_{RA9479}$/$R34_{RA787}$ but not observed with $P34_{Rsh}$. It is anticipated that these differences act in concert to drive differential RshTT200 binding across these ACE2 alleles.

We subsequently evaluated the ability of each of the corresponding full-length (membrane-anchored) ACE2 constructs of *R. affinis*, *R. sinicus* as well as ACE2 orthologs from other Rhinolophus species *R. alcyone*, *R. landeri* and *R. pusillus*, which also harbor sarbecoviruses, to promote entry into HEK293T cells of non-replicating vesicular stomatitis virus (VSV) pseudotyped with RShTT200 S, SARS-CoV-1 S, or SARS-CoV-2 S. Our data show that RShTT200 S VSV can utilize both of the *R. affinis* ACE2 contact alleles tested as well as the *R. sinicus* ACE2 WJ1 allele as entry receptors (**Fig. 3.2, Supplemental Fig. S3.11,S3.12)**, in line with our binding data. None of the other ACE2 orthologs tested supported RShTT200 S VSV cellular entry. As expected, SARS-CoV-1 and SARS-CoV-2 S VSV pseudoviruses exhibited a broader ACE2 tropism within our panel and entered cells transfected with a broader set of *R. sinicus* ACE2s as well as with *R. alcyone* and *R. landeri* ACE2s (**Fig. 3.2B-C**).

These results show that RshTT200 can utilize multiple bat ACE2s as entry receptors and underscore the ability of *R. affinis,* and *R. alcyone* ACE2 alleles to promote S-mediated entry into cells of several sarbecoviruses (Starr et al. 2022; Roelle et al. 2021). However. it also highlights the relatively limited promiscuity of RshTT200 across our Rhinolophus ACE2, panel compared to SARS-CoV-1 and SARS-CoV-2. Both of

which demonstrate particularly wide tropism; a feature likely contributing to their spillover. However, when comparing this tropism to what is observed in other bat borne sarbecoviruses, specifically PRD-0038 (Lee et al. 2023) against an identical ACE2 panel the limited tropism of RshTT200 begins to stand out for its limited tropism.

A likely explanation is the markedly shorter 430-434 loop (SARS-CoV-2 residues 443-451) as deletions in this loop (indel site i) and/or the 455-491$_{\text{SARS-CoV-2}}$ loop (indel site 2) which have been correlated with limited tropism and a more specialist interaction with its host ACE2 (Si et al. 2024). While the limited tropism may reduce the risk of spillover for this sarbecovirus the ability of it to utilize one ACE2 allele of *R. sinicus*, and two alleles of *R. affinis* bats, both species with ranges overlapping that of R. shameli suggestive that this virus could spread and evolve potentially increasing spillover risk.

3.6 Deep Mutational Scanning Identifies mutation of concern

While Rhinolophus bats act as a natural reservoir for sarbecoviruses, past spillovers of coronaviruses have resulted from transmission into an intermediate host such as palm civets in SARS-CoV-1 and camels in MERS-CoV. To assess the ability of RshTT200 to use ACE2 orthologs from other species, we assessed entry of RshTT200 S pseudotyped VSV particles into HEK293T cells transfected with a panel of 24 ACE2 orthologs from domesticated and wildlife species. We anticipated limited tropism due to the loss of key electrostatic interactions resulting from the shortened 430-434 loop than the equivalent SARS-CoV-2 RBD loop (residues 443-451) at indel site i (**Fig. 3.1J**) due to previous studies demonstrating that deletions at this site correlates with limited tropism (Si et al. 2024). The relatively limited tropism across our Rhinolophus ACE2

panel (**Fig. 3.3I-L**) appeared to confirm that assumption. However, RshTT200 S VSV particles facilitated efficient entry in all but 7 ACE2 alleles (marmoset, chicken, duck, snake, mouse, treeshrew, horse and whale) in our panel (**Fig. 3.4**). Although serological evidence of SARS-CoV-2 infection has been documented in horses (Lawton et al. 2022), horse ACE2 did not support RshTT200 S VSV entry into cells (i.e. entry was equivalent or less than that of mock VSV - VSV particles harboring no S). Marmoset, mouse, duck and chicken ACE2 have been previously demonstrated to be resistant to SARS-CoV-2 Wuhan (Liu et al. 2021; Wu et al. 2020; Suarez et al. 2020), however the Omicron variant gained mouse ACE2 binding and functional entry (McCallum et al. 2022; Cameroni et al. 2022) demonstrating that current tropism level may be modulated over time.

Notably the ACE2 of pangolins, from which several related clade 1b viruses have been isolated (Lam et al. 2020), and palm civets, which acted as intermediate hosts in the spillover of SARS-CoV-1 (Wang et al. 2005; Song et al. 2005) both facilitate efficient entry by RshTT200 S VSV (**Fig. 3.4**). Several ACE2 receptors in our panel have evidence for the capacity to be infected by SARS-CoV-2. Specifically, natural infections of SARS-CoV-2 or serological evidence of them have been found in cats (Ozer et al. 2022), dogs (Sit et al. 2020), rabbits (Fritz et al. 2022), mink (Oreshkova et al. 2020), camels (El Masry et al. 2023), tigers (Grome et al. 2022), and ferrets (Račnik et al. 2021). White tailed deer have been demonstrated to be particularly susceptible to SARS-CoV-2 infection (Hale et al. 2022; Chandler et al. 2021). Serological evidence of SARS-CoV-2 infection has been identified in red deer and fallow deer (Encinas et al. 2023). In our panel red deer ACE2 are likewise functional receptors to RshTT200 (**Fig.**

3.4). Experimental infections by SARS-CoV-2 in several other species in our panel have been confirmed including syrian hamster (Sia et al. 2020), cows (Ulrich et al. 2020), raccoon dogs (Freuling et al. 2020), rhesus monkeys (Deere et al. 2021), and crab eating macaques (Stammes et al. 2022).

These results demonstrate the challenges in predicting spillover potential of sarbecoviruses. As the shortened loop in indel site i of the RBM has been correlated with reduced tropism, RshTT200 could have been predicted to have a limited host range. As such, findings of such breadth of cross mammalian species ACE2 tropism were unexpected. Furthermore pigs, despite efficient ACE2 binding of SARS-CoV-2 (Wu et al. 2020) are not susceptible to SARS-CoV-2 infection (Schlottau et al. 2020) demonstrating that ACE2 binding and S-pseudotyped VSV cellular entry does not provide the full picture of spillover risk. It is notable most of the ACE2s identified in our panel as facilitating cellular entry of RshTT200 S have previously shown susceptibility to SARS-CoV-2 infections; highlighting many of these species are susceptible to sarbecovirus infections and could act as intermediate hosts in a future spillover.

3.7 Broad Tropism of RshTT200 for ACE2 orthologs from Domesticated and Wildlife Species

Previous reports have indicated that RshTT200 currently is unable to infect cells transfected with human ACE2 (Delaune et al. 2021), despite its ability to bind human ACE2, albeit weakly (Hu et al. 2023). To assess the basis of this inability to infect human cells we first tested RshTT200 RBD binding with human ACE2-Fc and compared it to other closely related bat born sarbecoviruses through BLI. This revealed both RatG13

and RshTT200 display markedly dampened human ACE2 binding compared to SARS-CoV-1, SARS-CoV-2, and WIV1 (**Fig.**), in line with previous reports. However, RaTG13 differs from RshTT200 in that it can facilitate human ACE2 dependent cellular entry (Liu et al. 2021). To further characterize their relative association with human ACE2, we next measured binding kinetics between the immobilized RBD and the natively dimerized human ACE2 by BLI, revealing a approximate the KD_{app} between human ACE2 and the RaTG13 RBD of 243nM (**Fig. S3.17, Table S3.3)** approximately 10 fold stronger than the KD_{app} between human ACE2 and the RshTT200 RBD (~4μM) (**Fig. S3.6; Table S3.3**). This difference in affinity may explain why the RaTG13 S but not RshTT200 S pseudotyped particles can efficiently enter cells with human ACE2.

Furthermore, the ability of RshTT200 S VSV to mediate entry with bat ACE2 receptors for which it has strong associations under conditions of avidity as seen with both *R. affinis* alleles demonstrates that RshTT200 S is capable of facilitating cellular entry when its RBD has high affinity binding with a receptor (**Fig. 3.2B, 32D**, **Fig. 3.3A-C, Fig. 3.3I-K, Table S3.3**). We expected then that mutations that enhanced the affinity of the RshTT200 RBD would likewise enable human ACE2 dependent cellular entry. To identify mutations enabling enhanced human ACE2 binding, we utilize a high-throughput RBD-yeast display system for deep mutational scanning of the RshTT200 RBD to assess the impact of every single amino acid mutation across the RBD on ACE2 binding and RBD expression. One mutation across our deep mutational scanning panel stood out Y453F as both the strongest binding enhancement as well as having been observed in the course of the SARS-CoV-2 pandemic in mink adapted variants. This finding was validated by BLI, with

RshTT200 453F and human ACE2 with a KD_{app} of 115nM a significant improvement over the 4μM KD_{app} observed between human ACE2 and RshTT200.

To assess whether this mutation was sufficient to enable human ACE2- dependent entry of RshTT200, we tested RshTT200 Y453F S pseudotyped VSV particles for entry in human ACE2 and demonstrated that functional entry of RshTT200 453 S VSV was observed. The mutation of tyrosine to phenylalanine involves only a single nucleotide mutation with the RshTT200 genome. The remarkable rate of emerging variants of concern during the SARS-CoV-2 pandemic illustrates the ability of sarbecoviruses to mutate and alter their tropism demonstrated the single nucleotide mutation required for Y453F in RshTT200 is well within the scope of the evolutionary potential of sarbecoviruses.

3.8 SARS-CoV-2 Wu vaccinated polyclonal sera neutralization potency against RshTT200 Y453F

In the event of the introduction of the RshTT200 Y453F mutation, we wondered if SARS-CoV-2 vaccination would provide sufficient protection to generate neutralizing antibodies effective against RshTT200 Y453F. To test this hypothesis, we performed neutralization assays with a panel of sera from a cohort of 14 patients vaccinated either 3x with Moderna or Pfizer, or 2x with Johnson and Johnson. Across all members of this cohort, neutralization titers against RshTT200 453F S VSV were equivalent to or better than those against SARS-CoV-2 Wuhan (**Fig. 3.5; S3.18, Table S3.4**). Suggestive that current vaccinations would provide protection against this virus.

We then asked whether Y453F would be a mutation that would be favored in RshTT200 host species *R. shameli* or in other species of Rhinolophus that share

territories with *R. shameli* like *R. affinis* and *R. sinicus* bats for which RshTT200 can utilize the ACE2 receptor for entry. To answer this question we determined and assessed KD_{app} of RshTT200 WT, RshTT200 453F and SARS-CoV-2 for the natively dimerized ACE2 ectodomains of *R. shameli, R. affinis* isolate 9479, *R. affinis* isolate 787, and *R. sinicus* WJ1. This revealed approximately equivalent KD_{app} for RshTT200 WT, 453F in R. affinis 9479, R. affinis 787, R. sinicus WJ1 (**Fig. X**). In R. shameli RshTT200 453F appears to bind modestly worse than RshTT200 WT, but still binds in the nanomolar affinity. For human ACE2, the Y453F mutation resulted in almost a 2-fold shift in binding. Fascinatingly, SARS-CoV-2 bound R. shameli ACE2 worse than RshTT200 WT bound human ACE2. However, we and others have observed SARS-CoV-2 S VSV facilitate entry in cells expressing membrane anchored *R. shameli* ACE2 (**Fig. 3.2D**) ((Delaune et al. 2021). Together with the observation that RshTT200 WT fails to enter in cells expressing human ACE2 (**Fig. 3.7A**) is suggestive that RshTT200's affinity for human ACE2 is not the only barrier for human ACE2 dependent cellular entry.

3.9 RBD conformational regulation modulates receptor tropism

Based on this apparent inconsistency, whereby low affinity interaction with R. shameli ACE2 results in cell entry of SARS-CoV-2, but roughly equivalent low affinity interactions between the RshTT200 RBD and human ACE2 do not we hypothesized that while affinity was a factor it is not the only barrier to successful cellular entry of RshTT200 S with human ACE2.

We began to suspect that conformational dynamics of the RBD were also at play; this was driven largely by the observation of only closed RBDs during 3D classification of the RshTT200 S. Coronaviruses rely on conformational masking to regulate exposure of the

receptor-binding domains dynamically, thereby balancing the necessity to engage host receptors and to limit recognition by neutralizing antibodies (Walls et al. 2016; Pallesen et al. 2017; Yuan et al. 2017; Walls et al. 2019; Tortorici et al. 2022). Whereas open and closed RBDs were observed in S ectodomain trimer structures of highly pathogenic coronaviruses (SARS-CoV-1, SARS-CoV-2 and MERS-CoV), all other S cryoEM structures feature exclusively closed RBDs, as is the case with RshTT200 S.

These observations contrast with the ability of all coronaviruses to enter host cells, suggesting that at least a fraction of the RBDs on each S trimer sample open conformations whether at the surface of authentic viruses or pseudotyped viruses. However, some coronaviruses may require binding with a co-receptor for RBD opening to occur. Notably HKU1, which by cryo-EM was shown to exist in a closed conformational state (Kirchdoerfer et al. 2016) has recently been shown to adopt open RBD conformations upon binding with sialic acid (Pronker et al. 2023). It seems likely then that coronaviruses revealing exclusively closed RBDs during structural reconstruction (RaTG13, Pangolin GX, OC43) (Zhang et al. 2021; Tortorici et al. 2019) either have small fractions of total spike adopting open RBD conformation or may require initial contacts with their receptor or a co-receptor to initiate RBD opening and enable full receptor engagement.

RshTT200 RBD binds the natively dimerized R. shameli, R. affinis alleles 9479 and 787, and R. sinicus ACE2 allele WJ1 ACE2 (**Fig. 3.2, 3.3, Table 3.3**) with nanomolar affinity suggesting that a co-receptor is not required for receptor engagement for RshTT200 S, which is congruent with its ability to enter with transfection of either *R. shameli* or other bat ACE2.Notably however the weak binding observed in this assays

are only at concentrations that are expected to exceed what would be expected physiologically, which may help reconcile the apparent incompatibility of the concept that the RshTT200 S ectodomain is both capable of binding human ACE2 as detected by BLI but also its RBD conformation is reducing it's spillover potential. While only closed conformations were revealed for the RshTT200 ectodomain, RshTT200 S pseudotyped VSV entered HEK293T cells transiently transfected with both of the *R. affinis* ACE2 contact alleles, the *R. sinicus* ACE2 'WJ1' allele for which tighter binding has been demonstrated (**Fig. 3.3**). Collectively, these findings suggest that limited RBD exposure and weak 1:1 binding affinity could act in concert to hinder RshTT200 S-mediated entry into human ACE2-expressing cells.

To evaluate further evaluate if the lack of RShTT200 S VSV entry into human ACE2-expressing HEK293T cells was a reflection of limited RBD exposure, and the resultant damping of ACE2 binding or of other factors (RBM accessibility or S protease processing), we generated chimera S VSV pseudotypes harboring the RShTT RBD in the SARS-CoV-1 S and SARS-CoV-2 S backgrounds as well as the SARS-CoV-1 RBD or the SARS-CoV-2 RBD in the RshTT200 S background and normalized them by S before assessing entry in HEK293T cells with human ACE2. RshTT200 S VSV chimeras harboring the SARS-CoV-1 or the SARS-CoV-2 RBDs promoted efficient entry into HEK293T cells with stable expression of human ACE2, indicating that affinity is a potentially contributing factor to RShTT200 S inability to utilize human ACE2. These data also confirmed that the RShTT200 S_2 subunit is fusion-competent and therefore functional, as previously demonstrated with Rhinolophus ACE2 mediated cellular entry (**Fig. 3.2**). We further found that swapping the RShTT200 RBD in the SARS-CoV-1 S

VSV chimera, but not in the SARS-CoV-2 S VSV chimera, enabled entry into human ACE2-expressing HEK293T cells.

Collectively, these data suggest that besides the intrinsic RBD affinity for human ACE2, conformational differences in these S trimers, such as differences in RBD opening propensity, altered prefusion S metastability, or distinct number of S incorporated per virion account for these alterations of entry. While we observed differential S incorporation into VSV pseudotypes particles for this assay (**Fig. S3.14**), the S incorporation if RshTT200 S and RshTT200 RBD in SARS-CoV-1 S were virtually identical ruling this out as a potential cause. These results likely reflect that the RShTT200 RBD-containing SARS-CoV-1 S chimera more readily accesses open RBD conformations compared to wildtype RShTT200 S and the RShTT200 RBD-harboring SARS-CoV-2 S chimera. It seems likely due to the high degree of conservation that the interactions maintaining a closed RBD conformation are conserved in the RshTT200 RBD in the SARS-CoV-2, but due to the higher degree of divergence, these interactions are lost in the RshTT200 RBD in the SARS-CoV-1 chimera enabling more open RBD conformations. Notably, human infecting SARS-CoV-1, and SARS-CoV-2 VSV PV showed the highest levels of S incorporation, which along with the broad tropism demonstrated by these S glycoproteins (**Fig. 3.3**), may suggest that evolution for these virus have largely favored efficient mechanisms for infecting a wide variety of hosts, whereas RshTT200 appears to favor features that make it a “specialist virus” uniquely adapted to infect its host species at cost of a loss of broader tropism.

To further validate that the chimeric RshTT200 RBD in SARS-CoV-1 S was capable of adapting RBD open conformations we generated the RshTT200 RBD in SARS-CoV-1

S ectodomain. While the hexapro mutation (Hsieh et al. 2020) used initially for the RshTT200 S were insufficient for expression and prefusion stabilization of the RshTT200 RBD in SARS-CoV-1 chimera, the addition of disulfide bonds to this construct rescued expression of a prefusion stabilized S. We then characterized this construct using cryoEM. In addition, we characterized a RshTT200 S construct containing the same additional disulfide bonds by cryoEM to ensure changes in RBD conformation could be attributed to the S background and not the stabilization mutations. Open and closed RBD conformations for the RshTT200 RBD in SARS-CoV-1 S ectodomain were observed during 3D classification, but were absent in the RshTT200 disulfide bond stabilized S control. Together this demonstrates that the prevalence of RBD down conformation in the RshTT200 is a factor in its inability to enter human cells through the use of human ACE2.

However, were the RShTT200 S to acquire mutations that enhance RBD opening it may become a public health threat. One such mutation, D614G, has been demonstrated to allow for enhanced entry and a more open RBD conformation in the SAR-CoV-2 S. This mutation became a prevalent across SARS-CoV-2 variants throughout the pandemic. To explore whether this mutation could play a similar role in the RshTT200 the SARS-CoV-2 D614, RshTT200 and RshTT200 D614G S pseudotyped VSV particles were used to infect human ACE2 HEK293T cells. However no entry was observed for human ACE2 293T cells for RshTT200 WT or the RshTT200 D614G mutant (**Fig, S20**), and this mutation resulted in roughly equivalent cell entry when higher affinity receptors were transfected into HEK293T cells (R. shameli, R. sinicus WJ1) (**Fig, S20**) suggesting that the impact of the D614G in the RshTT200 S has little influence on entry.

Differences in proteolytic cleavage of the RShTT200 S, the SARS-CoV-1 S, and the SARS-CoV-2 S may also play a role in the differences in RBD binding and cellular entry seen in the chimeric S utilized, but no differences in proteolytic processing of the RshTT200 WT, RshTT200 RBD in SARS-CoV-1, or the SARS-CoV-1 S were detected in the pseudovirus (**Fig. S3.14**) prior to infection. We wondered if perhaps like SARS-CoV-2, RshTT200 could undergo processing by TMPRSS2 and if so, if such processing could alter the RBD conformational dynamics and in doing so enable human ACE2 utilization. While, as expected, entry was altered for SARS-CoV-2, when TMPRSS2 was co-expressed with human ACE2, no changes in RshTT200 entry were observed.

This also illustrates some of the complexities of predictions of spillover potential, it can neither be predicted by sequence or receptor affinity alone, but rather a broader approach is required to understand the interplay of the many factors that determine receptor utilization including receptor association to include both receptor affinity and conformational accessibility. The latter may also be driven in part by proteolytic processing.

3.10 Chapter Discussion

The discovery of RShTT200, broadened our understanding of the geographical range and the timeline of when SARS-CoV-2 like (clade 1B) sarbecoviruses were circulating in nature. Functional receptor usage of *R. affinis* and *R. sinicus* ACE2 alleles demonstrate two potential pathways RShTT200 evolved and illuminates potential routes of spread. Furthermore, the differences in binding observed between *R. affinis* and *R. sinicus* ACE2 interface alleles highlights the role of polymorphism of ACE2 in bat

populations in determining potential spread of sarbecoviruses among Rhinolophus bat populations. This point highlights the importance of receptor sequencing in bats when screening for viruses as well as the importance of improved wildlife sequencing, through which we may be able to trace the evolutionary history of sarbecoviruses in bats and better understand their evolution and spread among bat species.

Due to the absence of paired sequencing of sarbecoviruses and receptors, bat ACE2 receptor polymorphism likely exists beyond the panel tested as this is a largely unexplored space. Even with this caveat, the capacity of RShTT200 S to utilize the ACE2 of several geographically dispersed species gives insight into potential tropism across Rhinolophus bats. This combination of range overlap, tropism, and ACE2 polymorphism provide sarbecoviruses many opportunities to spread geographically and evolve allowing for the introduction of ever-changing traits, which likely plays an important role in the adaptation of new traits including the transition to human ACE2 usage. We identified in our panel the ability of RshTT200 to enter with *R. affinis* ACE2 and suggested that this could enable cross species transmission. In fact, in November 2022, Ra22QT10 was isolated from *R. affinis* bats in Vietnam. This virus has 99% sequence identify with RshTT200 across its genome, 98.9% amino acid sequence identity across its S and 100% amino acid sequence identity across the RBD, showing it is closely related to RshTT200 and giving evidence of these types of cross species transmission in Rhinolophus bats. This highlights that this viral family is not isolated in Cambodia but has foothold in neighboring countries. Additionally, this gives us an idea of the timescale of evolution of these viruses in bats, RshTT00 was isolated 2010, 12 years later Ra22QT10 was isolated in another species with only 1% of the genome

changed. This is very different from the rate of evolution observed throughout the SARS-CoV-2 pandemic and highlights how evolution is very different in these native reservoirs, than it is once sarbecoviruses spillover and face new selective pressure in humans.

In our structural determination of the RShTT200 S we identified only RBD “down” conformations. The presence of predominantly down conformations has been an increasingly common trend in sarbecoviruses that have been isolated from animal reservoirs (Zang et al. 2021, Worbel et al 2021) and may be an indicator of reduced spillover potential as the RBD up conformation is necessary for advantageous contacts with the host receptor (Zhang et al 2021 ,Worbel et al 2021). Therefore, the propensity to adopt open conformations is expected to have implications on transmission and spillover potential.

These findings concur with observations that all four human endemic coronavirus S trimers (HCoV-NL63, HCoV-229E, HCoV-HKU1 and HCoV-OC43). It is noteworthy however that HKU1, OC43, 229E and NL63 are all infectious in humans but have no “up” RBD conformations identified by cryo-EM. More broadly, S glycoprotein structures of all three highly pathogenic coronaviruses determined to date, namely SARS-CoV-2, SARS-CoV and MERS-CoV, are characterized by the presence of both open and closed RBDs (Walls et al 2020, Gui et al. 2017, Kirchdoerfer et al 2018, Pallesen et al 2017, Song et al 2018, Walls et al 2019, Yuan et al 2017).

We demonstrate that the limiting factor in human ACE2 utilization by the RShTT200 S is likely both its low association with human ACE2 and the prevalence of a closed

RBD conformation that limits its capacity to bind human ACE2. Therefore, it is likely that either increases in binding affinity or increases in RBD openness will improve cellular entry. As the primary barrier to RShTT200 human ACE2 utilization appears to be RBD conformation at this time, then mutations that open the RBD may allow the RShTT200 S to facilitate functional entry though human ACE2, such as what was observed with the RShTT200 RBD SARS-CoV-1 S chimera. There are several possible explanations for this, including RShTT RBD contacts that maintain closed conformation being disrupted in the RShTT RBD- SARS-CoV-1 S chimera and alternatively, differences in proteolytic processing of the SARS–CoV-1 S and the RShTT200 S during cellular entry.

It is curious that despite the apparent low abundance of open conformations, this spike can readily mediate entry both *R. affinis* ACE2 alleles (9479 and 787) and with *R. sinicus* ACE2 allele WJ1. It appears in these cases that the higher degree of association observed between the RShTT RBD and these ACE2s is sufficient to overcome the limitations imposed by the low abundance of open RBD. Due to the complex balance between these features, proactive pandemic preparedness and screening of coronaviruses in wildlife populations should study not only cellular entry in human cell lines but also assess the capacity to gain cellular entry through enhancement of receptor affinity and/or shifting RBD accessibility through mutations that promote RBD “up” conformations. It is however important in such work to utilize safe systems, such as with our VSV pseudovirus as the potential for human transmission may be increased with such mutations. We have identified one such mutation equivalent to the SARS-CoV-2 Y453F that enables efficient human ACE2 utilization.

The finding that RaTG13, which enters with human ACE2, has similar affinity for the human as the RshTT200 Y453F mutant clearly demonstrates that at some threshold affinity for a receptor can overcome closed RBD conformations (as RaTG13 like RshTT200 was observed exclusively in closed RBD conformations by cryo-EM). However, the capacity of the RshTT200 RBD in the SARS-CoV-1 chimeric S to have function cellular entry with human ACE2 and the subsequent finding this construct can be observed to adopt open RBD conformations demonstrates that open RBD conformations lowers the affinity needed to enable cellular entry. It therefore makes sense that all highly pathogenic CoV discovered to date have readily adopted open RBD conformations, which likely reduced the buried giving more room for adaptations. In this way closed RBD conformations may be considered a barrier to spillover, whereas open ones are a hallmark of them.

Furthermore, while the use of chimeric S glycoproteins in research studies offer cost savings, this work highlights one way that experimental design and interpretation of results with chimeric S glycoproteins should be done cautiously. The capacity of chimeric RShTT RBD in the SARS-CoV-1 background to gain functional entry, combined with the increase binding of human ACE2 for this chimera indicates a potential “RBD mismatch” which presumably disrupts interactions between the RBD and other components of the spike. An alternative pathway by which this may be occurring is through differences in proteolysis between the SARS-CoV-1 and RShTT200 S, though the entry observed for the SARS-CoV-1/SARS-CoV-2 RBD chimeras. But either way the use of the chimera fundamentally alters our understanding of RshTT200 tropism compared to the full-length RshTT200 WT S.

Broad species tropism of RshTT200 was observed across a panel of animal ACE2, and of particular note is its ability to enter cells with mink ACE2. Minks in particular have been repeatedly infected with SARS-CoV-2 during the course of the pandemic (Oreshkova et al. 2020; Jahid et al. 2024; Lu et al. 2021) and several mink adapted SARS-CoV-2 variants have been isolated from mink farms both in mink and in the farmers that work with them (Lu et al. 2021). Several of which have borne SARS-CoV-2 Y453F mutations. In light of the favorability of the RshTT200 Y453F mutation for enabling successful human ACE2 mediated cellular entry, the ability of RshTT200 to mediate cellular entry in minks, this is of particular concern. As Cambodia, the origin of this virus, is a major supplier of fur farming including minks, the potential for spread of RshTT200 or closely related viruses into minks where they are more likely to pick up mink adapted changes, could put this virus at higher risk of spillover.

We demonstrate that, RBD targeting, monoclonal antibodies targeting site 1A S2X259 and S2X35 have neutralizing potency against RShTT in SARS-CoV-1 S VSV pseudotyped virus particles. These antibodies would prove useful should RShTT200 or a closely related virus cross the species barrier. Furthermore, these tools are likely to be useful against a broad range of bat origin sarbecovirus and should be screened to assess how broadly they would be effective and may act as a first line of defense against casualties during future spillover events. Had such tools been widely available at the onset of the SARS-CoV-2 pandemic the loss of life may have been reduced. In addition, we show that RshTT200 453F S VSV particles are potently neutralized by sera from 2-3x vaccinated patients demonstrating that while RshTT200 could spillover, with

the development of ACE2 affinity enhancing mutations or RBD opening mutations, that current vaccination strategies would be effective against it if it did.

3.11 Methods

Transient Expression and Purification of biotinylated RBDs

The DNA encoding RShTT (182/200) RBD was synthesized with codon optimization for mammalian expression by GenScript. The RShTT RBD (residues 315-513) was cloned in CMVR vector with a N-terminal signal peptide tag (MGILPSPGMPALLSLVSLLSVLLMGCVA) and a C-terminal 8x His tag for purification (HHHHHHHH). RBDs were produced in Expi293F cells in Gibco Expi293 Expression Medium at 37°C in a humidified 8% CO2 incubator rotating at 130 rpm. The cultures were transfected using PEI-25K with cells grown to a density of 3 million cells per mL and cultivated for 3-5 days. Proteins were purified from clarified supernatants using a 1mL HisTrap HP affinity column (Cytiva), concentrated, and biotinylated with a BirA biotinylation kit. Storage at -80 °C in 1x PBS was utilized prior to use.

RShTT200 S Ectodomain Expression and Purification

The DNA encoding RShTT200 Spike (S) glycoprotein (GISAID: EPI_ISL_852605) was synthesized with codon optimization for mammalian expression. The ectodomain of RShTT200 (residues 14-1189) was cloned in the CMVR vector with a N-terminal signal peptide tag (MGILPSPGMPALLSLVSLLSVLLMGCVA). A C-terminal trimerization tag (SGRENLYFQGGGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLG) followed by 8x His tag (HHHHHHHH) were also incorporated. RShTT200 S were produced in Expi293F cells in Gibco Expi293 Expression Medium at 37°C in a humidified 8% CO2 incubator

rotating at 130 rpm. The cultures were transfected using PEI-25K with cells grown to a density of 3 million cells per mL and cultivated for 5-6 days. Proteins were purified from clarified supernatants using a 1mL HisTrap HP affinity column (Cytiva), concentrated, and snap frozen in 1x TBS (50 mM Tris-HCL, pH 8.0, 150 mM Sodium Chloride) prior to storage at -80 °C. RShTT200 S containing 2P (K832P, F833P) (Pallesen et al. 2017) or Hexapro mutations (D596P, I812P, Y815P, K832P, F833P, T837P) (Hsieh et al. 2020) were attempted, however only the Hexapro mutation was capable of stabilizing the S in the prefusion conformation (**Supplemental 1A**).

RShTT200 S Data Collection

Three microliters of RShTT200 S Hexapro were loaded onto freshly glow discharged R 2/2 UltrAuFoil grids (200 mesh). Double blotting of RShTT200 S at 1.3 mg/ml was performed prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 1 and 6.5–7.5 sec blot time at 100% humidity and 22°C.

Data were acquired using Leginon to control a Glacios transmission electron microscope operated at 200 kV and equipped with a Gatan K2 summit direct detector. and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Each movie was acquired with a pixel size of 1.16 Å. In a single session ~ 1100 micrographs were collected with a defocus range between -0.5 and -1.2 uM. WARP was used for movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking, and extraction were carried out using Warp.

Reference free 2D classification were performed using CryoSPARC to select particle images. The selected particles were subjected to ab initio reconstruction followed by

non-uniform refinement in cryoSPARC to generate an initial model. The initial model was

Cryo-EM Model Building and Refinement:

UCSF Chimera and Coot were utilized to fit atomic models into the cryo-EM map. The RShTT200 S model was refined and relaxed using Rosetta with sharpened maps.

Bio-layer Interferometry analysis of RBD/ACE2 associations

BLI assays were performed as previously described. Briefly RBD was immobilized on a SA biosensor tip (Forte Bio) and RBD load was normalized to 1 nm. The tip was then dipped into 10x KB (Forte Bio) to establish a baseline signal for 60-120 sec. To assess for association, the tip was dipped into 1 μM ACE2 Fc for 300 seconds and dipped back into a 10X KB buffer to assess dissociation of ACE2 from RBD. Measurements shown are measurements of avidity not affinity but demonstrate relative association with ACE2. Biolayer interferometry experiments were repeated in biological replicates of biotinylated RBD having been generated on different days.

3.12 Figures & Tables

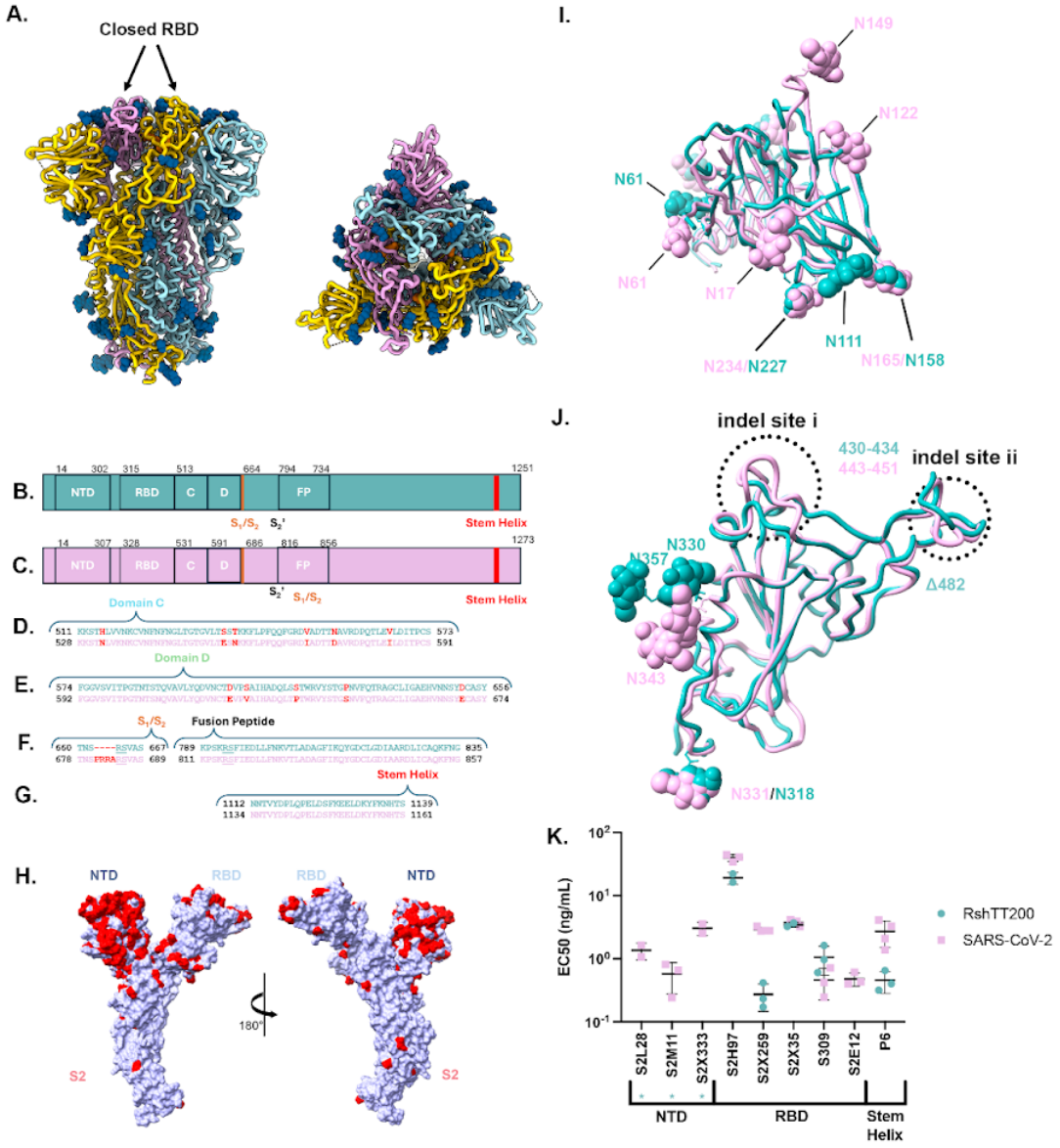
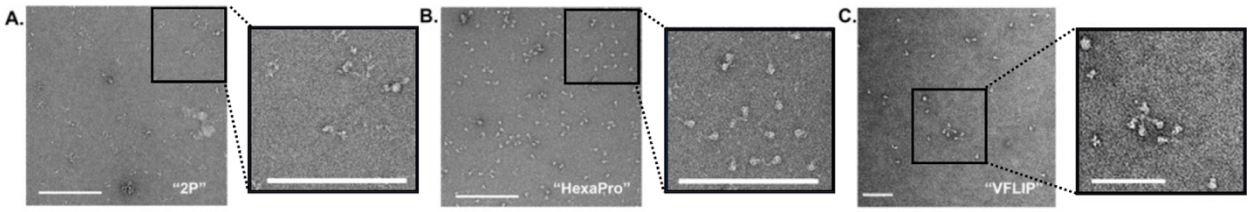


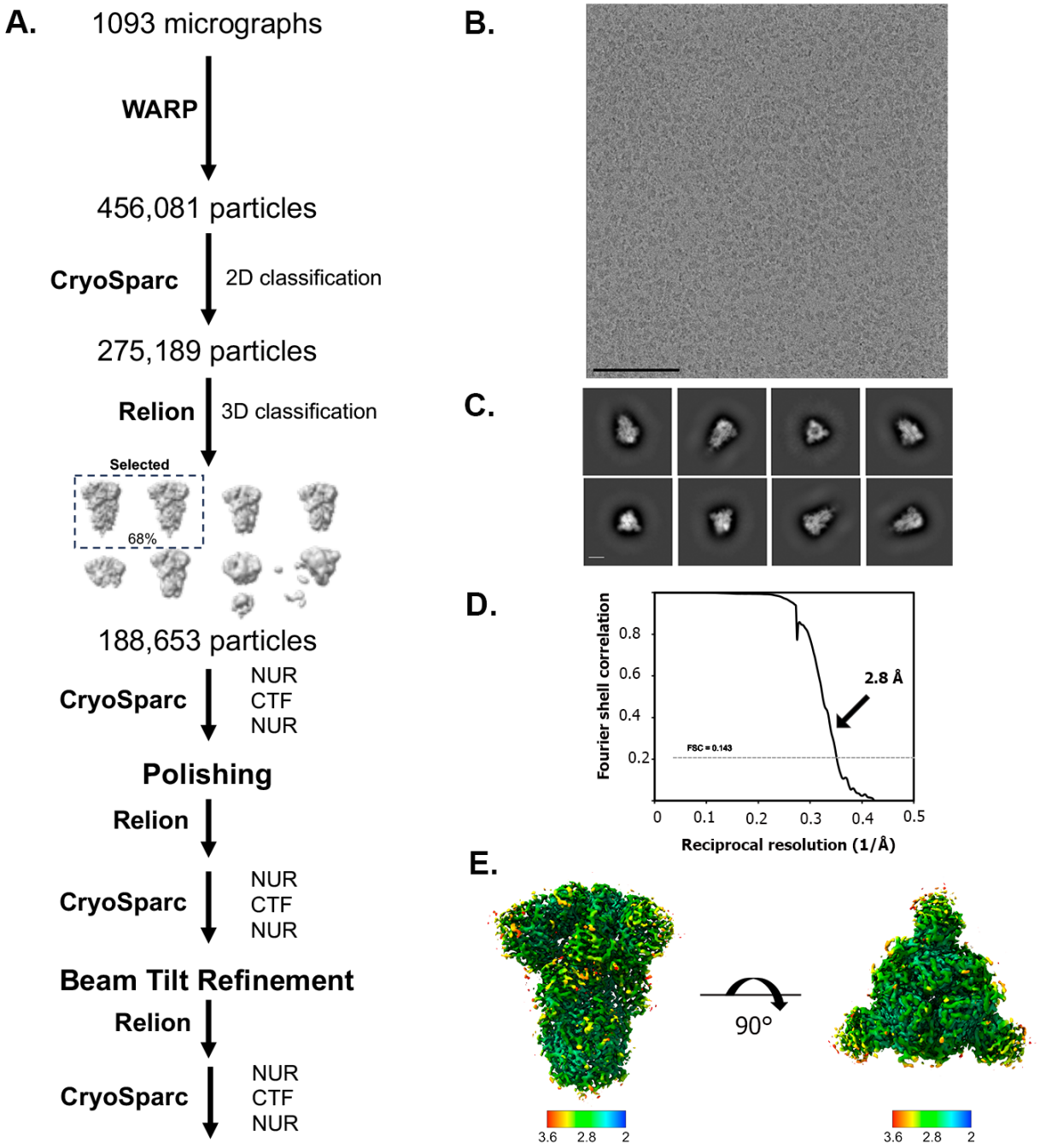
Figure 3.1: Architecture of the RshTT200 S glycoprotein trimer.

Ribbon diagram of the RShTT200 S trimer viewed along two orthogonal orientations, side (A, left) and top (A, right). Domain architecture and boundaries of RshTT200 (B) and SARS-CoV-2 (C) spike glycoproteins with sequence alignments for domain C (D), domain D (E), S1/S2 cleavage site, S2' cleavage site (underlined), and fusion peptide (F), and stem helix sequence amino acid sequence divergence indicated in red (G). RshTT200 S monomer regions of amino acid sequence conservation (blue) and divergence (red) with SARS-CoV-2 (H) with differences in the NTD RBD and S2 domain are labeled by region and differences in domain C (blue) and domain D (green) are circled. Structural and glycosylation differences in RshTT200 (sea green) and SARS-

CoV-2 (pink) PDB:7LXY NTD (I). Alterations in N- linked glycosylation sites between RshTT200 and SARS-CoV-2 PDB:6VXX are labeled. Indel sites i and ii along the RBM/ACE2 interface indicated by circles, with differences between RshTT200 and SARS-CoV-2 notated (J). ELISA binding of RshTT200 (sea green) and SARS-CoV-2 Hexapro S with monoclonal antibodies targeting the NTD, RBD, and S2 subunit (K).



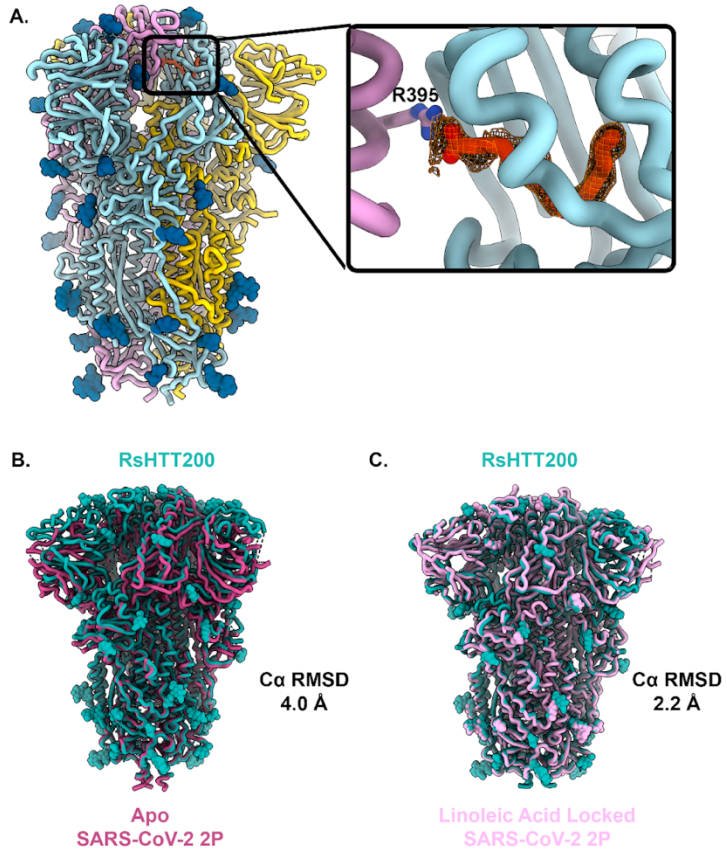
Supplemental Figure 3.1: Stabilization of RshTT200 S ectodomain.
Negative stain micrographs of RshTT200 S with 2P (K964P, V965P)(A), HexaPro (F795P, A870P, A877P, A920P, K964P, V965P) (B) and VFLIP del (658-668)+ insertion GGSGGGS linker (F795P, A870P, A877P, A920P, K964P, V965P, Y685C, T861C) mutations (C).



Supplemental Figure 3.2: Data Processing Pipeline for Cryo-EM structural reconstruction of the RshTT200 S.

Cryo-EM Data processing**.** Data Processing Pipeline of RShTT200 HexaPro S with 3D classes selected for subsequent refinements are indicated with boxes with black dotted borders (A). Representative electron micrograph scale bar correlates to 100 nm (B), 2D class averages of RshTT200 S embedded in vitreous ice with scale bar corresponding to 100 Å (C). Gold Standard Fourier shell correlation curves for the RshTT200 S. The

0.143 cutoff is indicated by horizontal dashed lines (D) Local resolution map calculated using cryo-SPARC with side (left) and top (right) views shown (E).

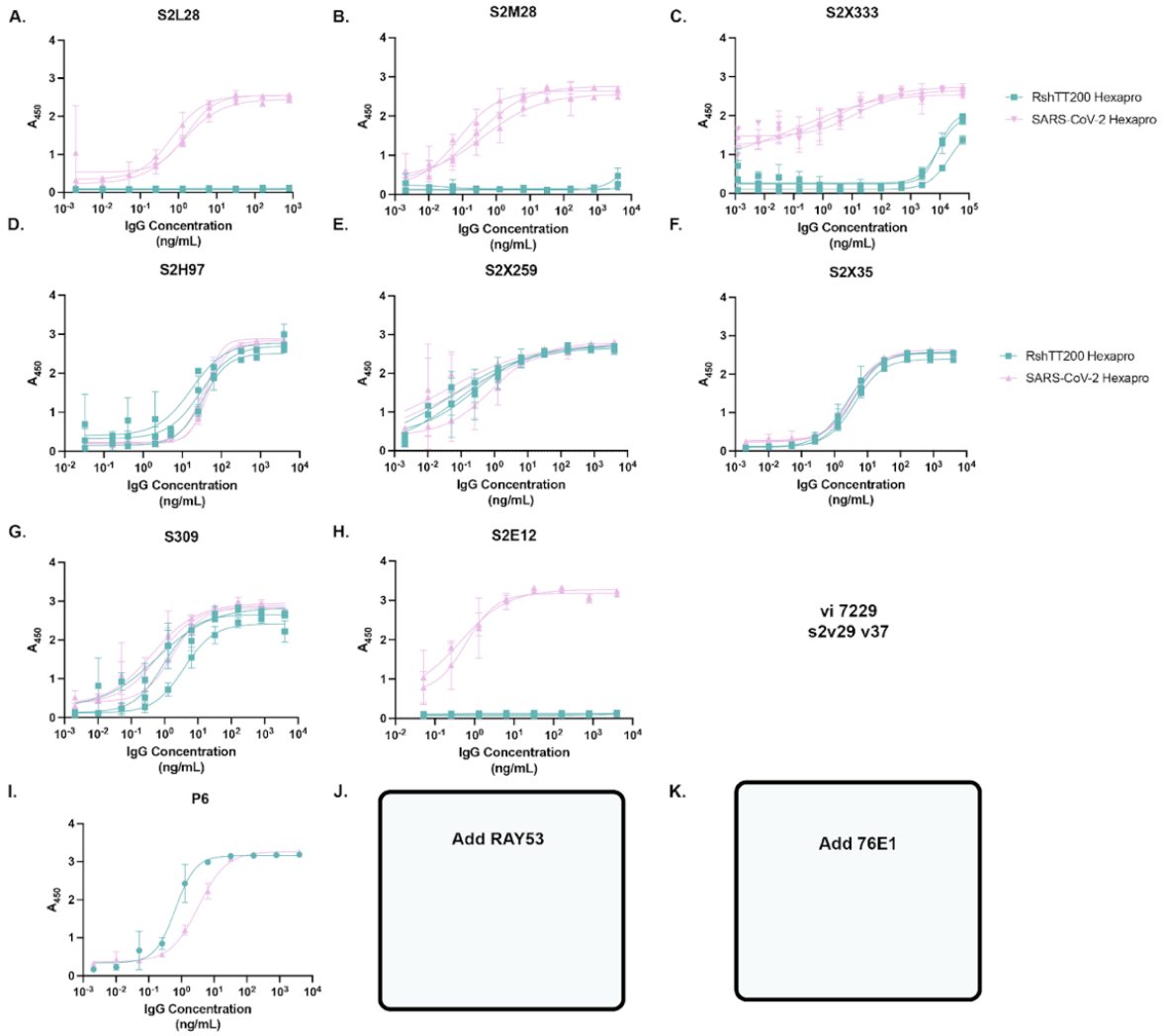


Supplemental Figure 3.3: ***Conservation of Linoleic Acid Binding Pocket.***

Ribbon diagram of the RShTT200 S trimer with linoleic acid (red orange) occupying fatty acid binding site in and each protomer indicated in plum, gold, or light blue (**A**). Overlay of Linoleic acid bound RshTT200 Hexapro (sea green) with Apo SARS-CoV-2 2P (pink, PDB:6VXX) (**B**). Overlay of Linoleic acid bound RshTT200 Hexapro (sea green) with linoleic acid bound SARS-CoV-2 2P (pink, PDB:6ZB5) **(C**).

Table S3.1*: Cα RMSD between RshTT200 and SARS-CoV-2 Wuhan (PDB: 6zb5).*

	pruned atoms	RMSD (Å)	all atoms	RMSD (Å)	Residues (SARS-CoV-2 #)	Residues (RshTT200#)
S	902	1.129	1004	2.224	1-1273	1-1251
S1	504	1.146	575	2.583	14-685	14-663
S2	424	0.813	430	1.248	686-1273	664-1251
NTD	208	0.783	224	2.813	14-307	14-302
RBD	194	0.649	199	1.033	328-531	315-513
C	63	0.586	63	0.586	528-591	511-573
D	62	0.722	70	3.384	592-674	574-686



Supplemental Figure 3.4: ***Monoclonal antibody binding to RshTT200 and SARS-CoV-2 Spikes.***

ELISA assays performed with RshTT200 Hexapro and SARS-CoV-2 Hexapro immobilized on a 384 well plate with binding measured across a titration of IgG concentrations at 450nm.

Table S3.2*: Conservation/Divergence of Glycosylation Sequon.*

RShTT200	SARS-CoV-2	SARS-CoV-1
V19	N17	T21
K27	P25	N29
N61	F59	Y63
V71	H68	N73

D73	N74	-
N111	S112	N109
K148	N149	T146
N158	N165	N158
N227	N234	N227
N269	N282	N269
N318	N331	N318
N330	N343	N330
N357	N370	N357
N585	N603	N589
N598	N616	N602
N639	N657	D643
N687	N709	N691
N695	N717	N699
N779	N801	N783
N1052	N1074	N1056
N1076	N1098	N1080
N1112	N1158	N1116

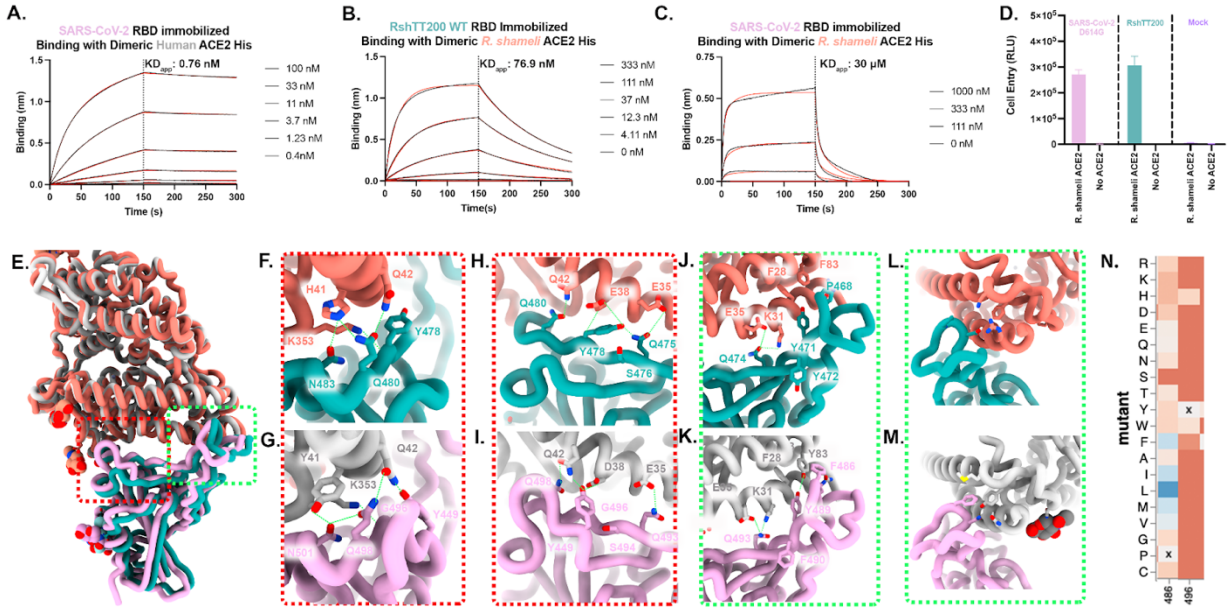
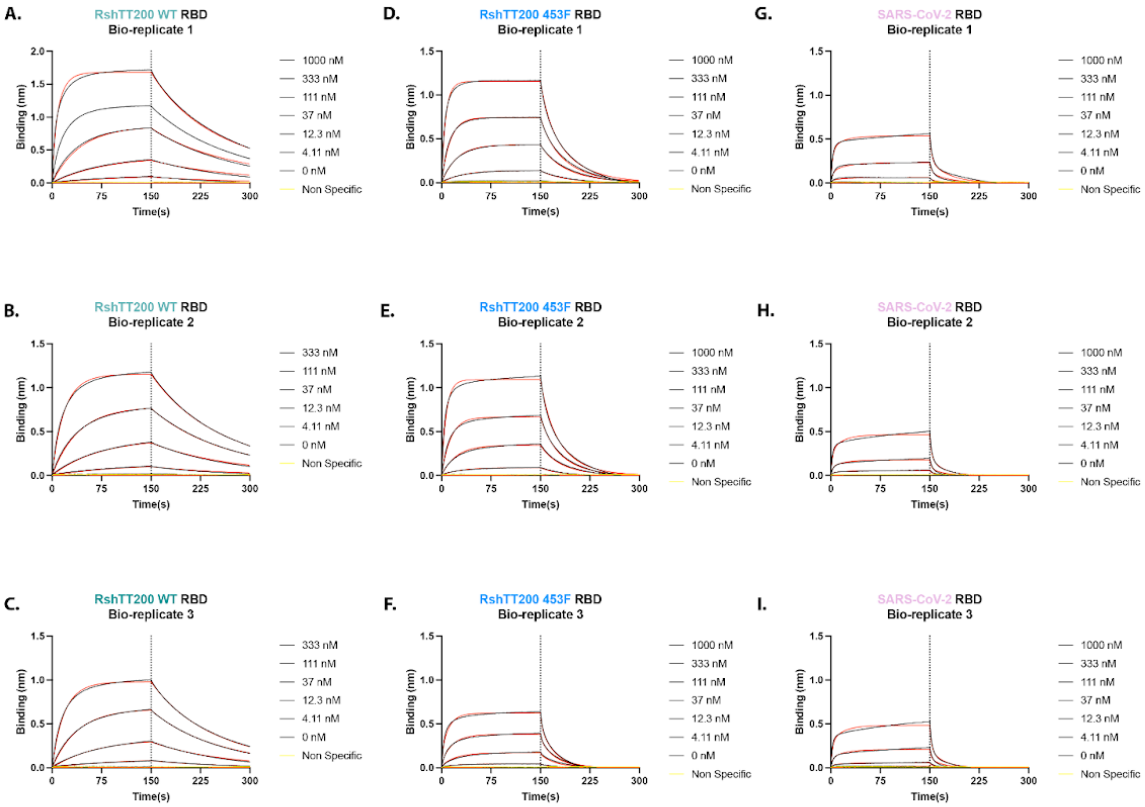


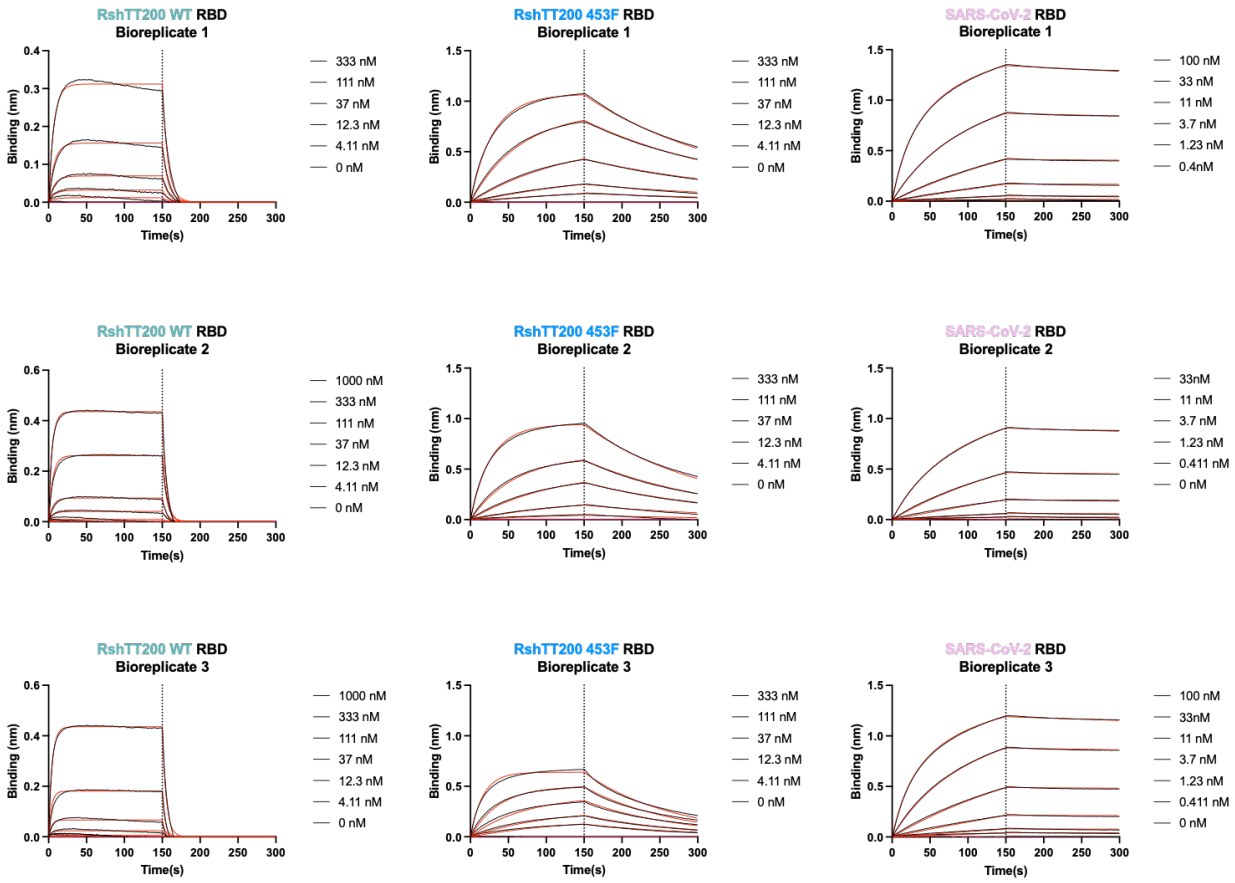
Figure 3.2: RshTT200 binding and key molecular interactions with host R. shameli ACE2

Representative biolayer interferometry measuring binding (avidity) of immobilized RBD with natively dimerized ACE2: SARS-CoV-2 (pink) with its native receptor human ACE2 (**A**), RshTT200 (sea green) with *R. shameli* ACE2 (**B**), and SARS-CoV-2 (pink) with *R. shameli* ACE2 (**C**). Additional Replicates, non-specific binding controls and kinetic analysis of replicates are reported in **Fig. S3.5- S3.6**, **Table S3.2**. Cell entry of RshTT200 (sea green) and SARS-CoV-2 (pink) S pseudotyped VSV (**D**). Comparison for cell entry were S normalized. Normalization reported in **Fig. S3.7**. Ribbon diagram of the RShTT200 RBD (sea green) in complex with *R. shameli* ACE2 (salmon) compared with SARS-CoV-2 RBD in complex with human ACE2 (PDB: 6M0J). Two critical sites for interactions are indicated by squares red and green dotted line(**E**). Close up views of these sites in RshTT200 RBD- *R. shameli* ACE2 (**F; H; J; L**) and SARS-CoV-2 RBD-Human ACE2 (**G; I; K; M**). DMS heatmaps of change in binding with mutation at two sites of divergence between RshTT200 and SARS-CoV-2 (**N**).



Supplemental Figure 3.5: ***Biolayer Interferometry Measuring Dimeric R. shameli ACE2 Binding. Binding.***

Binding curves (black) and fits (red) are shown for three biological replicates of RshTT200 WT RBD (A-C), RshTT200 453F RBD (D-F), and SARS-CoV-2 (G-I) immobilized on a SA biosensor and dipped into dimeric *R. shameli* ACE2. In yellow, non-specific binding is shown (no RBD loaded on a biosensor dipped into 1 μM *R. shameli* ACE2). Transition from association to dissociation phase is shown as a dotted line. Kinetic analysis of these replicates are reported in Table S2.



Supplemental Figure 3.6: ***Biolayer Interferometry Measuring Dimeric Human ACE2 Binding.***

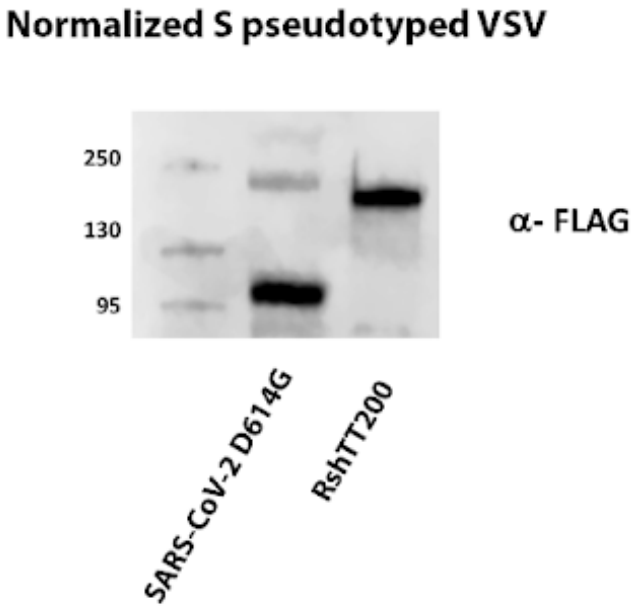
Binding curves (black) and fits (red) are shown for three biological replicates of RshTT200 WT RBD (A-C), RshTT200 453F RBD (D-F), and SARS-CoV-2 (G-I) immobilized on a SA biosensor and dipped into dimeric Human ACE2. Transition from association to dissociation phase is shown as dotted line. Kinetic analysis of these iterations are reported in Table S2.

Table S3.3*: Kinetic Analysis of ACE2 binding to sarbecovirus RBD by Biolayer Affinity*

R shameli Native Dimeric ACE2 Avidity			
	KD_{app}	K_{on} ($M^{-1}s^{-1}$)	K_{off} (s^{-1})
RshTT200 WT	152.6 ± 31.1 nM	1.17×10^{5}	9.02×10^{-3}
RshTT200 453F	2.9 ± 3.2 nM	8.3×10^{4}	6.9×10^{-1}

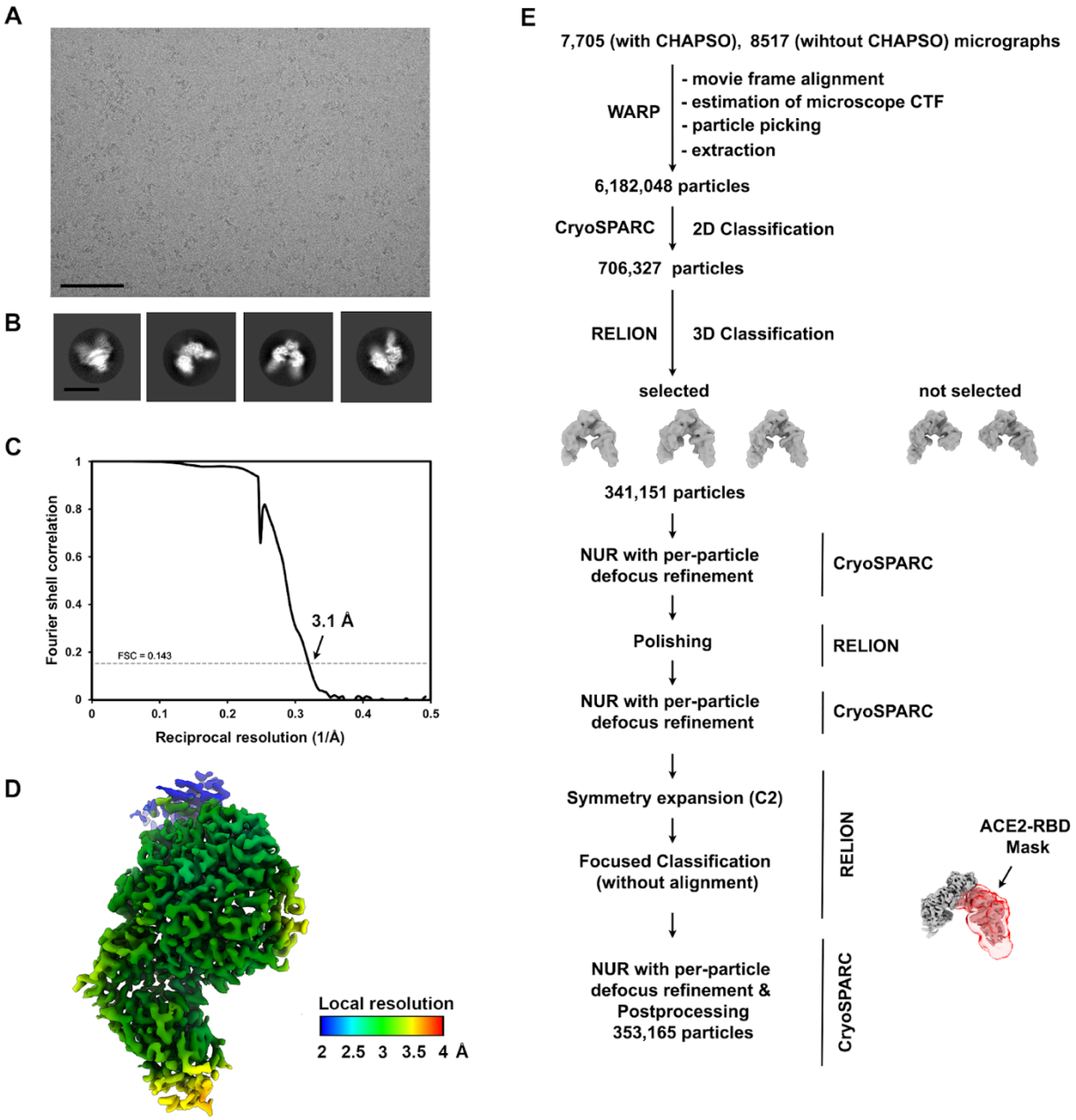
SARS-CoV-2	27.3 ± 12 μM*	1.5 x 10^{5}	8.5 x 10^{-3}
R. affinis 9479 Native Dimeric ACE2			
RshTT200 WT	207 ±60 nM	1.5 x 10^{5}	3.0 x 10^{-2}
RshTT200 453F	50.9 ± 19.4 nM	2.5 x 10^{4}	3.7 x 10^{-3}
SARS-CoV-2	107 ± 0.4 nM	1.5 x 10^{5}	3.0 x 10^{-2}
RaTG13	3± 2 μM*	5.0 x 10^{4}	1.5 x 10^{-2}
R. affinis 787 Native Dimeric ACE2			
RshTT200 WT	57.1 ± 4.6 nM	2.0 x 10^{5}	1.1 x 10^{-2}
RshTT200 453F	135 ± 54.8 nM	1.6 x 10^{5}	2.1 x 10^{-2}
SARS-CoV-2	417 ± 9.89 nM	1.2 x 10^{5}	5.1 x 10^{-2}
RaTG13	N.D.	N.D.	N.D.
R. sinicus WJ1 Native Dimeric ACE2			
RshTT200 WT	82.5 ± 0.4 nM	8.3 x 10^{4}	6.9 x 10^{-1}
RshTT200 453F	60.3 ± 12.7 nM	1.5 x 10^{5}	8.5 x 10^{-3}
SARS-CoV-2	77.5 nM	9.6 x 10^{4}	7.4 x 10^{-3}
Human Native Dimeric ACE2			
RshTT200 WT	4.5 ± 1.6 μM	4.1 x 10^{4}	4.1 x 10^{-4}
RshTT200 453F	115.6 ± 8.0 nM	5.8 x 10^{4}	7.0 x 10^{-3}
SARS-CoV-2	2.21 ± 0.9 nM	1.4 x 10^{5}	3.0 x 10^{-4}
RaTG13	243.0 ± 18.6 nM	2.2 x 10^{5}	5x 10^{-3}

*For these interactions we were unable to surround the KD values (best practice is to have max concentration to be ~ 10x the KD and the minimum concentration 1/10 of KD value) in our assay. This resulted from limitations of BLI, and the protein stability limits of our proteins. As such these should be considered to be approximate values.



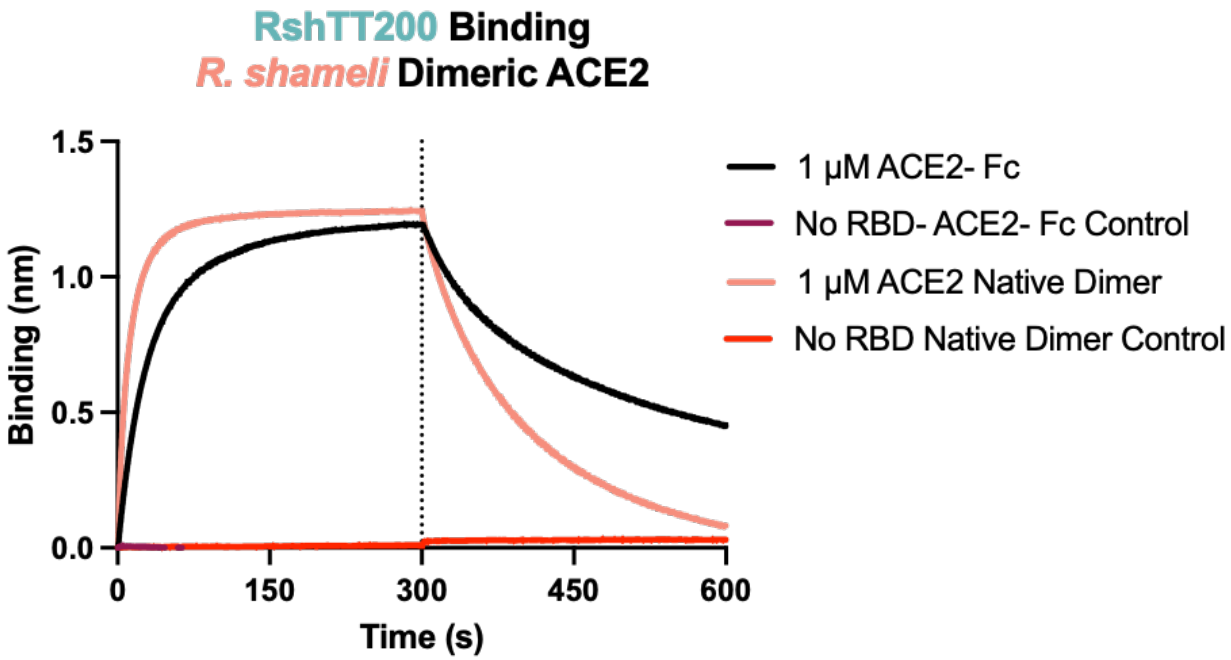
Supplemental Figure 3.7: ***Normalization of RshTT200 S and SARS-CoV-2 S VSV pseudotyped particles for cell entry.***

Anti-FLAG tag antibody detection of RshTT200 and SARS-CoV-2 containing a C-terminal del21 and a 3x FLAG tag S pseudotyped particles. Associated with **Fig. 2D**.



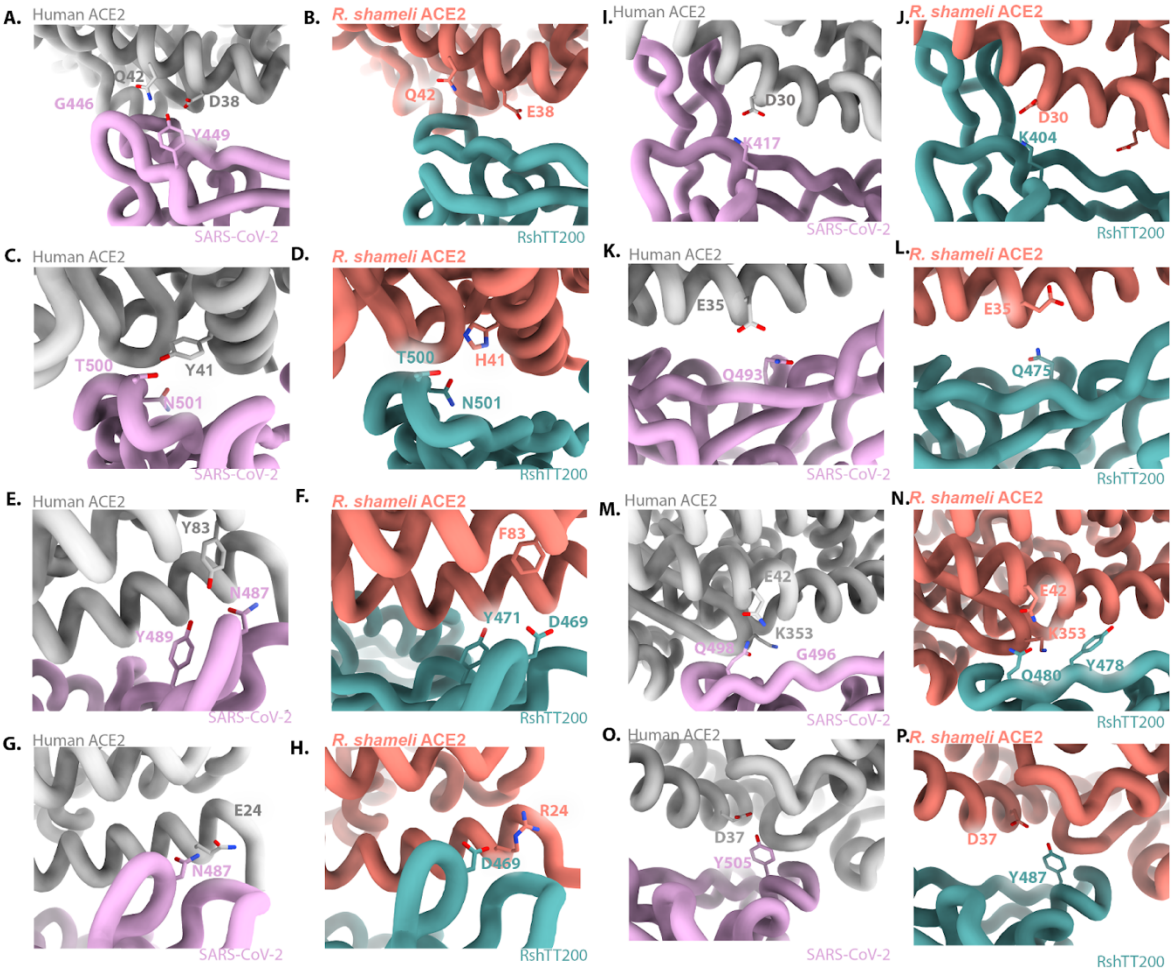
Supplemental Figure 3.8: ***Data Processing Pipeline for Cryo-EM structural reconstruction of the RshTT200 RBD in complex with R. shameli ACE2.***

Representative electron micrograph with scale bar correlates to 100 nm (A). 2D class averages of RshTT200 S embedded in vitreous ice with scale bar corresponding to 100 Å (B). Gold Standard Fourier shell correlation curves. The 0.143 cutoff is indicated by horizontal dashed lines (C) Local resolution map calculated using cryo-SPARC (E). Data processing pipeline.



Supplemental Figure 3.9: ***Biolayer Interferometry Comparison of the R. shameli ACE2- Fc dimerized, and natively dimerized R. shameli ACE2.***

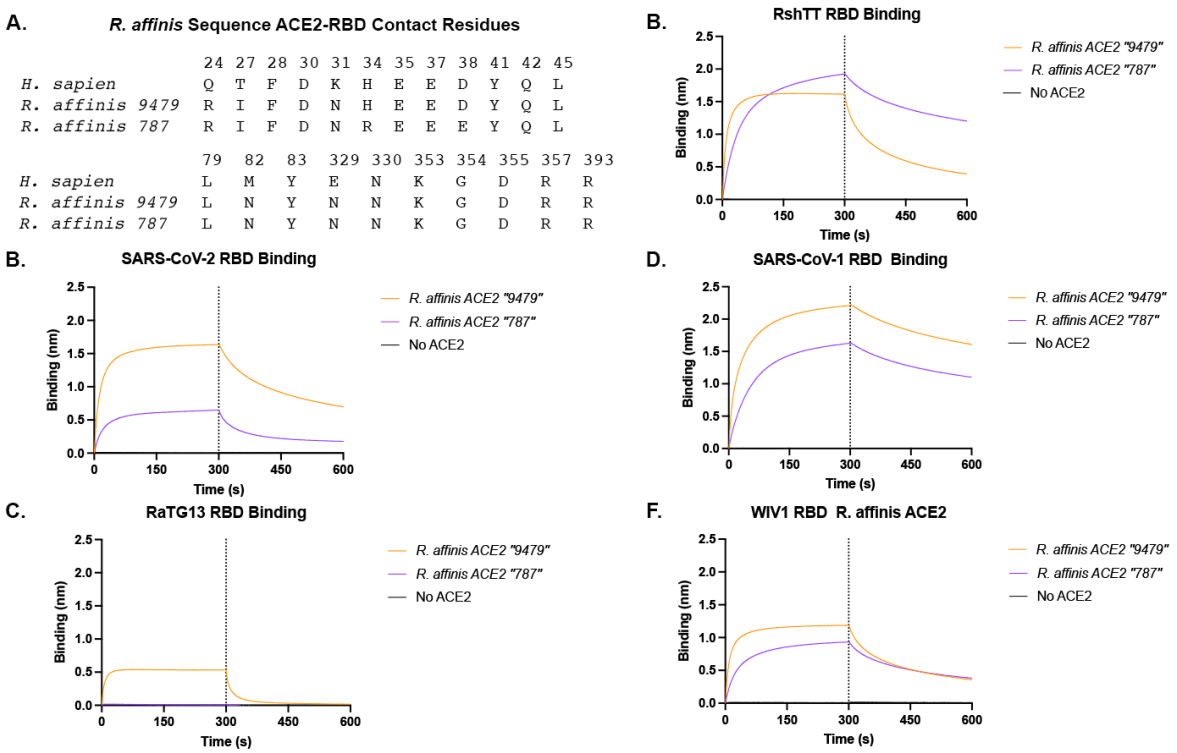
RshTT200 RBD immobilized on SA Biosensor tips, were dipped into 1 μM *R. shameli* ACE2 ectodomains dimerized by either the native dimerization (residues 19-740) (salmon) or with a C-terminal Fc-fusion with the dimerization domain truncated (residues 19-640) (black). No RBD loaded controls dipped into 1 μM ACE2 Fc (Maroon) or 1μM ACE Native Dimer (Red) were used to assess non-specific binding to the biosensor tip.



Supplemental Figure 3.10: ***RshTT200 interactions with R. shameli ACE2.***

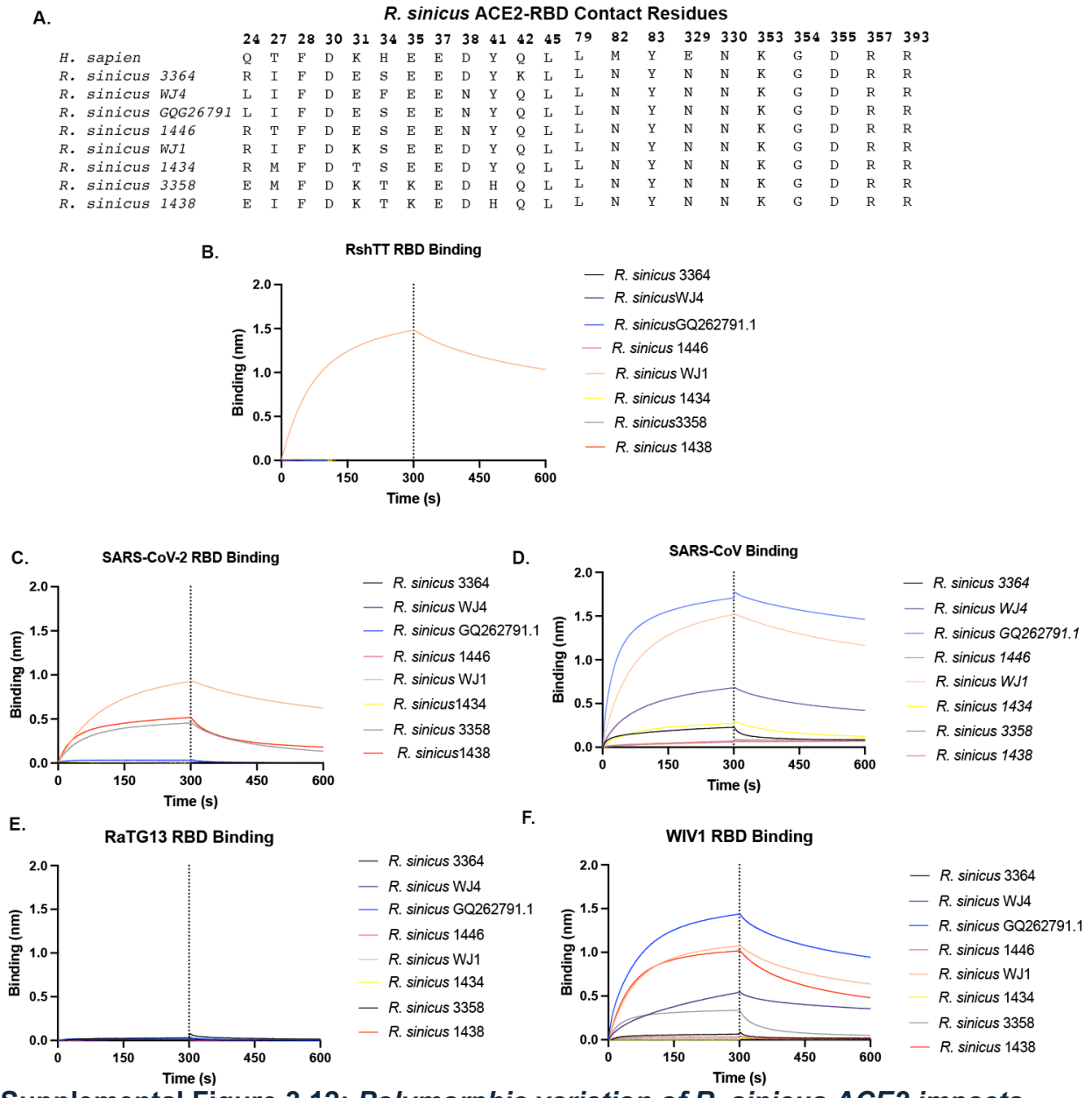
Table S3.4: **ACE2-RBD Interface Residues across Rhinolophus bat species.** Clustal Alignment of Rhinolophus ACE2 residues along the interface responsible for sarbecovirus RBD interactions; conserved residues (green) and divergent residues (red) are indicated. Residues are highlighted by their chemical character: hydrophobic (pink), positive charge (light orange), negative charge (dark cornflower blue), polar (light blue), prolines (yellow), aromatic (teal).

Species	Isolate	Accession	19	23	24	27	28	30	31	34	35	37	38	41	42	45	79	82	83	324	325	326	329	330	353	354	355	357	386	393
R. shameli	-	UBB59645.1	S	D	R	A	F	D	K	P	E	E	E	H	Q	L	L	N	F	T	E	G	N	N	K	N	D	R	A	R
R. affinis	9479	MT394208.1	S	D	R	I	F	D	N	H	E	E	D	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. affinis	787	MT394203.1	S	D	R	I	F	D	N	R	E	E	E	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	3364	MT394200.1	S	D	R	I	F	D	E	S	E	E	D	Y	K	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	WJ4	MT394181.1	F	D	L	I	F	D	E	F	E	E	N	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	-	GQ262791.1	F	D	L	I	F	D	E	S	E	E	N	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	1446	MT394194.1	S	D	R	T	F	D	E	S	E	E	N	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	WJ1	MT394187.1	S	D	R	I	F	D	K	S	E	E	D	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	1434	MT394197.1	S	D	R	M	F	D	T	S	E	E	D	Y	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	3358	MT394193.1	S	D	E	M	F	D	K	T	K	E	D	H	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. sinicus	1438	MT394184.1	S	D	E	I	F	D	K	T	K	E	D	H	Q	L	L	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. pusillus	-	GQ999938.1	S	D	K	K	F	N	D	S	E	E	D	Y	Q	L	I	N	Y	T	E	G	N	N	K	G	D	R	A	R
R. alcyone	-	KR559016.1	S	D	L	I	F	D	N	S	E	E	N	H	Q	L	H	N	F	T	E	G	N	N	K	G	D	R	A	R
R. landeri	-	KR559015.1	S	D	L	T	F	D	D	S	A	E	N	Y	Q	L	H	N	F	T	E	G	N	N	K	G	D	R	A	R
R. macrotis	-	ADN93471.1	S	D	E	K	F	D	K	S	K	E	D	Y	E	L	L	N	Y	T	E	G	N	K	K	G	D	R	A	R
Human	-	BAB40370.1	S	E	Q	T	F	D	K	H	E	E	D	Y	Q	L	L	M	Y	T	Q	G	E	H	K	G	D	R	A	R



Supplemental Figure 3.11: ***Polymorphic variation of R. affinis ACE2 impacts sarbecovirus-ACE2 binding.***

(A) *R. affinis* ACE2 allele variation at residues involved in sarbecovirus RBD associations. Biolayer Interferometry of biotinylated RBD immobilized on streptavidin (SA) biosensors and dipped into 1 uM *R. affinis* ACE2 Fc (dimerized) allele 9479 (orange), allele 787 (purple), and no ACE2 (black). Binding was assessed against (B) RshTT182/200, (C) SARS-CoV-2, (D) SARS-CoV-1 (E) RaTG13, (F) WIV1 RBDs.



Supplemental Figure 3.12: *Polymorphic variation of R. sinicus ACE2 impacts sarbecovirus-ACE2 binding.*

(A) *R. sinicus* ACE2 allele variation at residues involved in sarbecovirus RBD associations. Biolayer Interferometry of biotinylated RBD immobilized on streptavidin (SA) biosensors and dipped into 1 uM *R. sinicus* ACE2 Fc (dimerized). Binding was

assessed against (B) RshTT182/200, (C) SARS-CoV-2, (D) SARS-CoV-1 (E) RaTG13, (F) WIV1 RBDs.

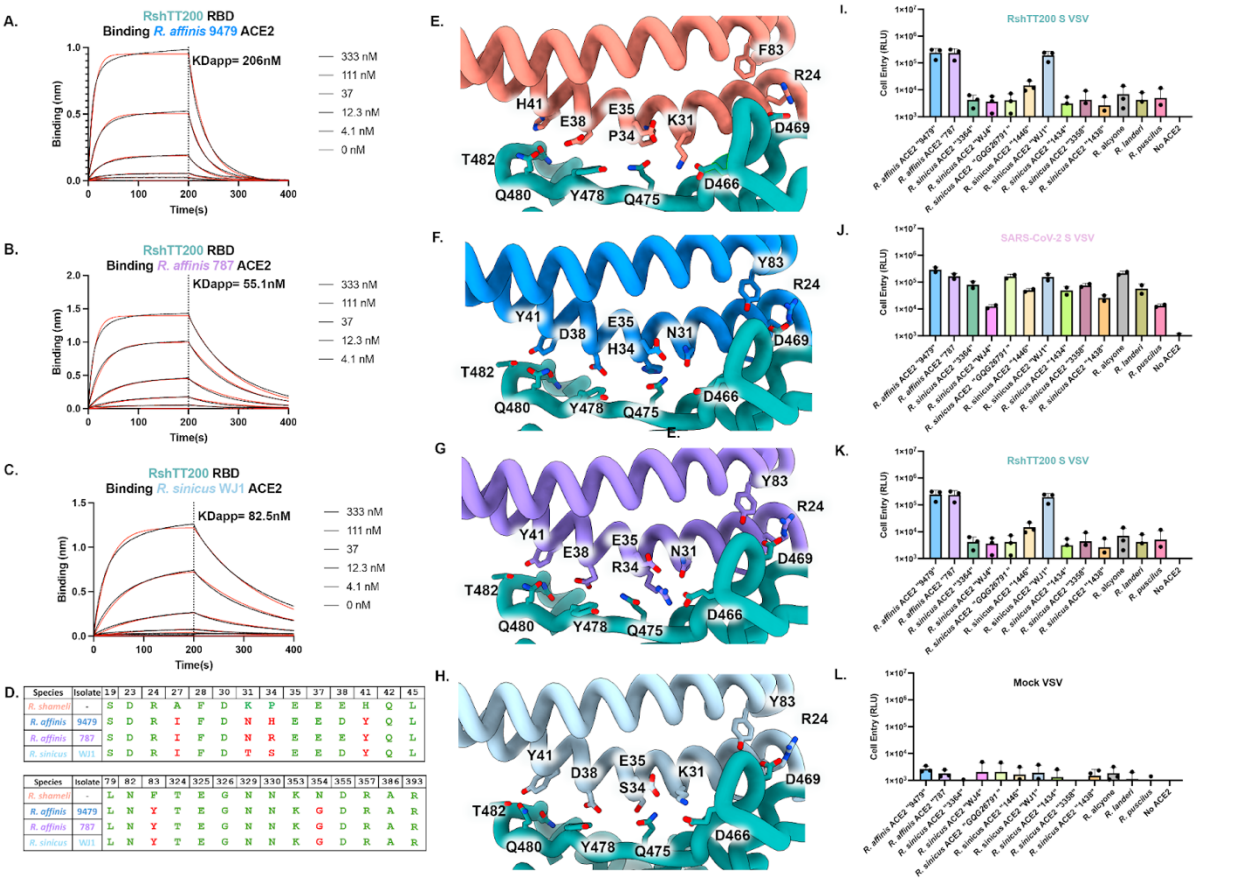
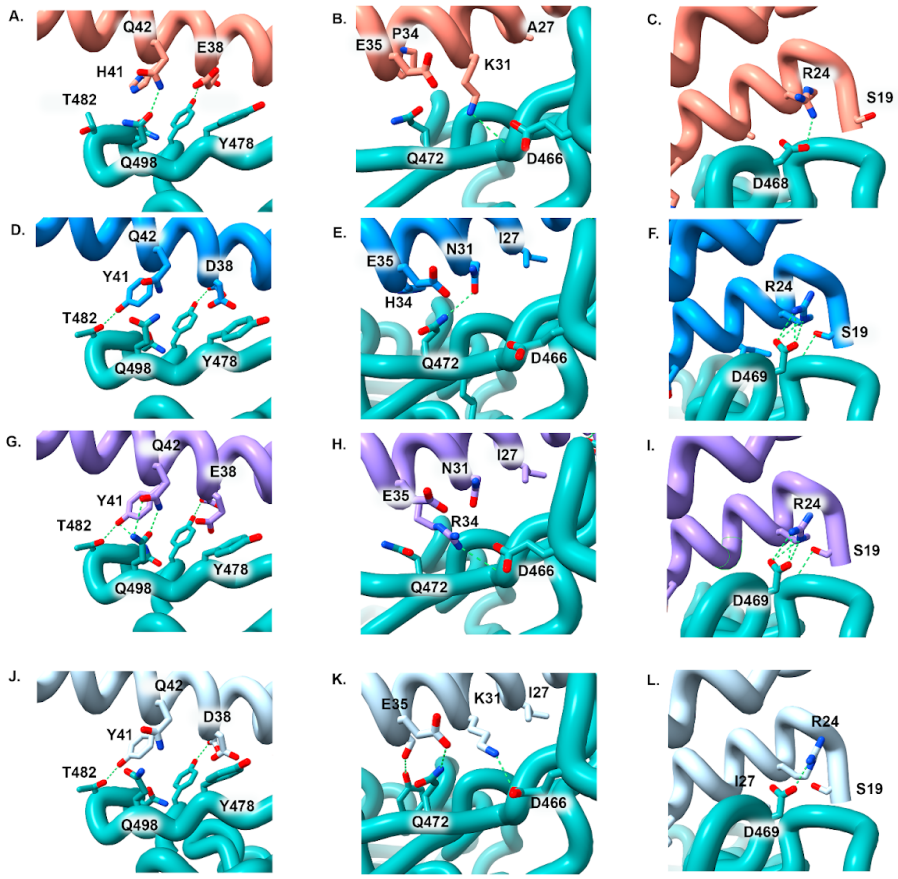


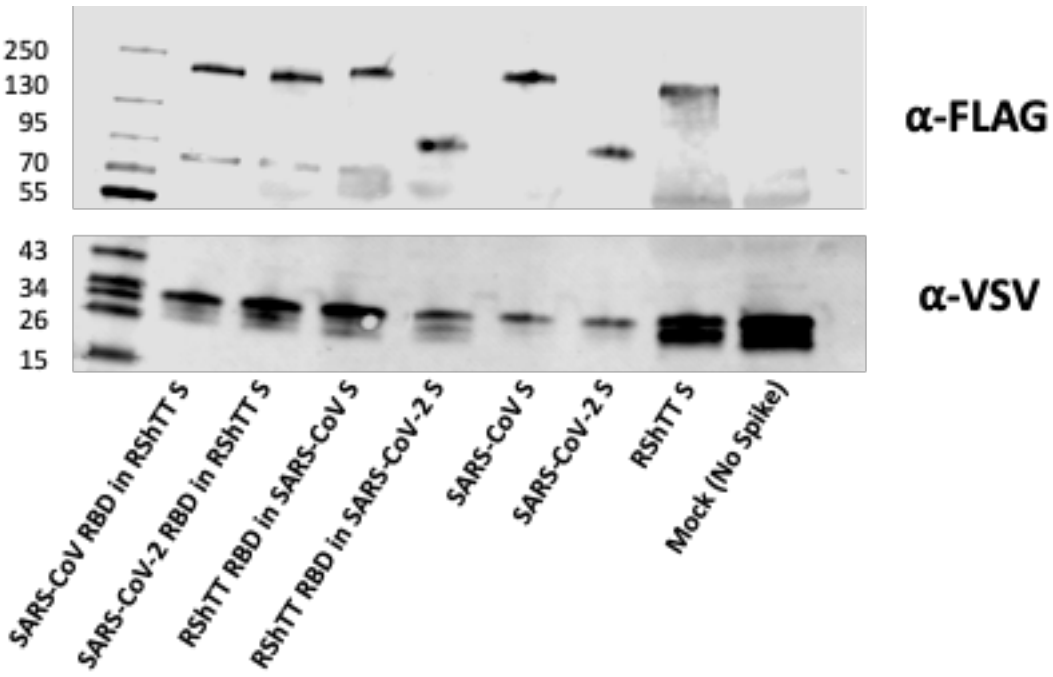
Figure 3.3: RshTT200 binding and key molecular interactions with Rhinolophus ACE2 panel.

Representative biolayer interferometry binding of immobilized RshTT200 RBD with natively dimerized *R. affinis* allele 9479 (A), *R. affinis* allele 787 (B), *R. sinicus* allele *WJ1 (C).* Sequence alignment of key residues along the ACE2-RBD interface for *R. shameli* ACE2, *R. affinis* ACE2 alleles 9479/787, and *R. sinicus* ACE2 allele WJ1 (D). Residues that differ from host species, *R. shameli* ACE2 sequence are shown in red. Expanded sequence alignment of key interacting residues of Rhinolophus ACE2 can be found in **Table S3.4.** RshTT200 RBD interface with *R. shameli* (E) from our structure compared to alphafold models of *R. affinis* allele 9479 (F), *R. affinis* allele 787 (G), and *R. sinicus* allele WJ1 (H) with key interacting residues notated. Spike normalized (**Fig. S3.13**) RshTT200 S (I), SARS-CoV-2 S (J) , and SARS-CoV-1 (K) S, and Mock (no spike) (L) pseudotyped VSV particles entry were tested across a panel of Rhinolophus ACE2 transfected in 293T cells. Each point represents a biological replicate that was tested in technical triplicates and repeated across two separate days.

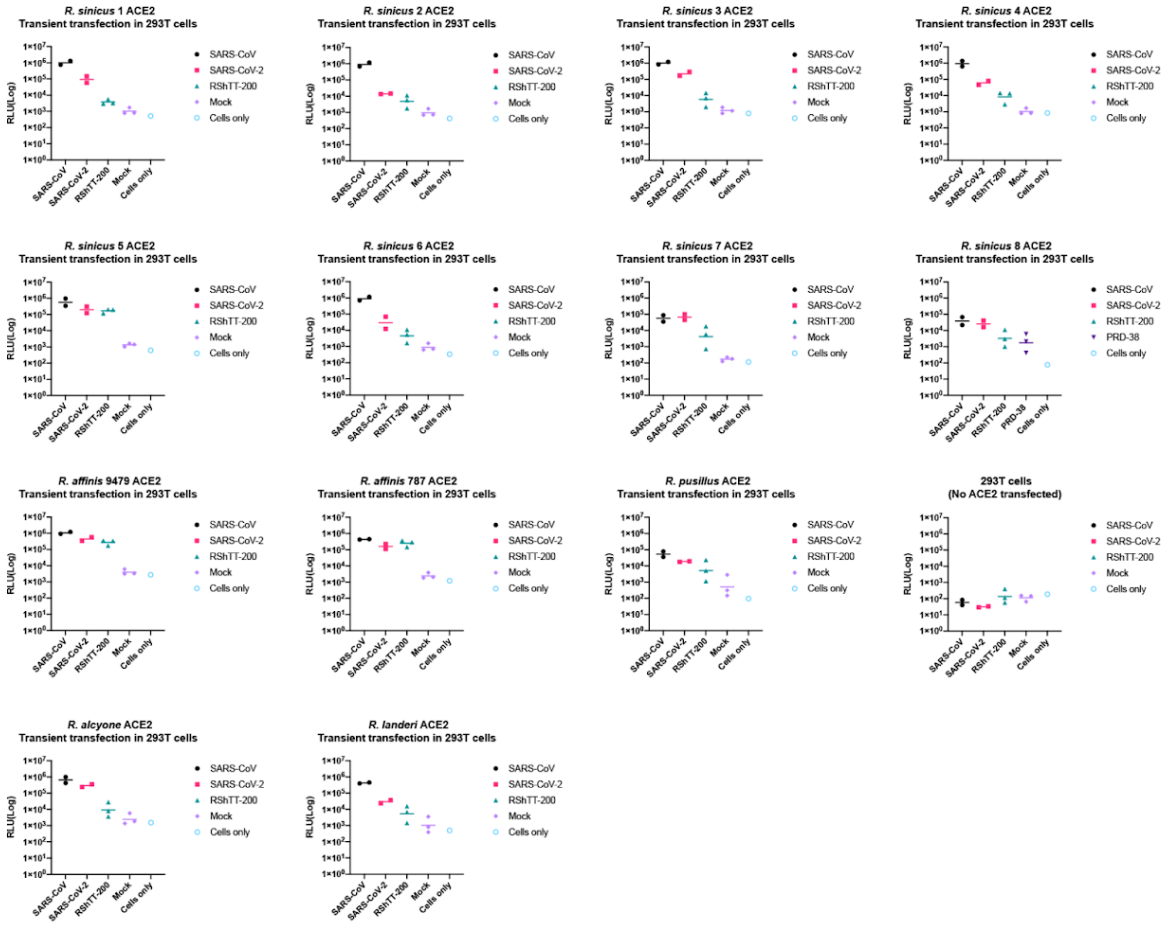


Supplemental Figure 3.13: ***Molecular Basis of RshTT200 RBD Interactions with Rhinolophus bat ACE2 receptors.***

RshTT200 RBD (sea green) interactions with *R. shameli* ACE2 (salmon) (A-C), *R. affinis* 9479 ACE2 allele (blue) in panels D-F, *R. affinis* 787 ACE2 allele (purple) in panels G-I, and *R. sinicus* WJ1 ACE2 allele (light blue) in panels J-L. Hydrogen bonds are indicated with green dotted lines.

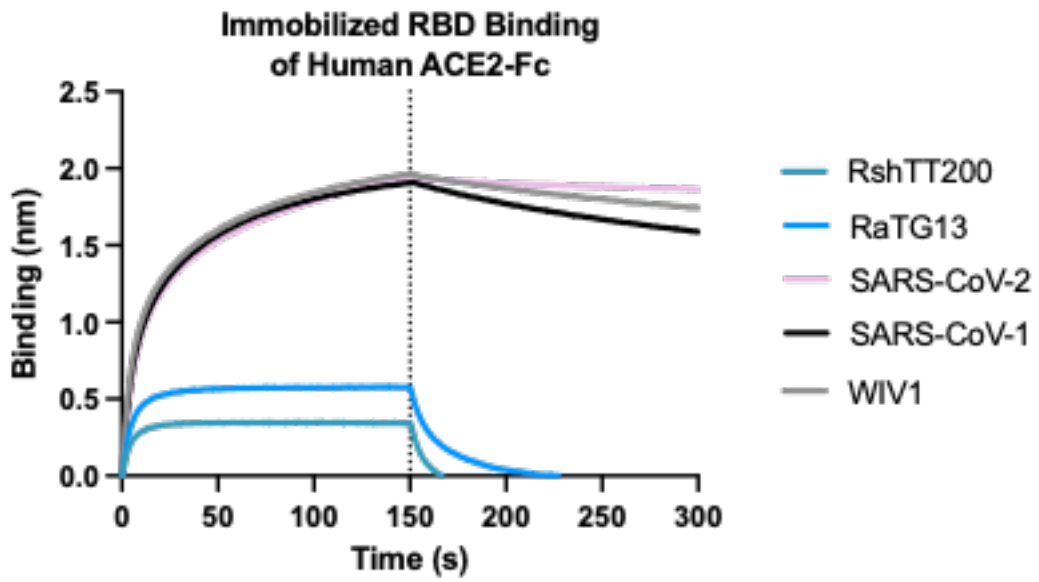


Supplemental Figure 3.14: ***Representative Western Blot of Spike Normalized RshTT200 Chimera S VSV.***

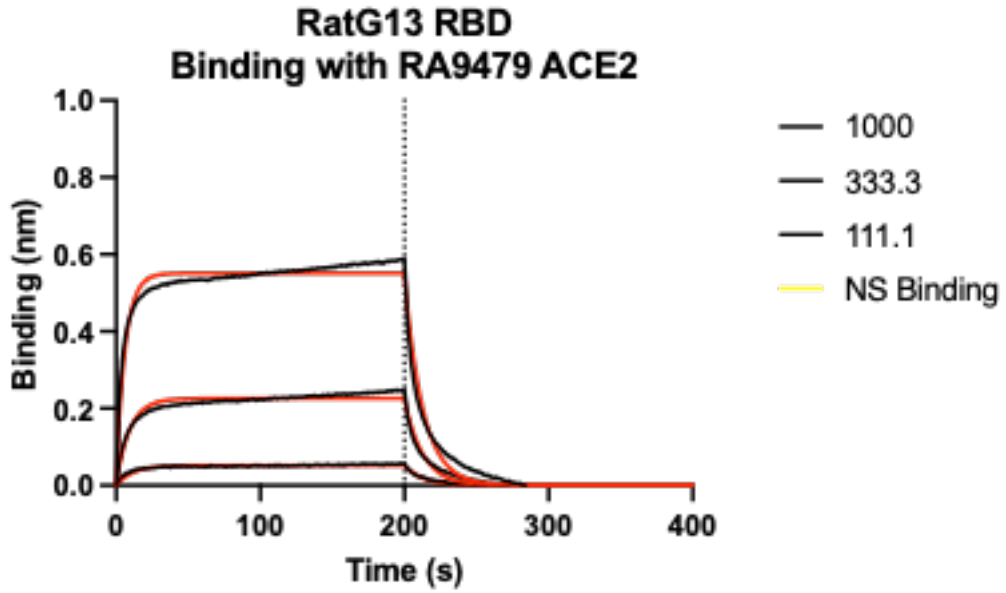


Supplemental Figure 3.15: ***Polymorphic variation of R. affinis and R. sinicus ACE2 impact on cellular entry.***

Cell entry in HEK293T cell transiently transfected with Rhinolophus ACE2 and infected with SARS-CoV-1 (black), SARS-CoV-2 (pink), and RshTT200 (sea green) S or mock (no S) pseudotyped particles.



Supplemental Figure 3.16: ***Biolayer Interferometry Immobilized Sarbecovirus Binding of hACE2 Fc Fusion ectodomains at 1 μM.***



Supplemental Figure 3.17: ***Representative BLI Binding of immobilized RaTG13 RBD with titration of RA9479 ACE2.***

Binding in nm shown in black with 1:1 Binding model fits shown in red. In yellow we show a no load RBD control dipped in 1 μM ACE2 to demonstrate the absence of nonspecific binding to the SA biosensor tip.

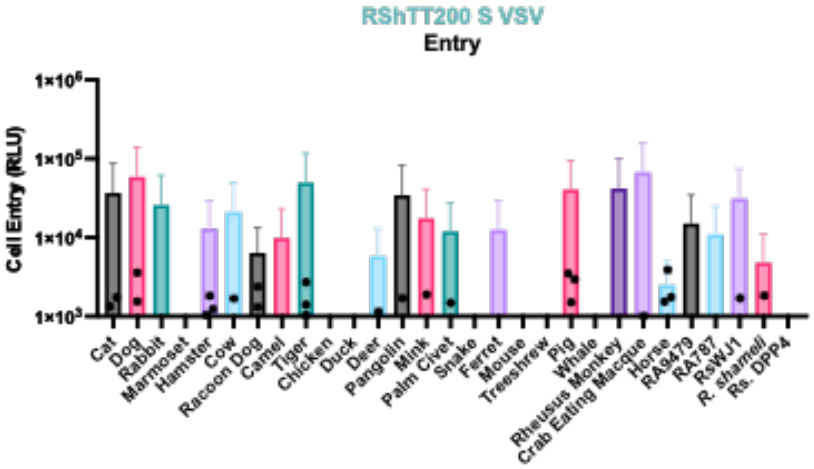


Figure 3.4: RshTT200 Tropism across Domesticated and Wildlife Species ACE2.

RshTT200 S pseudotyped VSV particles entry in HEK293T cells transiently transfected with the ACE2 orthologs of domesticated and wildlife species.

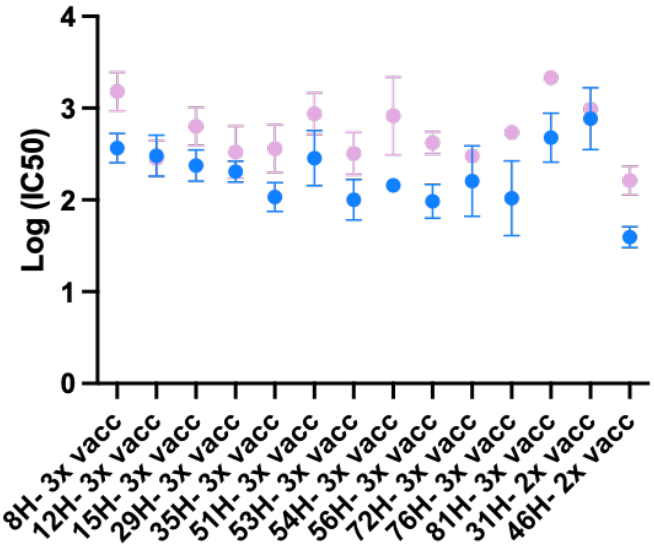
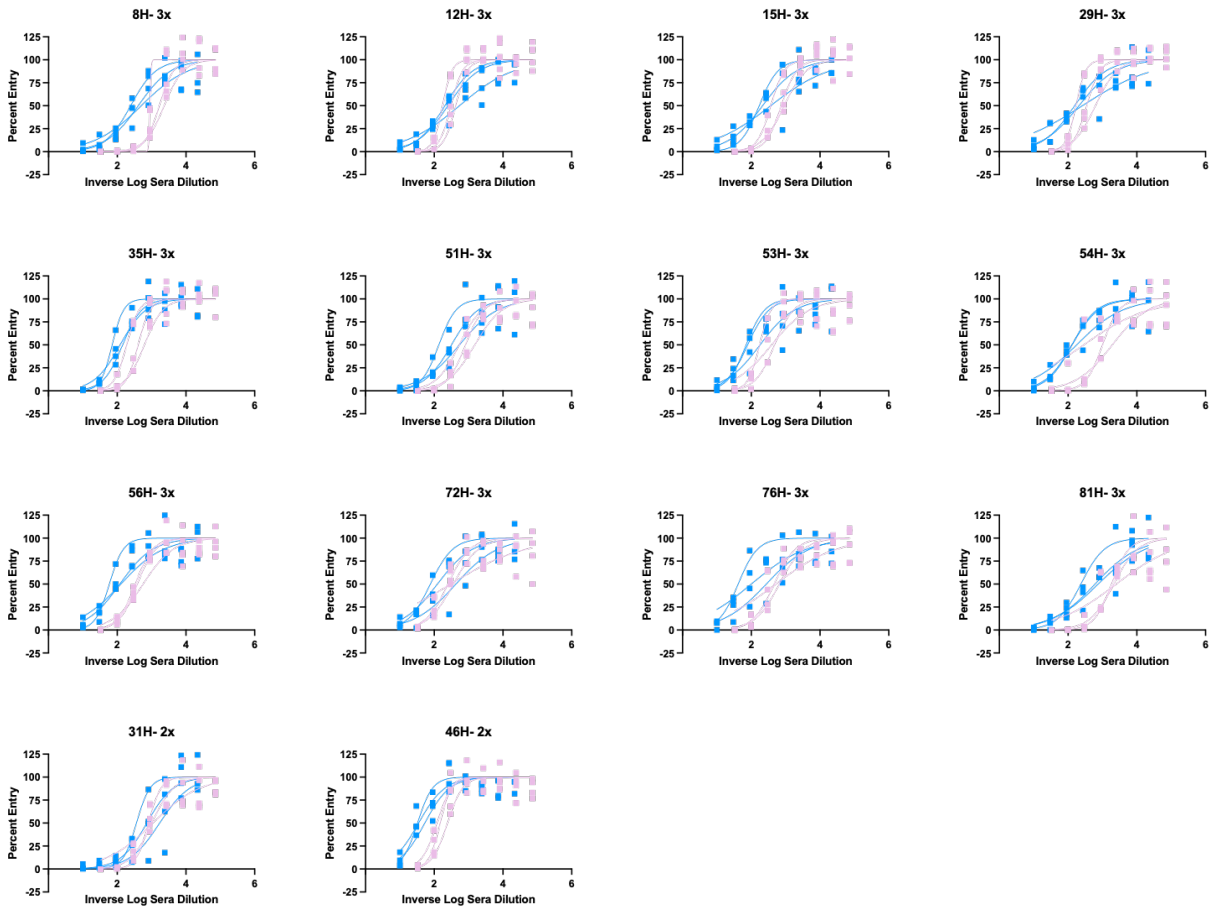


Figure 3.5: Polyclonal Sera from patients vaccinated 2-3x against SARS-CoV-2 Wuhan demonstrates equivalent or improved neutralization potency against RshT200 453F S VSV.

Each point represents a mean of three biological replicates log IC50 values. Patient cohort information can be located in **Table S3.4**. Full curves for all biological replicates in **Fig. S3.18.**



Supplemental Figure 3.18: ***Polyclonal Sera Neutralization of RshTT200 453F (royal blue) and SARS-CoV-2 D614G S VSV.***

Patient cohort of 14 patients (demographic information in **Table S4**) …

Table S1:Patient Cohort for polyclonal neutralization studies (in Fig. 3.5, S3.18).

ID	Age	Sex	BMI	Hispanic or latino?	Race/ethnicity	Post Boost Draw Date	Initial Vaccine Type	Draw Time Post Dose 1	Draw Time Post Dose 2	Booster Vaccine Type	Draw Time Post Dose 3
8H	79	Male	30.84592	No	White	2/2/2022	Moderna	364	332	Moderna	78
12H	42	Female	34.17361	No	White	11/17/2021	Moderna	320	291	Moderna	18
15H	52	Male	37.22973	No	White	11/12/2021	Pfizer	221	199	Pfizer	13
29H	64	Male	25.79944	No	White	10/15/2021	Pfizer	216	195	Pfizer	13
31H	49	Female	19.57563	No	Asian	11/16/2021	Johnson & Johnson	204	N/A	Pfizer	13
35H	60	Male		No	Other	9/21/2021	Pfizer	268	247	Pfizer	47
46H	60	Male	24.39189	No	White	11/10/2021	Johnson & Johnson	217	N/A	Pfizer	16
51H	50	Female	25.42066	No	White	10/27/2021	Pfizer	311	292	Pfizer	25
53H	38	Female		No	White	10/13/2021	Pfizer	296	265	Pfizer	14
54H	34	Female		No	White	10/22/2021	Pfizer	308	289	Pfizer	15
55H	52	Female	21.2846	No	White	11/2/2021	Pfizer	307	289	Pfizer	18
56H	46	Male		Yes	Other	2/8/2022	Moderna	409	383	Moderna	97
72H	35	Female		No	White	10/21/2021	Pfizer	259	231	Pfizer	15
76H	38	Female		No	White	11/19/2021	Pfizer	261	242	Pfizer	14
80H	35	Female		No	White	1/11/2022	Moderna	363	335	Moderna	25

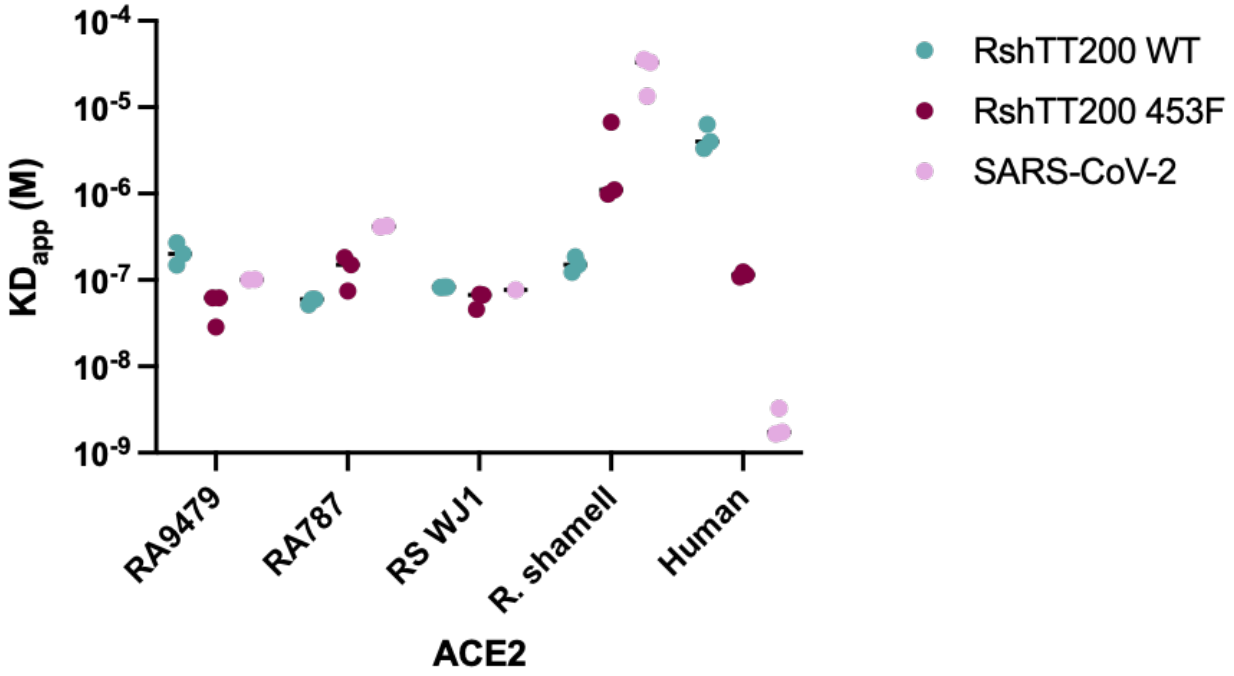
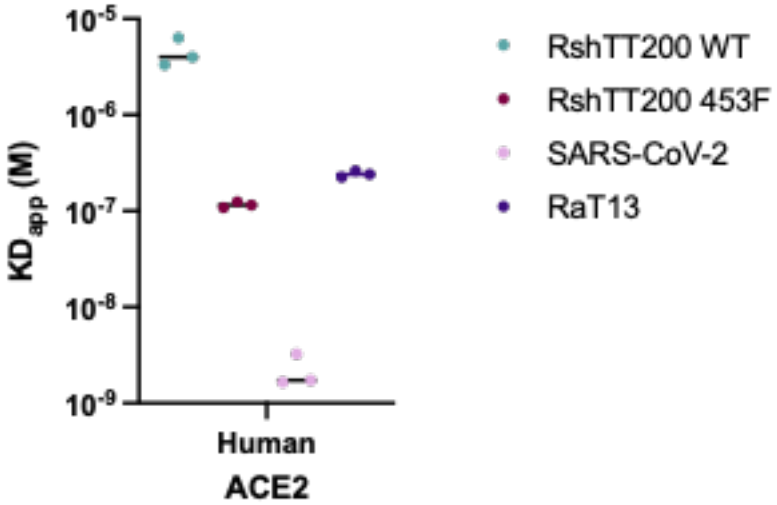


Figure 3.6: Impact of RshTT200 Y453F mutation on binding to Rhinophus and Human ACE2.

KD_{app} of RshTT200 RBD (sea green), RshTT200 453F (maroon) and SARS-CoV-2 (pink)



Supplemental Figure 3.19: ***RaTG13 RBD binds human ACE2 roughly equivalent to RshTT200 453F.***

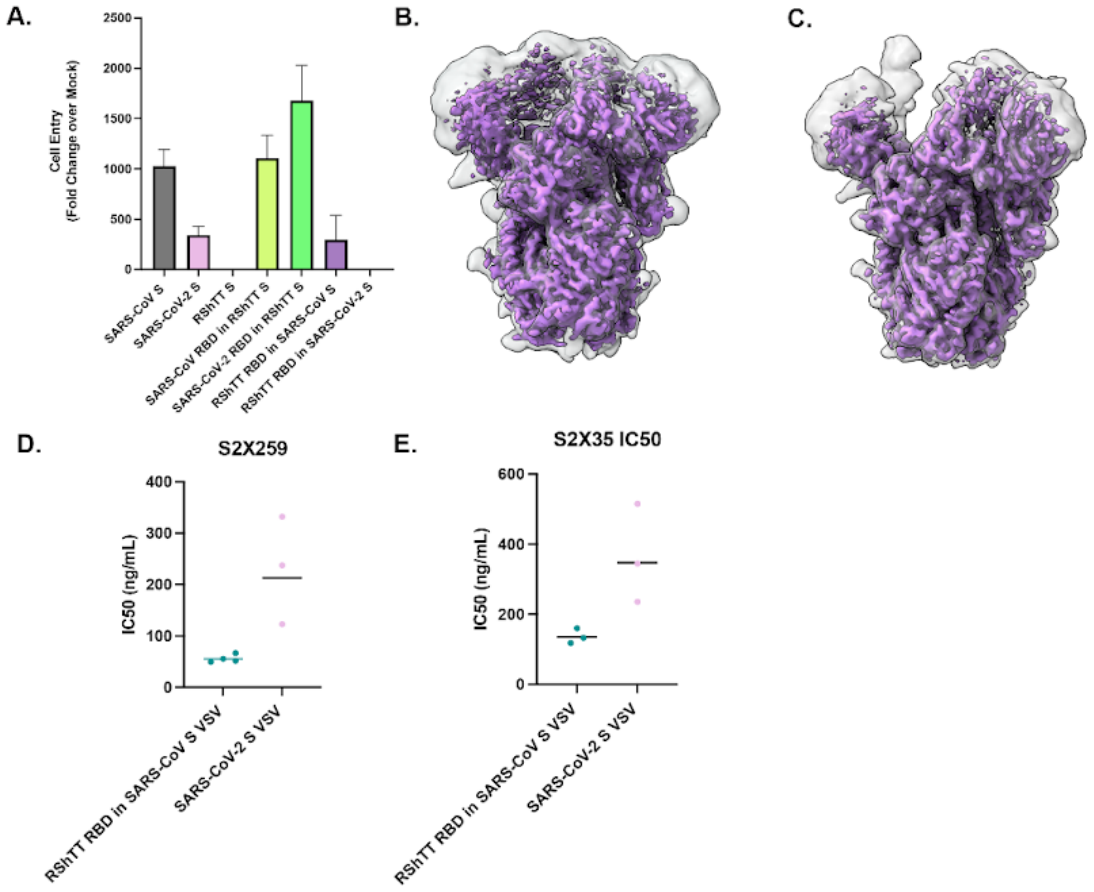
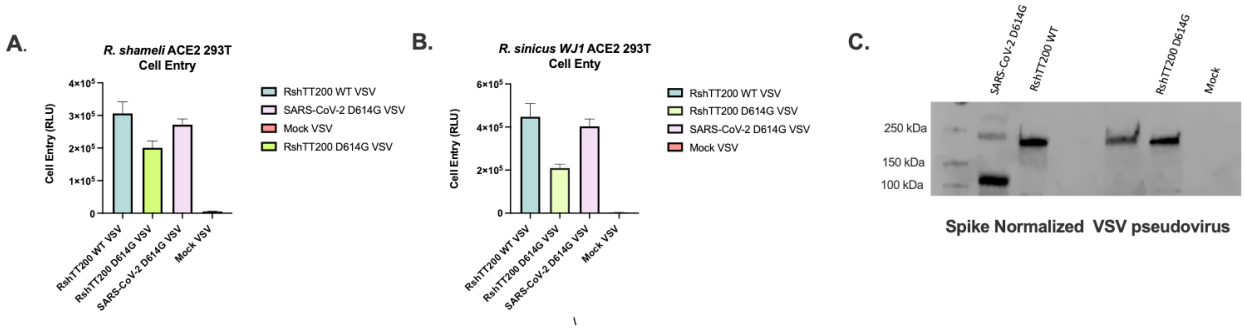


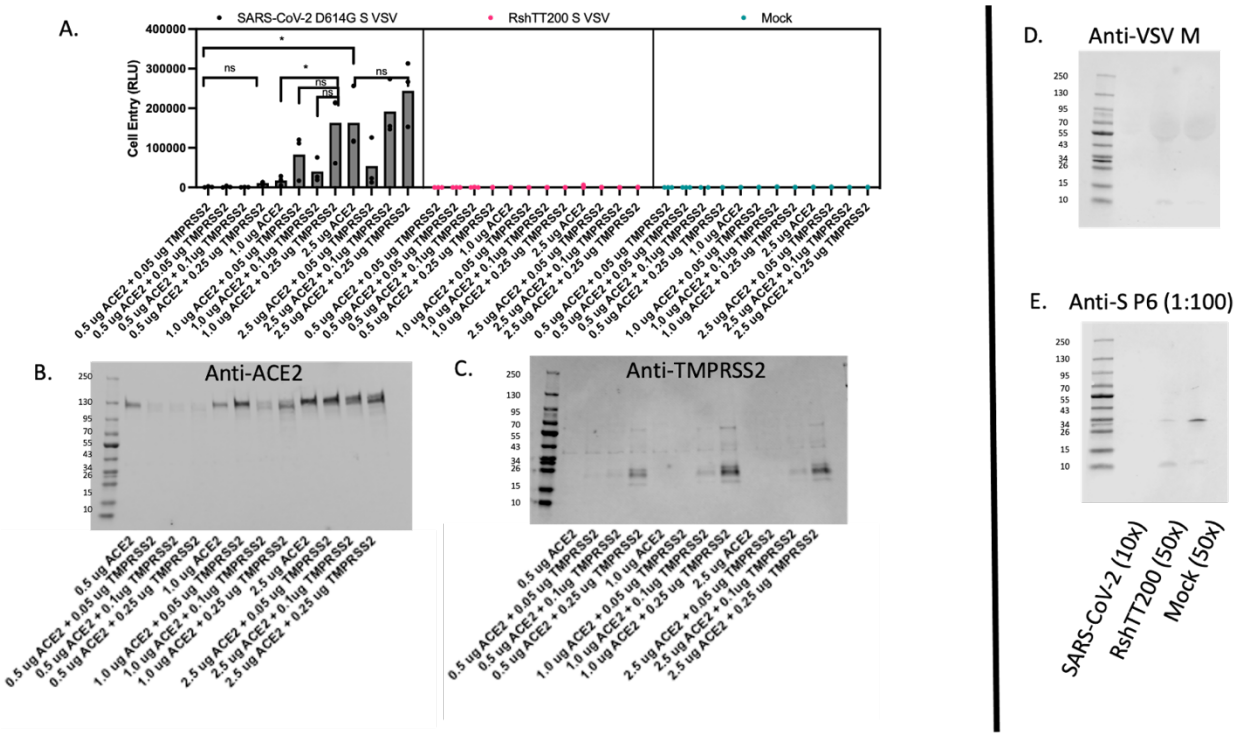
Figure 3.7: RshTT200 RBD in the SARS-CoV-1 S enables human ACE2 mediated cellular entry due to RBD conformation

SARS-CoV-1, SARS-CoV-2, RshTT200 WT, and Chimeric swapped RBD S (SARS-CoV-1 RBD in RshTT200 S; SARS-CoV-2 RBD in RshTT200 S, RshTT200 RBD in SARS-CoV-1 S, RshTT200 RBD in SARS-CoV-2 S pseudotyped VSV particles normalized for S (**Fig. S3.14**) entry in HEK293T cells with stable expression of human ACE2 (A). RshTT200 RBD in SARS-CoV-1 S ectodomain cryo-EM dataset 3D classifications reveal a mix of closed (B) and open (C) RBD conformations. Monoclonal antibody neutralization of RshTT200 RBD in SARS-CoV-1 S, and SARS-CoV-2 S pseudotyped VSV particles with S2X259 (D) and S2X35 (E).

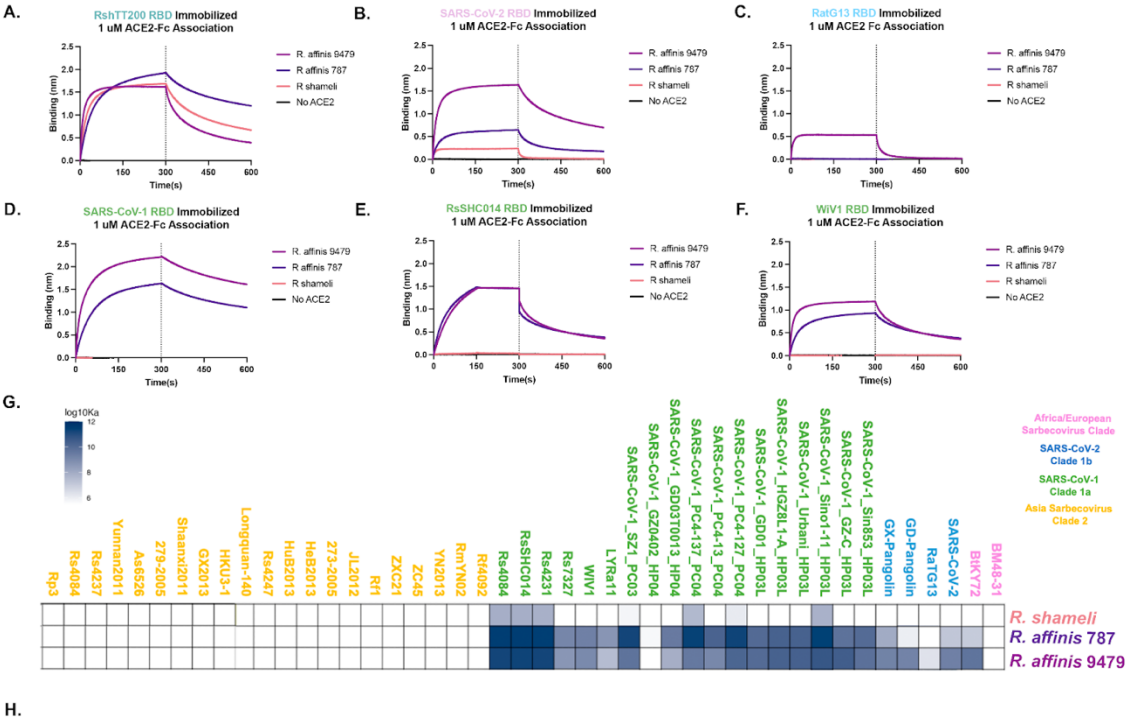


Supplemental Figure 3.20: *D614G mutation has no impact on cell entry in RshTT200 S.*

RshTT200 WT, RshTT200 D614G, and SARS-CoV-2 D614G entry in HEK293T cells with transient transfection of R. shameli ACE2 (A) or R. shameli ACE2 (B). VSV pseudotyped particles were normalized by S (C).



Supplemental Figure 3.21*:* Co-transfection of TMPRSS2 does not enable human ACE2 cellular entry by RshTT200 VSV.
A) HEK293T cells co-transfected in 6-well plate with varying levels of ACE2 and TMPRSS2 demonstrate were infected with VSV pseudotyped SARS-CoV-2 S, RshTT200 S, or mock particles. B) Western blot analysis of HEK293T ACE2 expression C) Western blot analysis of TMPRSS2 expression D) Western blot analysis of VSV particles during infection E) Western blot analysis of S in pseudotyped particles.



Supplemental Figure 3.22*:* Limited Promiscuity of *R. shameli* ACE2 across sarbecovirus clades.

CHAPTER 4. BROAD RECEPTOR TROPISM AND IMMUNOGENICITY OF A CLADE 3 SARBECOVIRUS

In this next story, I continue with our assessment of spillover prevention and preparedness through the characterization of a sarbecovirus that is more distal to SARS-CoV-2 then the one detailed in the previous story. Although, *Rhinolophus* bats harbor diverse clade 3 sarbecoviruses, the structural determinants of receptor tropism along with the antigenicity of their spike (S) glycoproteins had remain uncharacterized. Here, I describe how the African *Rhinolophus* bat clade 3 sarbecovirus PRD-0038 S has a broad angiotensin-converting enzyme 2 (ACE2) usage and that receptor-binding domain (RBD) mutations further expand receptor promiscuity and enable human ACE2 utilization. I detail our PRD-0038 RBD bound to *Rhinolophus alcyone* ACE2, explaining receptor tropism and highlighting differences with SARS-CoV-1 and SARS-CoV-2. I characterize the PRD-0038 S using cryo-EM and monoclonal antibody reactivity revealing its distinct antigenicity relative to SARS-CoV-2 and identifies PRD-0038 cross-neutralizing antibodies for pandemic preparedness. Finally, I discuss PRD-0038 S vaccination, which elicits greater titers of antibodies cross-reacting with vaccine-mismatched clade 2 and clade 1a sarbecoviruses compared with SARS-CoV-2 S due to broader antigenic targeting, motivating the inclusion of clade 3 antigens in next-generation vaccines for enhanced resilience to viral evolution.

Adapted From:

Lee, J.*, **Zepeda, S.K**.*, Park, Y., Taylor, A.L., Quispe, J., Leaf, E.M., Treichel, C., Corti, D., King, N.P., Starr, T.N., Veesler, D. (2023) Broad receptor tropism and immunogenicity of a clade 3 sarbecovirus. Cell Host and Microbe. 31(12):1961-1973.e11

4.1 Chapter Introduction

Two sarbecoviruses have crossed the species barrier and spilled over to humans in the past two decades. SARS-CoV-1 emerged in 2002 and spread worldwide through air travel routes, leading to an epidemic with 8,098 cases and 774 deaths (Drosten, 2003; Ksiazek, 2003). SARS-CoV-2 emerged at the end of 2019 and led to the devastating COVID-19 pandemic that claimed millions of lives worldwide(Zhou, 2020; Zhu, 2020). Both viruses enter human cells via spike (S)-mediated fusion of the viral and host membranes upon binding to the angiotensin-converting enzyme 2 (ACE2) receptor (Zhou, 2020; Walls, 2020; Letko, 2020; Hoffmann, 2020; Li, 2003).

Reports of additional sarbecovirus spillovers to humans(Evans, 2023; Wang, 2018) along with detection of numerous sarbecoviruses in bats and other wild animals (Zhou, 2020; Wacharapluesadee, 2021; Temmam, 2022; Lam, 2020; Ge, 2013; Hu, 2017; Yang, 2015) underscore the recurrent zoonotic threat to public health posed by these viruses. The S glycoprotein of some of these sarbecoviruses harbors a receptor-binding domain (RBD) that utilizes the human ACE2 receptor to enter host cells, indicating that they could possibly cross the species barrier to infect humans (Letko, 2020; Temmam, 2022; Wrobel, 2021; Niu, 2022, Menachery, 2016; Menachery, 2015).

Phylogenetic classification of sarbecoviruses based on their RBD sequences led to the definition of at least four clades: clade 1a (e.g., SARS-CoV-1), clade 1b (e.g., SARS-CoV-2), clade 2 (e.g., RmYN02), and clade 3 (e.g., BtKY72)(Starr, 2021; Tortorici, 2021). Clade 3 sarbecoviruses have been identified in bats in Europe and Africa,(Tong, 2009; Tao, 2019; Wells, 2021;Drexler, 2010; Crook, 2021; Alkhovsky, 2022) such as BtKY72 and PRD-0038 for which sequences were found in Kenya and Rwanda, respectively. We recently showed that the S glycoprotein of one of them

(BtKY72) could utilize two *Rhinolophus affinis* (*R. affinis*) ACE2 alleles to promote entry into cells (Starr, 2022). Furthermore, two amino acid residue substitutions in the BtKY72 RBD enabled S-mediated entry into human ACE2-expressing cells, broadening the range of sarbecoviruses with spillover potential(Starr, 2022). The importance of this observation was further underscored by the recent discovery of the clade 3 Khosta-2 virus, (Alkhovsky, 2022) which independently acquired the ability to bind(Starr, 2022) and enter cells (Seifert, 2022) using the human ACE2 receptor. Studying the structure and functional properties of clade 3 sarbecovirus S glycoproteins is therefore crucial to understand spillover risk and assist in pandemic preparedness.

Here, we report that the S glycoprotein of the clade 3 sarbecovirus PRD-0038, which is a member of the largely uncharacterized African bat-borne sarbecoviruses, has a broad ACE2 usage and that PRD-0038 RBD mutations further expand entry receptor tropism to additional *Rhinolophus* bat species and human ACE2. We determined structures of the PRD-0038 RBD bound to *Rhinolophus alcyone* (*R. alcyone*) ACE2 and of the PRD-0038 S trimer, explaining receptor tropism and the distinct antigenicity of clade 3 sarbecoviruses relative to SARS-CoV-2 and SARS-CoV-1. Evaluation of a panel of monoclonal antibodies enabled identification of PRD-0038 cross-neutralizing antibodies that could be deployed for outbreak response. Vaccination of mice with PRD-0038 S elicited greater titers of antibodies cross-reacting with vaccine-mismatched clade 2 and clade 1a sarbecoviruses, relative to SARS-CoV-2 S immunization, indicating that addition of clade 3 antigens in vaccine formulations could enhance the resilience of antibody responses to viral evolution. Our findings highlight a molecular

pathway for possible zoonotic spillover of a clade 3 sarbecovirus and the necessity of developing pan-sarbecovirus vaccines and countermeasures.

4.2 PRD-0038 S can utilize a broad spectrum of Rhinolophus bat ACE2 orthologs as entry receptors

To investigate the promiscuity of clade 3 sarbecovirus host receptor usage, we first assessed binding of a panel of *Rhinolophus* bat ACE2 orthologs harboring a C-terminal human Fc fusion to the immobilized PRD-0038 RBD using biolayer interferometry (BLI) (**Figs. 4.1A and 4.1B**). We selected PRD-0038 as a representative member of African bat-borne sarbecoviruses due to its more ancestral phylogenetic positioning relative to the other two sarbecoviruses isolated on the same continent[29] (BtKY72 and PDF-2370) and the high sequence similarity of their S glycoproteins. Our ACE2 panel comprised eight distinct *Rhinolophus sinicus* (*R. sinicus*) alleles and two distinct *R. affinis* alleles, which were defined based on polymorphisms within the region recognized by sarbecovirus RBDs, (Starr, 2022; Guo, 2020) as well as *R. alcyone* and *Rhinolophus landeri*(*R. landeri*) orthologs. *R. sinicus* Asian bats are probable reservoir hosts for SARS-CoV-1, (Ge, 2013) *R. affinis* bats have been shown to host closely related viruses to SARS-CoV-2 (Zhou, 2020), whereas *R. alcyone* and *R. landeri* bats are found in sub-Saharan Africa, overlapping with the regions of sampling of several clade 3 sarbecoviruses (the exact *Rhinolophus* species from which PRD-0038 and BtKY72 have been sampled is unknown)(Tao, 2019; Wells; 2021) (**Fig. 4.1C**). We observed the strongest binding to the PRD-0038 RBD with *R. alcyone* ACE2, which exhibited the slowest dissociation kinetics in our panel (**Fig. 4.1A**). The PRD-0038 RBD also interacted with both *R. affinis* ACE2s, albeit more tightly with the 9479 allele than the 787 allele, and

with *R. landeri* ACE2 (**Fig. 4.1A**). Finally, we detected binding to four out of the eight *R. sinicus* alleles evaluated (**Fig. 4.1B**).

To further evaluate receptor tropism, we pseudotyped vesicular stomatitis virus (VSV) particles with PRD-0038 S and assessed entry into HEK293T cells transiently transfected with the corresponding set of full-length, membrane-anchored ACE2s (Fig. 1D). Concurring with our binding data, *R. alcyone* enabled robust entry of PRD-0038 S VSV (**Fig. 4.1D**), in line with a previous study (Roelle, 2022). Moreover, we also detected efficient entry into cells expressing *R. affinis* 9479 and 787 ACE2 alleles, *R. landeri* ACE2, and the four *R. sinicus* ACE2 alleles for which binding was detected by BLI (**Fig. 4.1D**). Collectively, these data show that PRD-0038 S recognizes and can utilize a broad spectrum of *Rhinolophus*bat ACE2 orthologs as entry receptors, including those from bat species known to be found in geographic areas proximal to the site of PRD-0038 discovery.

4.3 Molecular basis of PRD-0038 RBD engagement of the *R. alcyone* ACE2 receptor

To reveal the structural determinants of ACE2 recognition by clade 3 sarbecoviruses, we determined a reconstruction of the PRD-0038 RBD bound to the natively dimeric *R. alcyone* ACE2 (RaACE2) ectodomain using single-particle cryoelectron microscopy (cryo-EM) (**Figs. 4.1E and S4.1; Table S4.2**). Symmetry expansion and local refinement yielded a structure at 3.2 Å resolution of the ACE2 peptidase domain bound to the RBD revealing the molecular interactions mediating complex formation. An average surface of ~750 Å (Ksiazek, 2003) is buried at the ACE2/RBD interface as compared with ~840 Å (Ksiazek, 2003) for the complexes of human ACE2 (hACE2) bound to the SARS-CoV-2 RBD (Lan, 2020) or to the SARS-

CoV-1 RBD (Li, 2005). The relative orientation of the binding partners is similar for these three structures, likely due to the conservation of several key ACE2-interacting residues including $L444_{PRD\text{-}0038}$/$L455_{SARS\text{-}CoV\text{-}2}$ and $F445_{PRD\text{-}0038}$/$F456_{SARS\text{-}CoV\text{-}2}$, $L475_{PRD\text{-}0038}$/$L472_{SARS\text{-}CoV\text{-}1}$, $Y478_{PRD\text{-}0038}$/$Y489_{SARS\text{-}CoV\text{-}2}$/$Y475_{SARS\text{-}CoV\text{-}1}$, $G485_{PRD\text{-}0038}$/$G496_{SARS\text{-}CoV\text{-}2}$/$G482_{SARS\text{-}CoV\text{-}1}$, $G491_{PRD\text{-}0038}$/$G502_{SARS\text{-}CoV\text{-}2}$/$G488_{SARS\text{-}CoV\text{-}1}$, and $Y494_{PRD\text{-}0038}$/$Y505_{SARS\text{-}CoV\text{-}2}$/$Y491_{SARS\text{-}CoV\text{-}1}$ (**Figs. 4.1F–4.1I**). However, the salt bridge formed between $K417_{SARS\text{-}CoV\text{-}2}$ and $D30_{hACE2}$ is absent due to substitution to $V408_{PRD\text{-}0038}$ and $V404_{SARS\text{-}CoV\text{-}1}$ (**Figs. 4.1F, 4.1G, and 4.1J**). Furthermore, the electrostatic interactions involving residues $D38_{hACE2}$ and $Q42_{hACE2}$ with $Y449_{SARS\text{-}CoV\text{-}2}$ and $Q498_{SARS\text{-}CoV\text{-}2}$ or $Y436_{SARS\text{-}CoV\text{-}1}$ and $Y484_{SARS\text{-}CoV\text{-}1}$ are lost due to truncation and remodeling of the PRD-0038 433–440 loop along with substitution of $Q498_{SARS\text{-}CoV\text{-}2}$ to $T487_{PRD\text{-}0038}$ (**Figs. 4.1F–H, 4.1K, and S4.2**). The interface between T500/$N501_{SARS\text{-}CoV\text{-}2}$ (or Y501 in currently circulating variants) and $Y41_{hACE2}$ is replaced by tenuous contacts between the topologically equivalent residues T489/$V490_{PRD\text{-}0038}$ and $H41_{RaACE2}$ (**Figs. 4.1F, 1G, and 4.1K**). $Q493_{SARS\text{-}CoV\text{-}2}$ optimally interacts with $K31_{hACE2}$/$E35_{hACE2}$, whereas the topologically equivalent residue $K482_{PRD\text{-}0038}$ is better adapted to $N31_{RaACE2}$/$E35_{RaACE2}$ due to swapping of the position of a positively charged amino acid side chain across the interface (**Figs. 4.1F, 4.1G, and 4.1L**). These findings concur with (1) the enhanced entry of the closely related BtKY72 RBD-harboring pseudovirus into cells expressing a human ACE2 K31D mutant relative to wild-type ACE2 (Roelle, 2022); (2) the isolation of a mouse-adapted SARS-CoV-2 isolate harboring the Q493K substitution (Leist, 2020) that promotes favorable interactions with mouse ACE2, the latter ortholog also possessing residues N31/E35 (Cameroni, 2021;

McCallum, 2022); and (3) the emergence of $R493_{SARS\text{-}CoV\text{-}2}$ in Omicron BA.1 and BA.2 (Cameroni, 2021; McCallum, 2022), which was subsequently reverted to the more favorable $Q493_{SARS\text{-}CoV\text{-}2}$ in subsequent variants, likely due to relieving electrostatic repulsion with $K31_{hACE2}$ (Starr, 2022). Overall, most of the PRD-0038 RBD binding interface with *R. alcyone* ACE2 is remodeled as compared with human ACE2 bound to SARS-CoV-2 or SARS-CoV-1, thereby explaining the shift in receptor species tropism.

4.4 PRD-0038 RBD mutations enable human ACE2 utilization and expand receptor tropism to additional geographically relevant bat species

To investigate how viral evolution could alter the PRD-0038 receptor species tropism for pandemic preparedness, we evaluated the impact of RBD mutations on utilization of several ACE2 orthologs. Using BLI, we observed that human ACE2-Fc did not bind to the wild-type PRD-0038 RBD (**Fig. 4.2A**). However, we found that two amino acid residue substitutions promoted binding of human ACE2-Fc to the immobilized PRD-0038 K482Y/T487W RBD mutant (**Fig. 4.2A**, SARS-CoV-2 numbering 493Y/498W) (Starr, 2022), in line with their ability to enable BtKY72 S-mediated utilization of human ACE2 (Starr, 2022). Moreover, we found that human ACE2-Fc interacted with the T487W RBD, albeit weakly (**Fig. 4.2A**). We validated these findings using deep mutational scanning (DMS) of the yeast-displayed PRD-0038 RBD showing that T487W is the only single amino acid mutation enabling detection of human ACE2-Fc binding (**Figs. 4.2B and S4.3**). The geographically relevant *R. alcyone* and *R. landeri* ACE2 orthologs exhibited enhanced binding to the PRD-0038 T487W RBD relative to the wild-type RBD, with 1:1 binding affinity improvements corresponding to 2- and greater than 1 order of magnitude, respectively (**Figs. 4.2C and 4.2D; Table S4.3**). These data concur with our DMS measurements showing that T487W mutation had the most marked

positive effect on *R. landeri* ACE2 binding (**Figs. 4.2B and S4.3**). Furthermore, the PRD-0038 T487W RBD mutation enhanced binding to the *R. affinis* 787 allele as well as to the *R. sinicus* alleles recognized by the wild-type PRD-0038 RBD and enabled detectable binding to the *R. sinicus* WJ1 allele (**Figs. S4.4A and S44.B**). By contrast, the K482Y mutation was deleterious for binding to *R. alcyone*, *R. landeri*, *R. affinis*, and *R. sinicus* ACE2 orthologs (**Figs. 4.2E, 4.2F, S3, S4.4C, and S4.4D**). We observed that wild-type PRD-0038 S and all three mutants (K482Y, T487W, and K482Y/T487W) promoted entry of VSV pseudotypes in HEK293T cells transiently transfected with *R. alcyone* ACE2 or with *R. landeri* ACE2 except for PRD-0038 K482Y S VSV that did not enter *R. landeri* ACE2-expressing cells (**Figs. 4.2G, 4.2H, and S4.4E),** suggesting that binding avidity could overcome to some extent the observed differences in affinity. To broaden our understanding of PRD-0038 tropism, we also examined cell entry promoted by transient transfection of ACE2 alleles from *Rhinolophus ferrumequinum* (*R. ferrumequinum*) bats, which are found in northern Africa, southern Europe, and southeast Asia. We observed that *R. ferrumequinum* ACE2 allele XM_033107295.1 promoted entry of wild-type and T487W PRD-0038 S VSV, whereas *R. ferrumequinum* ACE2 allele FJ598617.1 enabled entry of T487W and even more so of K482Y/T487W PRD-0038 S VSV (**Figs. 4.2I and 4.2J**). Finally, we observed entry of PRD-0038 T487W S and even more so of PRD-0038 K482Y/T487W S VSV pseudoviruses in HEK293T cells stably expressing human ACE2 (**Fig. 4.2K**). These findings indicate that a single RBD residue mutation is sufficient for broadening the *Rhinolophus* bat ACE2 tropism and for enabling PRD-0038 S-mediated entry into

cells overexpressing the human ACE2 receptor, highlighting the possible future zoonotic risk of this virus and related clade 3 sarbecoviruses.

4.5 Architecture of the PRD-0038 S trimer

The overall architecture of the PRD-0038 S trimer is similar to that of SARS-CoV-1 and SARS-CoV-2, and a PRD-0038 S protomer can be superimposed with root-mean-square deviation values of 4.9 and 7.2 Å to SARS-CoV-1 S and SARS-CoV-2 S with which it shares 75% and 72% amino acid sequence identity, respectively (**Fig. S4.6**). Glycosylation at RBD residue N360, which is present in most sarbecoviruses except SARS-CoV-2 (position N370), has been reported to favor the closed S conformation (Harbison, 2022; Zhang, 2022), which could possibly explain the sole observation of closed PRD-0038 S trimers in our cryo-EM dataset. However, glycosylation at this site is also present in SARS-CoV-1, which spontaneously adopts open RBD conformations (Walls, 2019; Kirchdoefer, 2018; Yuan, 2017). The closed PRD-0038 S RBD conformation is most similar to the linoleic acid (LA)-bound form of SARS-CoV-2 S previously described(Toelzer, 2020) (**Figs. 4.3D and S64.**). Indeed, the conformation of the RBD helix containing residues 354–359 (equivalent to SARS-CoV-2 residues 364–369) closely resembles that of the linoleic acid-bound SARS-CoV-2 S (**Fig. 4.3D**) or that of F371-harboring SARS-CoV-2 Omicron variants (Park, 2022; Addetia, 2023; Stalls, 2022; Cao, 2022). This conformation, however, is distinct from that observed in the structure of the isolated PRD-0038 RBD bound to *R. alcyone* ACE2 described above (**Fig. 4.3E**). No linoleic acid density is resolved in our cryo-EM map, although the RBD pocket that accommodates this ligand is conserved in the PRD-0038 S structure, including residues R399 and Q400 (equivalent to SARS-CoV-2 R408 and Q409 forming electrostatic interactions with the linoleic acid carboxylate). Instead, we found that the

Y355 side chain (equivalent to SARS-CoV-2 Y365) partially obstructs the hydrophobic pocket, which would have otherwise been occupied by the linoleic acid hydrocarbon tail (**Fig. 4.3F**). The PRD-0038 S Y355 side chain rotamer resembles that of apo SARS-CoV-2 S (PDB: 6VXX (Walls, 2020)) and apo SARS-CoV-1 S (PDB: 5X58 (Yuan, 2017)), suggesting that this rotameric configuration is accessible to SARS-CoV-2 Y365 and likely changes to allow linoleic acid binding (Toelzer, 2020).

4.6 Antigenicity of the PRD-0038 S trimer

To define the antigenic landscape of clade 3 sarbecoviruses, we probed binding of a panel of monoclonal antibodies with broadly neutralizing activity against sarbecoviruses (Starr, 2021, Tortorici, 2021; Park, 2022; Piccoli, 2020; Park, 2022) and α- and β-coronaviruses (Silva, 2023; Pinto, 2021; Sauer, 2021; Low, 2022) to prefusion-stabilized PRD-0038 S harboring the HexaPro mutations (Hseih, 2020) using an enzyme-linked immunosorbent assay (ELISA) (**Fig. 4.3G**). We found that S2X259(Tortorici, 2021) and S2X35(Piccoli, 2020) (antigenic site II) as well as S2H97 (Starr, 2021) (antigenic site V) cross-reacted with PRD-0038 S, whereas S309 (Pinto, 2020) (antigenic site IV) bound very weakly (**Fig. 4.3G**). The markedly dampened S309 binding likely results from the $E340_{SARS\text{-}CoV\text{-}2}$/$Q330_{PRD\text{-}0038}$ escape substitution previously identified by DMS of the yeast-displayed SARS-CoV-2 Wuhan-Hu-1 RBD (Starr, 2021) (**Fig. S4.7**). The S2X324 (Park, 2022) antibody (antigenic site Ib), which neutralizes a broad panel of SARS-CoV-2 variants and resembles LY-COV1404 (Westendorf, 2022), recognizes the SARS-CoV-2 437–448 loop, which is truncated and remodeled in the PRD-0038 RBD, explaining the lack of binding observed (**Fig. S4.7**). The S2K146 (Park, 2022) antibody (antigenic site Ia), which contacts an epitope sharing several residues with the ACE2-binding site,

bound very weakly to wild-type PRD-0038 S, whereas S2M11 (Tortorici, 2020) (antigenic site Ia) did not bind at all, as a result of extensive RBM (receptor-binding motif) mutations (**Fig. S4.7**). Nevertheless, we previously showed that S2K146 weakly neutralized VSV pseudotyped with BtKY72 S (clade 3) harboring the K482Y/T487W mutations (SARS-CoV-2 numbering 493Y/498W) (Park, 2022), underscoring the possible usefulness of this antibody if such mutations arose in related clade 3 sarbecoviruses.

The stem helix-targeting S2P6 (Pinto, 2021) antibody and the RAY53 (Silva, 2023) antibody (recognizing the fusion machinery apex) cross-reacted as efficiently with PRD-0038 S as they did with SARS-CoV-2 S, whereas the fusion peptide-directed 76E1 antibody did not, possibly as a result of the $F823_{SARS\text{-}CoV\text{-}2}Y806_{PRD\text{-}0038}$ epitope mutation (which is shared with other clade 3 sarbecoviruses) along with the presence of the $F800P_{PRD\text{-}0038}$ HexaPro stabilizing mutation (Low, 2022) (**Figs. 4.3G and S4.7**).

Consistent with the ELISA data, we found that PRD-0038 K482Y/T487W S VSV pseudovirus was neutralized in a concentration-dependent manner by S2X259, S2X35, and more weakly by 76E1, and the activity of the latter antibody is likely explained by the absence of the $F800P_{PRD\text{-}0038}$ HexaPro stabilizing mutation in our pseudovirus construct (**Figs. 4.3G and 4.3H**). Collectively, these data show that monoclonal antibodies targeting RBD antigenic sites or fusion machinery epitopes that are conserved across sarbecoviruses or α- and β- coronaviruses, respectively, retain neutralizing activity against PRD-0038 and are possible candidates for pandemic preparedness.

4.7 Immunogenicity of the PRD-0038 S trimer

To better understand the immunogenicity of clade 3 sarbecoviruses and the impact of their possible inclusion in next-generation vaccine candidates, we immunized groups of six mice with three 1 µg doses of either PRD-0038 S or SARS-CoV-2 S (**Fig. 4.4A**), both stabilized in the prefusion conformation using the HexaPro mutations (Hseih, 2020). Serum-neutralizing activity, expressed as half-maximum inhibition dilution (ID_{50}), was analyzed 2 weeks post dose 3 using VSV particles pseudotyped with clade 1a (SARS-CoV-1), clade 1b (SARS-CoV-2/G614, BA.2, BA.5, RaTG13) or clade 3 (PRD-0038 and Khosta-1) S glycoproteins. SARS-CoV-2 S-immunized mice had the strongest serum-neutralizing activity against SARS-CoV-2/G614 VSV S (ID_{50} = 8,730, vaccine-matched), which was reduced against BA.2 S VSV (ID_{50} = 247) and even more so against BA.5 S VSV (ID_{50} = 71) (**Figs. 4.4B and S4.8**). RaTG13 S VSV (vaccine-mismatched), however, was neutralized with almost comparable potency (ID_{50} = 7,752) to that against SARS-CoV-2/G614 S VSV (ID_{50} = 8,730), whereas no neutralization of PRD-0038 S VSV and Khosta-1 S VSV clade 3 sarbecoviruses (vaccine-mismatched) was detected except for a subset of animals with very weak neutralization (mouse 1-1, 1-3, 1-5) (**Figs. 4.4B and S4.8**). PRD-0038 S-immunized mice had the strongest serum-neutralizing activity against PRD-0038 S VSV (ID_{50} = 3,634, vaccine-matched) and Khosta-1 S VSV (ID_{50} = 863, vaccine-mismatched), whereas we could not detect any neutralization of clade 1a and 1b pseudoviruses tested besides RaTG13 inhibition (ID_{50} = 928) (**Figs. 4.4B and S4.8**). Although none of the sera could block SARS-CoV-1 S-mediated entry into target cells in standard experimental conditions, we observed very weak SARS-CoV-1 S VSV neutralization with greater dilution of the pseudovirus

stock when using PRD-0038 S- but not SARS-CoV-2 S-elicited sera (**Fig. S4.9**). These results suggest that this clade 3 S trimer immunogen induced slightly more broadly neutralizing antibody responses than SARS-CoV-2 S against a clade 1a pseudovirus. Serum binding titers positively correlate with neutralization potency (Piccoli, 2020; Park, 2022), and the RBD is the main target of neutralizing antibodies against vaccine/infection-matched and mismatched viruses (Piccoli, 2020; Bowen, 2022; Greaney, 2021; Stamatatos, 2021). We analyzed binding of selected vaccine-elicited sera to a panel of yeast-displayed RBDs spanning the known sarbecovirus phylogenetic diversity using two mice immunized with PRD-0038 S and four mice immunized with SARS-CoV-2 S. We selected two mice with greatest (mouse 1-3, 1-5) and two mice with the weakest (mouse 1-4, 1-6) serum-neutralizing activity against our panel of pseudoviruses from the group immunized with SARS-CoV-2 HexaPro S as well as two mice with the greatest SARS-CoV-1 cross-neutralization (mouse 2-3, 2-6) from the group immunized with PRD-0038 HexaPro S (**Figs. S4.8 and S4.9).** In line with the serum neutralization data, inclusion of an antigen in the vaccine formulation was associated with strong cross-reactivity with vaccine-matched and related antigens within the same clade among the sera analyzed (**Figs. 4.4C and S4.10**). Furthermore, PRD-0038 S immunization elicited greater titers of antibodies cross-reacting with clade 2 and clade 1a RBDs, as well as the RsYN04 RBD that branches independently of the four previously known clades (Zhou, 2021), likely explaining the weak but detectable SARS-CoV-1 neutralization (**Figs. 44.B, 4C, and S4.9).** These data indicate that inclusion of a clade 3 antigen in a vaccine formulation could not only elicit clade 3 serum-neutralizing activity but also enhance cross-reactive (and weakly neutralizing) antibody responses

against vaccine-mismatched antigens from distinct clades, which could participate in protection through direct neutralization and Fc-mediated effector functions (Addetia, 2023; Schäfer, 2021; Winkler, 2021; Adams, 2023; Mackin, 2023).

To understand the molecular basis for variation in breadth of serum cross-reactivity across mice and vaccine regimen, we mapped the dominant epitope specificities in these six sera using yeast-displayed DMS libraries in the vaccine-matched SARS-CoV-2 Wuhan-Hu-1 and PRD-0038 RBDs as well as in the vaccine-mismatched SARS-CoV-2 Omicron BA.2 and SARS-CoV-1 Urbani RBDs (**Figs. 4.4D and S4.11**). SARS-CoV-2 S-elicited polyclonal serum antibodies predominantly targeted the SARS-CoV-2 Wuhan-Hu-1 RBD region comprising residues 484–490, consistent with the strong antigenic pressure on this position that drove early variant evolution at residue E484 during the COVID-19 pandemic (Tegally, 2021; Freitas, 2021; Fuijino; 2021). These residues are highly variable among sarbecoviruses, consistent with the weaker cross-reactive breadth seen in these sera. By contrast, PRD-0038 S-elicited polyclonal serum antibodies showed no dominant targeting of specific PRD-0038 RBD antigenic sites, which could indicate a more balanced binding antibody response that would be less susceptible to single amino acid mutations. Epitope mapping to vaccine-mismatched RBDs revealed that PRD-0038 and SARS-CoV-2 S-elicited sera contained antibodies targeting antigenic sites II, IV, and V (Starr, 2021; Piccoli, 2020), which are typically recognized by subdominant broadly-reactive and neutralizing antibodies (Starr, 2021; Totorici, 2021; Pinto, 2020; Jette, 2021; Martinez, 2021) (**Fig. S4.11**). However, DMS using vaccine-mismatched RBD strains also revealed the presence of antibodies in SARS-CoV-2 S-elicited sera that target suboptimal sites for sarbecovirus breadth. For

example, we observed strong antibody responses to the N370 glycan hole (resulting from mutations of this glycosylation sequon in our DMS experiments) for two of the four SARS-CoV-2-immunized mice analyzed. Antibody responses targeting this site are elicited as a result of the absence of this oligosaccharide in the SARS-CoV-2 RBD immunogen but its presence in all other sarbecovirus RBDs. Restoration of the N370 glycan in SARS-CoV-2 vaccines could therefore possibly limit these off-target responses by reducing the elicitation of antibodies to this strain-specific epitope. In conclusion, the DMS data indicate that PRD-0038 S-elicited serum antibodies analyzed here target a broader spectrum of antigenic sites present on vaccine-matched and mismatched RBDs than SARS-CoV-2 S-elicited antibodies, providing a molecular basis for enhanced cross-reactivity (**Figs. 4.4D and S4.11**).

4.8 Chapter Discussion

Coronavirus S glycoproteins are evolutionary hotspots and can acquire amino acid substitutions, insertions, deletions, or even recombine distinct domains (Roemer, 2022). These mutational changes can alter host receptor tropism (Xiong, 2022), binding affinity(Starr, 2022; Bowen, 2022; Starr, 2020), entry route (Addetia, 2023 ;Meng, 2022; Willet), and immune evasion (Cameroni, 2021; McCallum, 2022; Bowen, 2022; Starr, 2022; McCallum, 2021; McCallum, 2021). Most mutations occur within the RBD, which engage the host receptor and account for most of the neutralizing activity against vaccine/infection-matched and mismatched sarbecoviruses (Piccoli, 2020; Bowen, 2022; Greaney, 2021; Stamatatos, 2021).

Spillover is a complex process involving multiple factors such as receptor recognition, proteolytic S activation, immune antagonism, and contact opportunity. To

examine potential spillover pathways of clade 3 sarbecoviruses, we evaluated binding of PRD-0038 RBD mutants to a panel of *Rhinolophus* bat ACE2 orthologs. Although human ACE2 cannot serve as entry receptor for wild-type PRD-0038 S (or the closely related BtKY72 S (Starr, 2022), introduction of a single amino acid RBD mutation (T487W) enabled binding and S-mediated entry into cells expressing human ACE2. Moreover, this point mutation broadened host receptor tropism by enabling utilization of the geographically relevant *R. landeri* and *R. ferrumequinum* ACE2 receptors without compromising binding to *R. alcyone*, *R. sinicus*, and *R. affinis* ACE2 alleles. Although the T487W RBD mutation requires 3 nucleotide substitutions, these findings point to a possible spillover pathway in which a single amino acid change expands host receptor tropism markedly. Indeed, acquisition of *R. landeri*, *R. ferrumequinum* (allele FJ598617.1), and human ACE2 tropism would allow PRD-0038 (and likely BtKY72 and related clade 3 viruses) to expand the geographic range of host reservoirs that can be infected and, in turn, the likelihood of zoonotic transmission. Broadly neutralizing antibodies with activity against sarbecoviruses (RBD-directed) and beyond (fusion machinery-directed) inhibited PRD-0038 S-mediated entry into cells and could be stockpiled as possible countermeasures for pandemic preparedness.

We observed that three immunizations with a clade 1b or with a clade 3 sarbecovirus S trimer predominantly elicited vaccine-matched serum-neutralizing antibody responses. We note that our neutralization data of clade 1b pseudoviruses underscore the distinction between antigenic and genetic distance: although RaTG13 harbors a greater number of RBD mutations than Omicron BA.2 or BA.5, neutralizing activity was higher against RaTG13 than against these SARS-CoV-2 variants, which accumulated

mutations to erode neutralizing antibody titers (Cao, 2023; Tan, 2022). Compared with SARS-CoV-2 S, we found that PRD-0038 S-elicited polyclonal serum antibodies were more broadly reactive with vaccine-mismatched antigens, including clades 1a and 2 RBDs, which likely account for the weak but detectable cross-neutralization of SARS-CoV-1 with sera from mice immunized with PRD-0038 S but not with SARS-CoV-2 S. However, we also note that sera from three SARS-CoV-2 S-immunized mice had weak but detectable neutralization against clade 3 sarbecoviruses, whereas none of the PRD-0038 S-elicited sera could neutralize the SARS-CoV-2 S pseudoviruses tested. As non- or weakly neutralizing monoclonal and polyclonal antibodies have been shown to participate in protection against SARS-CoV-2 challenge in small animal models through Fc-mediated effector functions (Addetia, 2023; Schäfer, 2021; Winkler, 2021; Adams, 2023; Mackin, 2023), our findings motivate the inclusion of clade 3 and other divergent RBDs in updated multi-antigen sarbecovirus vaccine formulations (Walls, 2021; Cohen, 2022; Cohen, 2021; Martinez, 2021). This would allow the elicitation of potent clade 3 neutralizing antibodies and cross-clade binding (and possibly neutralizing) antibodies with maximal breadth to achieve optimal protection against continuously evolving SARS-CoV-2 variants and sarbecoviruses found in wildlife.

4.9 Methods

Cell lines

Cell lines used in this study were obtained from DH10B competent cells (Thermo Fisher Scientific), HEK293T (ATCC, CRL-11268) and Expi293F (Thermo Fisher Scientific, A14527) except for the HEK293T cells with stable human ACE2 expression which was kindly provided by Jesse Bloom (Crawford, 2020). Cells were cultured in 10% FBS

(Fisher Scientific-Cytiva), 1% penicillin-streptomycin (Thermo Fisher Scientific) DMEM at 37°C, 5% CO_2. AWY101 *S. cerevisiae* strain was used for yeast surface-display experiments. None of the cell lines were authenticated or tested for mycoplasma contamination.

In vivo animal studies

Female BALB/c mice were purchased from Envigo (order code 047) at 7 weeks of age and were maintained in a specific pathogen-free facility within the Department of Comparative Medicine at the University of Washington, Seattle, accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). Animal experiments were conducted in accordance with the University of Washington's Institutional Animal Care and Use Committee.

Production of recombinant PRD-0038 RBDs

The PRD-0038 S glycoprotein sequence was obtained from Genbank (MT726045). Both wildtype and mutant PRD-0038 RBDs (residues 318-520) were synthesized by GenScript with an N-terminal mu-phosphatase signal peptide and C-terminal 8x His tag, a short linker (GGSS) followed by an Avi tag in a pCMVR plasmid and sub-cloned with E. coli DH10B Competent Cells. Expi293F cells were grown at 37°C with 8% CO_2 and DNA transfections were conducted with the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatants were harvested three days post transfection. RBDs were purified using nickel based affinity chromatography using HisTrap™ High Performance column (Cytiva). Proteins were first washed with 10 column volumes of a buffer containing 25mM sodium phosphate (pH8.0) and 300 mM NaCl, before elution with 8 column volumes of a buffer containing 25 mM sodium

phosphate (pH8.0), 300 mM NaCl, 500mM imidazole. Eluted proteins were buffer exchanged into 1x PBS (137mM NaCl, 2.7mM KCl, 10mM Na2HPO4, 1.8mM KH2PO4, pH 8.0) using Amicon Ultra-15 Centrifugal Filter Unit (10 kDa) (Millipore). Overnight biotinylation reactions were performed using the BirA Biotin-Protein Ligase Kit (Avidity) at 4°C in 1x BiomixA, BiomixB. Biotinylated proteins were once again affinity purified using the HisTrap column as previously described to get rid of BirA. Once purified, buffer exchanged into PBS, and concentrated, proteins were flash-frozen and stored at -80°C until use.

Production of recombinant ACE2 ectodomains

Genbank accession numbers for all ACE2s can be found in **Table S1**. Recombinant ACE2 ectodomain constructs were synthesized by GenScript. ACE2-His ectodomain constructs comprise of residues 19–615 with an N-terminal mu-phosphatase signal peptide and C-terminal 8x His tag, a short GGSS linker, and an Avi tag. ACE2-Fc ectodomain[76] constructs comprise residues 19–615 with an N-terminal mu-phosphatase signal peptide and C-terminal fusion to a sequence encoding a thrombin cleavage site, a short linker (GGGG) and a human Fc fragment and were cloned in a pCMV plasmid. The native *R. alcyone* ACE2 ectodomain dimer comprises residues 1-740 and a C-terminal 8x His tag, a GGSS linker and an Avi tag and was cloned in a pCMVR plasmid. All ACE2 orthologue ectodomains were produced in Expi293F cells at 37°C supplemented and 8% CO_2. Transfections were performed with the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatants were harvested four days after transfection and proteins were purified using HiTrap Protein A HP (Cytiva) or HisTrap™ High Performance column (Cytiva). ACE2-Fc proteins were first

washed with 10 column volumes of 20 mM sodium phosphate (pH8.0) then eluted with 0.1 M citric acid (pH 3.0) directly into tubes containing 1M Tris-HCl (pH 9.0). Purified proteins were buffer exchanged into PBS (pH 8.0), concentrated using Spin-X® UF 20 mL Centrifugal Concentrator, 100,000 MWCO Membrane (PES) (Corning), and flash-frozen. ACE2-His proteins were washed with 10 column volumes of a buffer containing 25 mM sodium phosphate (pH 8.0), 300 mM NaCl, then eluted with 8 column volumes of a buffer containing 25 mM sodium phosphate (pH 8.0), 300 mM NaCl and 500 mM imidazole. Eluted proteins were buffer exchanged into 1x PBS (pH8.0) using Amicon Ultra-15 Centrifugal Filter Unit (10 kDa) (Millipore) and flash-frozen.

Production of recombinant PRD-0038 PentaPro S, HexaPro S, and SARS-CoV-2 HexaPro S

The PRD-0038 S glycoprotein sequence was obtained from Genbank (MT726045). Recombinant PRD-0038 S glycoprotein ectodomain (residues 16-1194) constructs with pre-fusion stabilizing mutations (PentaPro: F800P, S882P, S925P, K969P, V970P or HexaPro: F800P, A875P, S882P, S925P, K969P, V970P) were synthesized by GenScript with an N-terminal mu-phosphatase signal peptide and C-terminal short linker (SG), TEV protease site (RENLYFQ), a short linker (GGGGSG), Foldon, 8x His tag, a short linker (GGSS) followed by an Avi tag in a pCMVR plasmid. The SARS-CoV-2 S glycoprotein ectodomain construct comprises residues 1-1208 with the native signal peptide, the HexaPro prefusion stabilizing mutations[55] (F817P, A892P, A899P, A942P, K986P, V987P), abrogation of the furin cleavage site (residues 682-685, GSAS) and a C-terminal short linker (GSG), followed by a foldon, HRV 3C site (LEVLFQGP), a short linker (GSG), an avi tag, a short linker (GSG), an 8x his tag in a pcDNA3.1(-) plasmid.

Expi293F cells were grown at 37°C with 8% CO_2 and DNA transfections were conducted with the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatants were harvested four days post-transfection and proteins were purified using HisTrap™ High Performance column (Cytiva). Proteins were first washed with 10-15 column volumes of a buffer containing 25 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, pH 8.0, followed by elution with 10-15 column volumes using 300 mM imidazole, pH 8.0. Eluted proteins were concentrated and buffer exchanged into 1x PBS or 1x TBS (20 mM Tris, 150 mM NaCl, pH 8.0) using Amicon Ultra-15 Centrifugal Filter Unit (100 kDa) (Millipore). Purified proteins were snap frozen and stored at -80°C. Purified proteins were checked for endotoxin level using Charles River Limulus Amebocyte Lysate (LAL) cartridges (Charles River, PTS201F). Endotoxin-free SARS-CoV-2 HexaPro S and PRD-0038 HexaPro S were flash-frozen and stored at -80°C until the day of immunization.

Production of PRD-0038 wild-type RBD- Natively Dimerized R. alcyone Complex

For complex formation, wild-type PRD-0038 RBD was mixed with natively dimerized R. alcyone ACE2 His at a 4:1 molar ratio, then incubated at room temperature for 5-10 min. Gel filtration was performed to remove excess RBD on a Superose 6 10/300 GL column (Cytiva) equilibrated in 50 mM Tris-HCl, 150 mM NaCl. Complex formation was confirmed by SDS-PAGE, and the PRD-0038 RBD- R. alcyone ACE2 complex was snap frozen and stored at -80°C until day of grid preparation.

Binding analysis using biolayer interferometry (BLI)

BLI binding assays were performed on an Octet Red (Sartorius) instrument operated at 30°C with shaking (1000 rpm). For biotinylated RBD and ACE2-Fc or ACE2-His binding assays, streptavidin biosensors were hydrated in water for 10 min prior to the

experiment. Biosensors were incubated in 10x kinetics buffer (Sartorius) for 60s followed by the loading of biotinylated RBDs to the tip, all to a final level of 1 nm. Loaded biosensors were equilibrated in 10x kinetics buffer for 120s which served as our baseline. For avidity binding assays, association with 1 μM ACE2-Fc (dimeric form) was performed for 300 s followed by 300 s of dissociation in 10x kinetics buffer. For affinity binding assays to determine K_D values, RBD-loaded tips were dipped into a concentration series of ACE2-His (2-fold serial dilution from 100 nM to 6.25 nM for *R. alcyone*ACE2-His, 3 fold serial dilution from 660 nM to 8 nM for *R. landeri* ACE2-His) for 600 s followed by 600 s of dissociation in 10x kinetics buffer. Global fits were used to calculate K_D values using a 1:1 binding fit model. Data were plotted using GraphPad Prism. Assays were replicated with three biological replicates (recombinant RBD proteins generated on different days) and representative graphs are shown.

Production of VSV pseudoviruses

Wildtype and mutant PRD-0038 S constructs consisting of residues 1-1235 and containing a 21 residue C-terminal deletion (del21) followed by a 3x FLAG tag were synthesized by GenScript and placed into an HDM plasmid. VSV pseudoviruses were produced using HEK293T cells seeded on BioCoat Cell Culture Dish : poly-D-Lysine 100 mm (Corning). Cells were transfected with PRD-0038 S-Flag constructs using Lipofectamine 2000 (Thermo Fisher Scientific) in Opti-MEM transfection medium. After 5h of incubation at 37 °C with 5% CO2, cells were supplemented with DMEM containing 10% of FBS. On the next day, cells were infected with VSV (G*ΔG-luciferase) (Crawford, 2020; Kaname, 2010) for 2h, followed by five time wash with DMEM medium before addition of anti-VSV G antibody (I1-mouse hybridoma supernatant diluted 1:40, ATCC CRL-2700) and medium. After 18-24 h of incubation at 37 °C with 5% CO_2,

pseudoviruses were collected and cell debris removed by centrifugation at 3,000xg for 10 min. Pseudoviruses were further filtered using a 0.45 µm syringe filter and concentrated 25-50x prior to storage at -80°C. Mock VSV pseudoviruses were prepared as above but without S transfection.

Cell entry assays comparing wildtype and mutant PRD-0038 S VSV pseudoviruses

HEK293T cells were transfected with full length *Rhinolophus* ACE2 orthologs using Lipofectamine 2000 (Thermo Fisher Scientific) in Opti-MEM five hours prior to plating into 96-well plates [3610] (Corning) coated with poly-lysine [P4707] (Sigma) and incubated 18-24 h before infection with VSV pseudoviruses. For human ACE2 entry assays, HEK293T cells with stable hACE2 expression were plated into poly-lysine-coated 96-well plates and incubated for 18-24h before infection with VSV pseudoviruses. The amount of pseudovirus used for infection was adjusted using Western Blot based on S incorporation across different mutants to use a constant input of S. Detection of VSV backbone was performed with 1:1,000 Anti-VSV-M [23H12] Antibody (Kerafast). Detection of 3x-FLAG tagged S was performed with 1:400 monoclonal ANTI-FLAG® M2 antibody [F3165] produced in mouse (Sigma). 1:50,000 Alexa Fluor® 680 AffiniPure Goat Anti-Mouse IgG [115-625-174] (Jackson ImmunoResearch) was used as the secondary antibody. A representative Western Blot is shown in **Fig. S4.4E**. Genbank accession numbers for all ACE2s can be found in **Table S1**. After 1h of infection, an additional 40 µL of DMEM supplemented with 20% FBS and 2% PenStrep was added to the cells. After 18–20h, 40 µL of One-Glo-EX substrate (Promega) was added to each well and incubated on a plate shaker in the dark for 5 min before reading the relative luciferase units using a BioTek Neo2 plate

reader. Fold change of relative luciferase units over mock VSV were plotted in Prism (GraphPad) with mock being cells that were not transfected with an S-encoding plasmid. 3 biological replicates, each of which with 3 technical replicates were carried out.

Cell entry assay for PRD-0038 S VSV with distinct Rhinolophus ACE2s

For wildtype PRD-0038 S VSV entry into HEK293T cells transfected with *R. affinis*, *R. sinicus*, *R. landeri*, and *R. alcyone* alleles full-length ACE2 alleles (**Fig. 4.1D**), HEK293T cells (ATCC) were cultured in 10% FBS, 1% penicillin–streptomycin DMEM at 37 °C in a humidified 5% CO_2incubator. Cells were plated 18-24 hours prior to transfection into 96-well plates [3610] (Corning) coated with poly-lysine [P4707] (Sigma). All transfections were performed using full-length *Rhinolophus* ACE2 placed into a HDM plasmid (synthesized by GenScript). Transfection of ACE2 alleles into HEK293T cells was performed using 0.2 µg DNA and 0.15 µL Lipofectamine 2000 (Thermo Fisher Scientific) per well in Opti-MEM. After a 5 h incubation at 37 °C in a humidified 8% CO_2 incubator, DMEM was added to obtain a final concentration of 10% FBS and 1% penicillin–streptomycin. Cells were incubated at 37 °C in a humidified 8% CO_2 incubator for 36–48 h prior to infection. For each infection test, 2-3 technical replicates were performed, and the assays were repeated on a second day, for a total of 4-6 technical replicates. Three biological replicates (pseudovirus generated on different days) were used for cell entry and each point shown represents the mean fold change for each biological replicate. Results were plotted using Graphpad Prism (**Fig. 1D**). Genbank accession numbers for all ACE2s can be found in **Table S1.**

Cryo-EM sample preparation and data collection

Cryo-EM grids of PRD-0038 PentaPro S were prepared using two separate methods and data were combined during data processing. The first dataset was collected from the grids prepared using a reverse grid-blotting method. 3 µL of sample was added to the carbon side of a glow discharged C-flat R2/2 copper grid and 1µL was added to the back side before addition of 1 µL of CHEMS lipid dissolved in chloroform on the front side. The sample was allowed to sit on the grid for 1 minute and then manually blotted from the back side using a strip of Whatman #1 filter paper and plunged into liquid ethane. The second dataset was collected from a lacey carbon grid with a thin home-made continuous carbon layer. 3 µL of 0.15 mg/mL PRD-0038 Pentapro S was loaded onto the glow discharged (6s at 20mA) grid followed by plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific). The grid was blotted with a blot force of -1, 3 second blot time, and 10 second wait time before the plunge freeze at 100% humidity and 25 °C. For the PRD-0038 RBD-ACE2 complex, grids were prepared by applying 3 µL of 4 mg/ml PRD-0038 RBD bound to the *R. alcyone* ACE2 dimer with 7 mM CHAPSO (Anatrace) were applied and blotted twice as previously described (Snijder, 2017), onto freshly glow discharged R 2/2 UltrAuFoil grids prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 5 sec blot time at 100 % humidity and 22°C. Data were acquired using an FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K3 direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon (Suloway, 2005) at a nominal magnification of 105,000x with a pixel size of 0.843 Å. The dose rate was adjusted to 15 counts/pixel/s, and each movie was acquired in counting mode fractionated in 75

frames of 40 ms. A total of 2,482 and 11,743 micrographs were collected for the PRD-0038 S and PRD-0038 RBD-ACE2 datasets, respectively.

Cryo-EM data processing, model building and refinement

Motion correction, contrast-transfer function (CTF) parameter estimation, automatic particle picking, and extraction were performed using Warp (Tegunov, 2019) for each data set. For the PRD-0038 S structure, particle images were extracted with a box size of 260 pixels with a pixel size of 1.686Å. After two rounds of 2D classification using cryoSPARC (Punjani, 2017), well-defined particles were selected and particles from each dataset were combined and binned to a box size of 130 pixels with a pixel size of 3.372Å for subsequent 3D classification using Relion (Zivanov, 2018; Scheres, 2012) with 50 iterations (angular sampling 7.5˚ for 25 iterations and 1.8˚ with local search for 25 iterations). 103,347 particles were selected and re-extracted with a box size of 260 and pixel size of 1.686Å for cryoSPARC non-uniform refinement (Punjani, 2020) with C3 symmetry and further subjected to Bayesian polishing (Zivanov, 2019) in Relion. Finally, another round of non-uniform refinement with C3 symmetry and optimized per-particle defocus was carried out to the polished particles. To improve the density of the NTD, we used symmetry expansion and local refinement using cryoSPARC. For the PRD-0038 RBD - *R. alcyone*ACE2 structure, two rounds of reference-free 2D classification were performed using cryoSPARC to select well-defined particle images. These selected particles were subjected to two rounds of 3D classification with 50 iterations each (angular sampling 7.5˚ for 25 iterations and 1.8˚ with local search for 25 iterations) using Relion with an initial model generated with ab-initio reconstruction in cryoSPARC. 3D refinements were carried out using non-uniform refinement along with per-particle defocus refinement in CryoSPARC. To improve the

density of the RBD-ACE2 dimer, the particles were subjected to cryoSPARC heterogeneous refinement. Particles belonging to classes with the best resolved RBD-ACE2 density were selected and subjected to the Bayesian polishing procedure implemented in Relion before performing another round of non-uniform refinement in cryoSPARC followed by per-particle defocus refinement and again non-uniform refinement. To further improve the density of the RBD-ACE2 domains, the particles were symmetry expanded and subjected to focus 3D classification without refining angles and shifts using a soft mask encompassing the RBD and monomer ACE2 using a tau value of 40 in Relion. Particles belonging to classes with the best resolved RBD-ACE2 density were selected and then subjected to local refinement using CryoSPARC. Local resolution estimation, filtering, and sharpening were carried out using CryoSPARC. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (Rosenthal, 2003; Chen, 2013). UCSF Chimera (Pettersen, 2004), UCSF ChimeraX (Goddard, 2018), and Coot (Emsley, 2010) were used to fit atomic models into the cryoEM maps. S and RBD-ACE2 models were refined and relaxed using Rosetta(Wang, 2016; Frenz, 2019) using sharpened and unsharpened maps and validated using Phenix (Liebschner, 2019), Molprobity (Chen, 2010) and Privateer (Agirre, 2015). Analysis of interface residues was assisted by PISA (Krissinel, 2007).

Monoclonal antibody ELISAs

For PRD-0038 HexaPro S and SARS-CoV-2 HexaPro S ELISAs, 30 µl of the proteins at 3 µg/mL were plated onto 384-well Nunc Maxisorp plate (ThermoFisher, 464718) in 1x TBS and incubated 1h at 37°C followed by slap drying and blocking with 80 µL of

Casein for 1 h at 37°C. After incubation, plates were slap dried and 1:4 serial dilutions of the corresponding mAbs starting from 0.1 mg/ml were made in 30 µl TBST, added to the plate and incubated at 37°C for 1 h. Plates were washed 4x in TBST and 30 µl of 1:5,000 Goat anti-Human IgG Fc Secondary Antibody, HRP (Thermo Fisher, A18817) or Goat anti-Syrian Hamster IgG (H+L) Secondary Antibody, HRP (Thermo Fisher, PA1-28823) were added to each well and incubated at 37°C. After 1 h, plates were washed 4x in TBST and 30 µl of TMB (SeraCare) was added to every well for 2 min at room temperature. Reactions were quenched with the addition of 30 µl of 1N HCl. Plates were immediately read at 450 nm on a BioTek Neo2 plate reader and data plotted and fit in Prism 9 (GraphPad) using nonlinear regression sigmoidal, 4PL, X is the concentration to determine EC_{50}values from curve fits.

Deep mutational scanning for mutational effects on ACE2 binding

The complete deep mutational scanning pipeline can be found at: https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/summary/summary.md. Deep mutational scanning libraries for sarbecovirus strains including PRD-0038 and SARS-Cov-1 Urbani were constructed as previously described (Starr, 2022). Briefly, site-saturation mutagenesis libraries spanning all RBD positions were produced by Twist Bioscience (or NNS mutagenesis for positions that failed Twist mutagenesis), tagged with an N16 barcode, and cloned into a yeast display vector backbone via Gibson Assembly. Libraries were electroporated into *E. coli* and plated at a target bottleneck of 40,000 unique barcodes per library to overrepresent the ~4,000 possible amino acid mutations. Colonies were scraped from each transformation plate, library plasmid purified, and transformed into the AWY101 *S. cerevisiae* strain (Wentz, 2007) for yeast surface-display experiments.

Library plasmids were sequenced using a PacBio Sequel IIe to generate long sequence reads spanning the N16 barcodes and RBD coding sequence. Raw CCS reads are available on the NCBI Sequence Read Archive, BioProject PRJNA962117, BioSample SAMN34384156. Reads were processed using alignparse (version 0.2.4) (Crawford, 2019) to generate a table linking each N16 barcode to its unique RBD mutant, available at: https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/variants/codon_variant_table_PRD0038.csv and https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/variants/codon_variant_table_SARS1.csv.

The RBD expression level and ACE2-binding avidity of each RBD mutant was determined via high-throughput FACS-seq assays as previously described (Starr, 2022). ACE2-binding titrations were performed by incubating induced yeast-display libraries with a concentration series of dimeric ACE2 ligands from 10^{-6} to 10^{-13} M at 1-log intervals, plus a 0 M ACE2 sample, with samples equilibrated overnight at room temperature with mixing. Yeast were washed with PBS-BSA (0.2 mg/L), labeled with 1:100 FITC-conjugated chicken anti-Myc (Immunology Consultants CMYC-45F) to detect yeast-displayed RBD and 1:200 PE-conjugated streptavidin (Thermo Fisher S866) or goat anti-human-IgG (Jackson ImmunoResearch 109-115-098) to detect binding of biotinylated (human, Acro Biosystems H82E6) or Fc-tagged (*R. alcyone*, *R. landeri*) ACE2. For each titration sample, RBD^+ yeast were fractionated into four bins of PE fluorescence (ACE2 binding), grown overnight, plasmid isolated, N16 barcode amplified, and barcodes counted via high-throughput sequencing on an Illumina NextSeq. RBD expression was measured by sorting cells into four bins on the

basis of Myc-FITC labeling, followed by outgrowth, plasmid isolation, N16 barcode amplification, and sequencing. Sequencing reads are available on the NCBI Sequence Read Archive, BioProject PRJNA962117, BioSample SAMN34384823. Demultiplexed Illumina barcode reads were aligned to library barcodes using dms_variants (version 0.8.9), yielding a table of counts of each barcode in each FACS bin which is available at https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/counts/variant_counts.csv.gz. Reads were downweighted by the ratio of total sequence reads from a bin to the number of cells sorted into that bin. For each barcode, we inferred the apparent dissociation constant for avid binding ($K_{D,app}$) by fitting the standard non-cooperative Hill equation to the mean FACS bin of a barcode variant as a function of ACE2 concentration. For each barcode, expression was determined via a maximum likelihood estimator of log-MFI based on the distribution of barcode counts across FACS bins and the known fluorescence boundaries of those bins. The computational pipelines for computing per-barcode binding constants and expression phenotypes are available at: https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/summary/compute_binding_Kd_huACE2.mdand https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/summary/compute_expression_meanF.md. Because most mutants in the library were independently associated with more than one N16 barcode, we derived the final mutant phenotype as the average of per-barcode measurements, as computed at: https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/summary/collapse_barcodes_lib40_41.md. The final per-mutant deep mutational scanning phenotypes are available

at: https://github.com/tstarrlab/SARSr-CoV-RBD_DMS/blob/main/results/final_variant_scores/final_variant_scores_lib40_41.csv.

Immunogenicity

Prior to each immunization, immunogens (endotoxin-free SARS-CoV-2 HexaPro S or PRD-0038 HexaPro S) were diluted to 20 µg/mL in 1x TBS (20mM Tris, 150mM NaCl, pH 8.0) and mixed with 1:1 vol/vol AddaVax (InvivoGen vac-adx-10) to reach a final dose of 1 µg of immunogen per injection. At 8 weeks of age, 6 mice per group were anesthetized and injected intramuscularly in the quadriceps with 50µL of immunogen per leg, 100µL total at weeks 0, 3, and 6. Mice were bled via the submental route at weeks 0, 2, 5, and 8. Blood was collected in serum separator tubes (BD # 365967) and rested for 30 min at room temperature for coagulation. Serum tubes were then centrifuged for 10 min at 2,000 x g and serum was collected and stored at -80°C until use.

Neutralization assays

For mAb neutralization against SARS-CoV-2 S VSV and PRD-0038 S K482Y/T487W VSV, HEK293T cells with stable human ACE2 expression in DMEM supplemented with 10% FBS and 1% PenStrep were seeded at 40,000 cells/well into 96-well plates [3610] (Corning) coated with poly-lysine [P4707] (Sigma) and incubated overnight at 37°C. After 16-20h of incubation, a half-area 96-well plate (Greiner) was prepared with 1:5 serial dilutions of S2X259 and S2X35 starting from 0.1 mg/ml in DMEM, and 1:3 serial dilutions of 76E1 starting from 0.5 mg/ml in DMEM, for a total of 22 µL per well. An equal volume of DMEM with diluted pseudoviruses was added to each well. All pseudoviruses were diluted between 1:3-1:27 to reach a target entry of ~10^6 RLU. The

mixture was incubated at room temperature for 45-60 minutes. Media was removed from the cells and 40 µL from each well of the half-area 96-well plate containing mAb and pseudovirus were transferred to the 96-well plate seeded with cells and incubated at 37°C for 1h. After 1h, an additional 40 µL of DMEM supplemented with 20% FBS and 2% PenStrep was added to the cells. After 18–20h, 40 µL of One-Glo-EX substrate (Promega) was added to each well and incubated on a plate shaker in the dark for 5 min before reading the relative luciferase units using a BioTek Neo2 plate reader. Relative luciferase units were plotted and normalized in Prism (GraphPad): 100% neutralization being cells lacking pseudovirus and 0% neutralizing being cells containing virus but lacking mAb. Prism (GraphPad) nonlinear regression with "[inhibitor] versus normalized response with a variable slope" was used to determine IC_{50} values from curve fits with 2 technical repeats. 3 biological replicates were carried out for each mAb. For SARS-CoV-2 D614G S VSV (Crawford, 2020), BA.2 S VSV, BA.5 S VSV (Bowen, 2022), RaTG13 S VSV (Walls, 2021), and SARS-CoV-1 S VSV (Millet, 2016) neutralization, HEK293T cells with stable human ACE2 expression in DMEM supplemented with 10% FBS and 1% PenStrep were seeded at 40,000 cells/well into 96-well plates [3610] (Corning) coated with poly-lysine [P4707] (Sigma) and incubated overnight at 37°C. For PRD-0038 S VSV and Khosta-1 S VSV neutralization, HEK293T cells were transfected with full length *R. alcyone* ACE2 using Lipofectamine 2000 (Thermo Fisher Scientific) in Opti-MEM five hours prior to plating into 96-well plates [3610] (Corning) coated with poly-lysine [P4707] (Sigma) and incubated overnight at 37°C. The following day, a half-area 96-well plate (Greiner) was prepared with 3-fold serial sera dilutions (starting dilutions determined for each serum and pseudovirus,

22uL per well). An equal volume of DMEM with diluted pseudoviruses was added to each well. All pseudoviruses were diluted between 1:3-1:27 to reach a target entry of ~10^6 RLU. The mixture was incubated at room temperature for 45-60 minutes. Media was removed from the cells and 40 µL from each well of the half-area 96-well plate containing sera and pseudovirus were transferred to the 96-well plate seeded with cells and incubated at 37°C for 1h. After 1h, an additional 40 µL of DMEM supplemented with 20% FBS and 2% PenStrep was added to the cells. After 18–20h, 40 µL of One-Glo-EX substrate (Promega) was added to each well and incubated on a plate shaker in the dark for 5 min before reading the relative luciferase units using a BioTek Neo2 plate reader. Relative luciferase units were plotted and normalized in Prism (GraphPad): 100% neutralization being cells lacking pseudovirus and 0% neutralizing being cells containing virus but lacking sera. Prism (GraphPad) nonlinear regression with "log[inhibitor] versus normalized response with a variable slope" was used to determine ID_{50} values from curve fits with 3 technical repeats. 3 biological replicates were carried out for each sample-pseudovirus pair.

Breadth- and epitope-mapping of vaccine sera via deep mutational scanning

The complete serum deep mutational scanning pipeline is described https://github.com/tstarrlab/SARSr-CoV_MAP_PRD0038-vaccine/blob/main/results/summary/summary.md. Binding of serum was evaluated against deep mutational scanning pools for PRD-0038 and SARS-CoV-1 whose construction is described above, previously published deep mutational scanning pools for SARS-CoV-2 Wuhan-Hu-1[79] and Omicron BA.2 (Starr 2022), and a previously published pan-sarbecovirus panel (Starr, 2022) that was supplemented with additional newly described sarbecovirus and SARS-CoV-2 variants. Serum was first depleted of

non-specific yeast-reactive antibodies as previously described (Greaney, 2021). Yeast-display RBD libraries were pooled, induced for yeast surface expression, and labeled with serum at 1:100 1:1000, 1:10,000, and 1:100,000 dilutions for one hour at room temperature. Yeast were washed with PBS-BSA and labeled with secondary Myc-FITC antibody and APC-conjugated goat anti-mouse-IgG (Jackson ImmunoResearch 115-605-008). As with ACE2-binding titrations, libraries were then partitioned into four bins of serum binding on a BD FACSAria, collecting a minimum of 6 million RBD+ cells per sample concentration across the four bins. Cells were grown post-sort, plasmid purified, N16 barcode amplified, and sequenced on an Illumina NextSeq. Raw Illumina sequencing data is available from the NCBI Sequence Read Archive, BioProject PRJNA714677, BioSample SAMN36715819. Barcode reads were mapped to library barcodes, with raw counts found at: https://github.com/tstarrlab/SARSr-CoV_MAP_PRD0038-vaccine/blob/main/results/counts/variant_counts.csv.

For each library barcode, an area under the curve (AUC) metric was derived from its distribution of sequence reads across sort bins. First, the strength of serum binding to each barcode at each serum dilution was determined as the simple mean bin from cell counts across integer-weighted bins, and subtracted by background mean bin determined from a sort from yeast libraries not incubated with sera. Any barcode with less than 3 cell counts at any sample concentration was eliminated from analysis. An AUC metric was then calculated from the relationship between mean bin and serum dilution. AUC calculation can be found at: https://github.com/tstarrlab/SARSr-CoV_MAP_PRD0038-vaccine/blob/main/results/summary/compute_AUC.md, and per-barcode AUC metrics are available at: https://github.com/tstarrlab/SARSr-

CoV_MAP_PRD0038-vaccine/blob/main/results/bc_sera_binding/bc_sera_binding.csv. We then computed the per-variant AUC as the robust mean of replicate barcodes linked with the identical RBD variant, by taking the mean per-barcode AUC after trimming tails of the top and bottom 2.5% of AUC values among the replicate barcodes. Because mutations that disrupt RBD expression artifactually decrease serum binding, we applied two final filters: first, we censored the AUC measurement for any mutant with a measured impact on RBD expression of greater than one log-MFI unit (RBD expression < -1) from DMS measurements described above; and second, we derived a normalization constant from the slope of the linear model relating serum AUC and expression globally across all library variants for variants with expression >-1, and normalized our raw AUC measurements by this constant. The final variant derivation can be found at: https://github.com/tstarrlab/SARSr-CoV_MAP_PRD0038-vaccine/blob/main/results/summary/collapse_barcodes_SARSr-DMS.md, and final per-variant serum-binding values are available at: https://github.com/tstarrlab/SARSr-CoV_MAP_PRD0038-vaccine/blob/main/results/final_variant_scores/final_variant_scores_wts.csv and https://github.com/tstarrlab/SARSr-CoV_MAP_PRD0038-vaccine/blob/main/results/final_variant_scores/final_variant_scores_dms.csv.

Quantification and statistical analysis

Quantification details of experiments can be found in the figure legends.

4.10 Figures & Tables

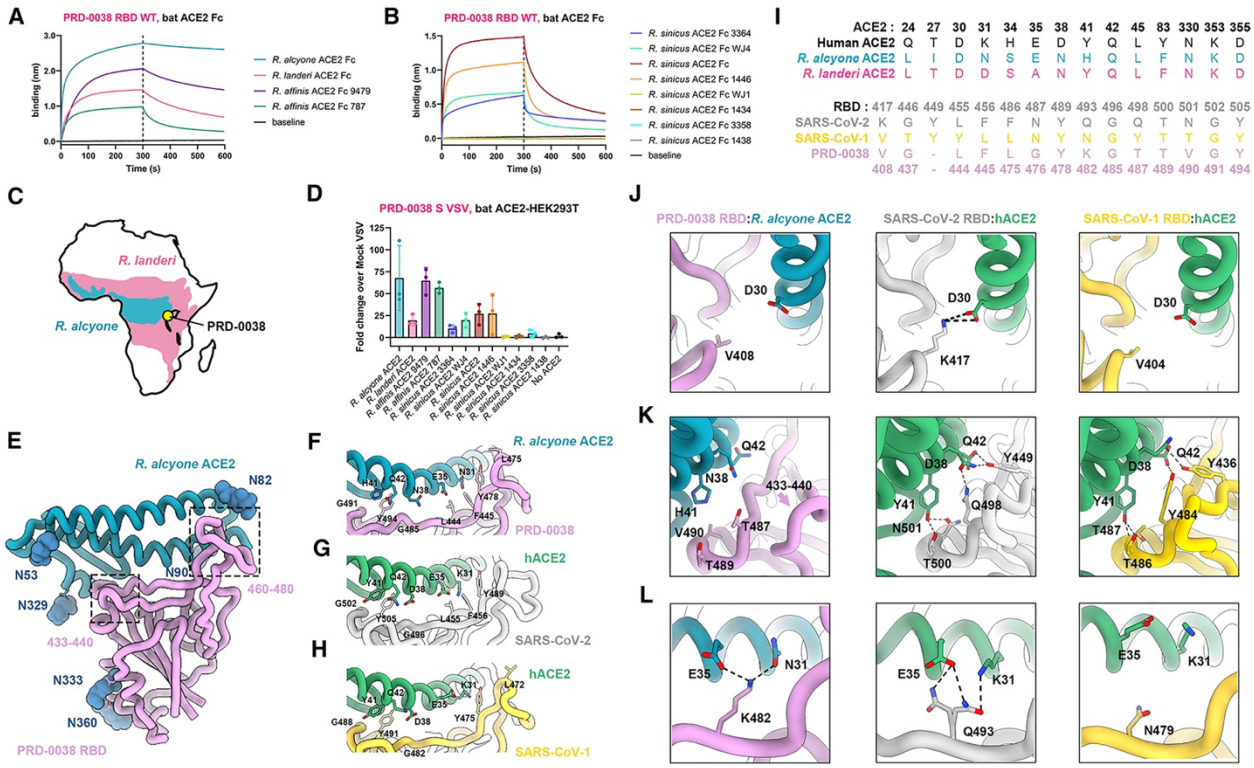


Figure 4.1: The clade 3 PRD-0038 sarbecovirus has a broad Rhinolophus bat ACE2 tropism

(A and B) BLI binding analysis of *R. affinis*, *R. alcyone*, and *R. landeri* (see also **Table S4.1**) (A) or *R. sinicus* (B) ACE2-Fc alleles at a concentration of 1 μM to the biotinylated PRD-0038 RBD immobilized on streptavidin biosensors. Baselines represent non-specific binding of uncoated streptavidin biosensors to 0.25 μM ACE2 Fc.
(C) Known geographic distribution of *R. alcyone* and *R. landeri* bats in sub-Saharan Africa (https://www.iucnredlist.org/). The yellow area indicates the site of PRD-0038 sampling (Rwanda).
(D) Entry of VSV pseudotyped with wild-type PRD-0038 S in HEK293T cells transiently transfected with the indicated *Rhinolophus* bat ACE2 orthologs. Each point represents the average of technical duplicates from each biological triplicate. Means and standard deviations shown as bars and error bars.
(E) Cryo-EM structure of the PRD-0038 RBD bound to *R. alcyone* ACE2. N-linked glycans are shown as dark blue spheres and labeled with the corresponding asparagine residue number. The dotted gray boxes highlight regions with major structural deviations from SARS-CoV-2 (see also **Fig. S4.2**).
(F–H) Close-up views of the interface between the PRD-0038 RBD and *R. alcyone*ACE2 (F, pink and blue, respectively), the SARS-CoV-2 RBD and human ACE2 (G, gray and green, respectively, PDB: 6M0J[33]), and the SARS-CoV-1 RBD and human ACE2 (H, gold and green, respectively, PDB: 2AJF[34]). Key conserved residues at the interface are rendered as sticks.
(I) Sequence alignments of the key ACE2 and RBD residues at the binding interfaces. Numberings used are for human ACE2 (black), SARS-CoV-2 (gray), and PRD-0038 (pink) (see also Table S1).

(J–L) Close-up views of selected key contact residues at the interface between the PRD-0038 RBD and *R. alcyone* ACE2, the SARS-CoV-2 RBD and human ACE2, and the SARS-CoV-1 RBD and human ACE2 colored as in (F). (J) shows that the salt bridge formed between $K417_{SARS\text{-}CoV\text{-}2}$ and $D30_{hACE2}$ is absent due to substitution to $V408_{PRD\text{-}0038}$ and $V404_{SARS\text{-}CoV\text{-}1}$. (K) shows that the electrostatic interactions involving residues $D38_{hACE2}$ and $Q42_{hACE2}$ with $Y449_{SARS\text{-}CoV\text{-}2}$ and $Q498_{SARS\text{-}CoV\text{-}2}$ or $Y436_{SARS\text{-}CoV\text{-}1}$ and $Y484_{SARS\text{-}CoV\text{-}1}$ are lost due to truncation and remodeling of the PRD-0038 433–440 loop along with substitution of $Q498_{SARS\text{-}CoV\text{-}2}$ to $T487_{PRD\text{-}0038}$(L) shows that $Q493_{SARS\text{-}CoV\text{-}2}$ optimally interacts with $K31_{hACE2}$/$E35_{hACE2}$, whereas the topologically equivalent residue $K482_{PRD\text{-}0038}$ is better adapted to $N31_{RaACE2}$/$E35_{RaACE2}$due to swapping of the position of a positively charged amino acid side chain across the interface.

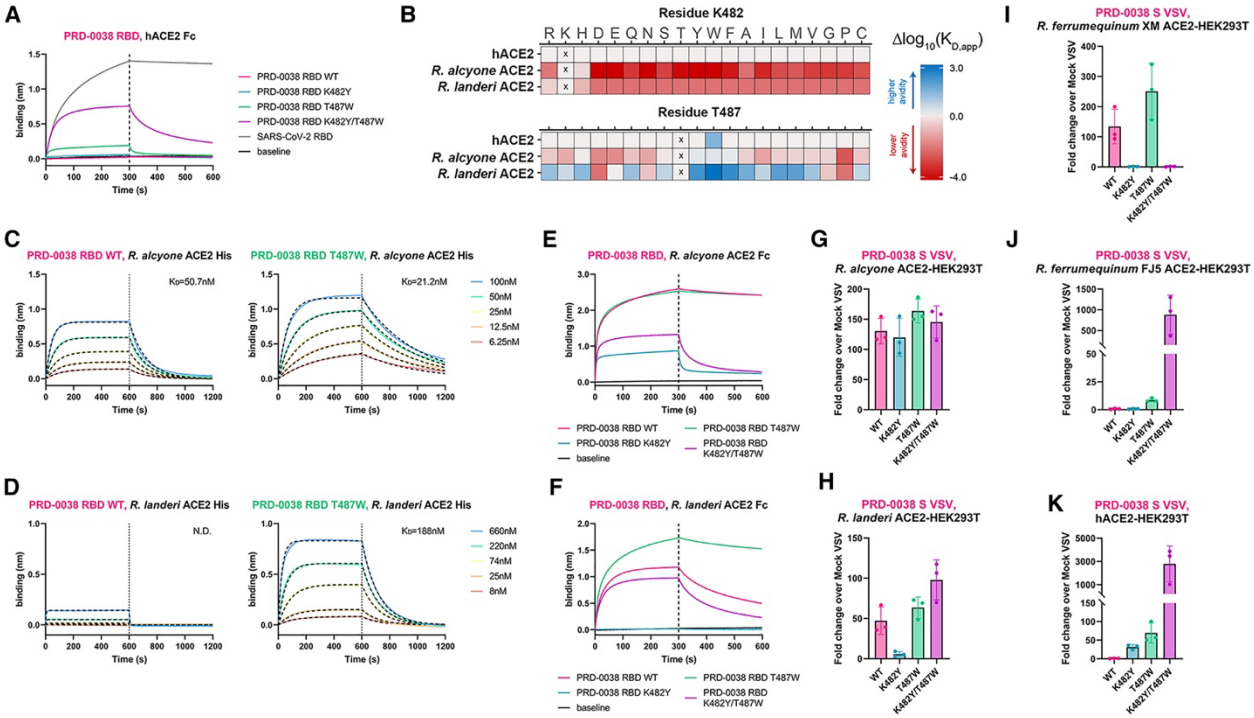


Figure 4.2: PRD-0038 RBD amino acid mutations broaden receptor tropism.

(A) BLI binding analysis of 1 μM dimeric hACE2-Fc to biotinylated wild-type and mutant PRD-0038 RBDs immobilized on streptavidin biosensors.
(B) DMS heatmaps of change in binding avidity to several ACE2-Fc orthologs caused by all possible mutations of the PRD-0038 RBD residues K482 and T487. An interactive version of the DMS data can be found at https://tstarrlab.github.io/SARSr-CoV-RBD_DMS/RBD-heatmaps_delta/.
(C) BLI binding analysis of various concentrations of monomeric *R. alcyone* ACE2 to biotinylated wild-type (left) and T487W (right) PRD-0038 RBDs immobilized on streptavidin biosensors.
(D) BLI binding analysis of various concentrations of monomeric *R. landeri* ACE2 to biotinylated wild-type (left) and T487W (right) PRD-0038 RBDs immobilized on streptavidin biosensors.

(E) BLI binding analysis of 1 μM dimeric *R. alcyone* ACE2-Fc to biotinylated wild-type and mutant PRD-0038 RBDs immobilized on streptavidin biosensors.
(F) BLI binding analysis of 1 μM dimeric *R. landeri* ACE2-Fc to biotinylated wild-type and mutant PRD-0038 RBDs immobilized on streptavidin biosensors.
(G–K) Entry of VSV pseudotyped with wild-type and mutants PRD-0038 S into HEK293T cells transiently transfected with *R. alcyone* ACE2 (G), *R. landeri* ACE2 (H), *R. ferrumequinum* XM_033107295.1 (I, XM), *R. ferrumequinum* FJ598617.1 (J, FJ5), or stably expressing human ACE2 (K). Each point represents the average of technical triplicates from each biological triplicate. Means and standard deviations shown as bars and error bars.
See STAR Methods and **Fig. S4.4E** for VSV S pseudotype normalization details.

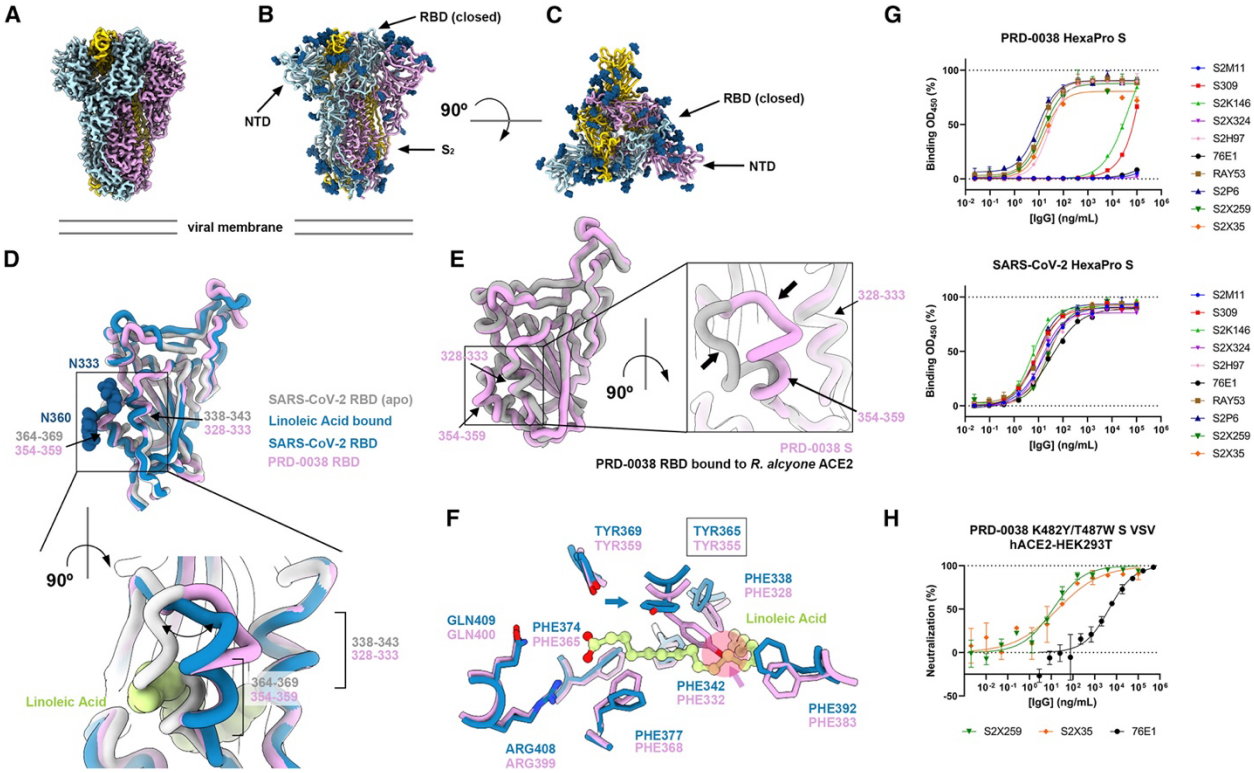


Figure 4.3: Architecture and antigenicity of the PRD-0038 S trimer.

(A) Unsharpened cryo-EM map of the closed PRD-0038 S trimer at 2.8 Å resolution.
(B and C) Ribbon diagram of the PRD-0038 S trimer atomic model viewed along (B, side) and normal (C, top) to the viral membrane. N-linked glycans are rendered as blue spheres. In panels (A–C), S protomers are colored blue, pink, and gold.
(D) Superimposition of the PRD-0038 S structure described here to the apo (gray, PDB: 6VXX[5]) and linoleic acid-bound (blue, PDB: 6ZB5[44]) SARS-CoV-2 S structures with a close-up view of the linoleic acid-binding pocket. N-linked glycans at positions N333 and N360 are rendered as blue spheres.
(E) Superimposition of the PRD-0038 RBDs from the apo S structure (pink) and from the *R. alcyone* ACE2-bound RBD structure (gray).

(F) Conservation of the SARS-CoV-2 linoleic acid-binding residues in PRD-0038 S. The PRD-0038 S Y355 side chain rotamer (Y365 in SARS-CoV-2 numbering) would sterically hinder linoleic acid binding (semi-transparent red circle denoted with an arrow).
(G) Evaluation of binding of a panel of monoclonal antibodies to PRD-0038 S Hexapro S and SARS-CoV-2 Hexapro S measured by ELISA. Each point represents the mean of technical triplicates. Standard deviations shown as error bars.
(H) Monoclonal antibody-mediated neutralization of PRD-0038 K482Y/T487W S pseudotyped VSV entry into HEK293T cells stably expressing human ACE2. Each point represents the mean of technical duplicates. Standard deviations shown as error bars.

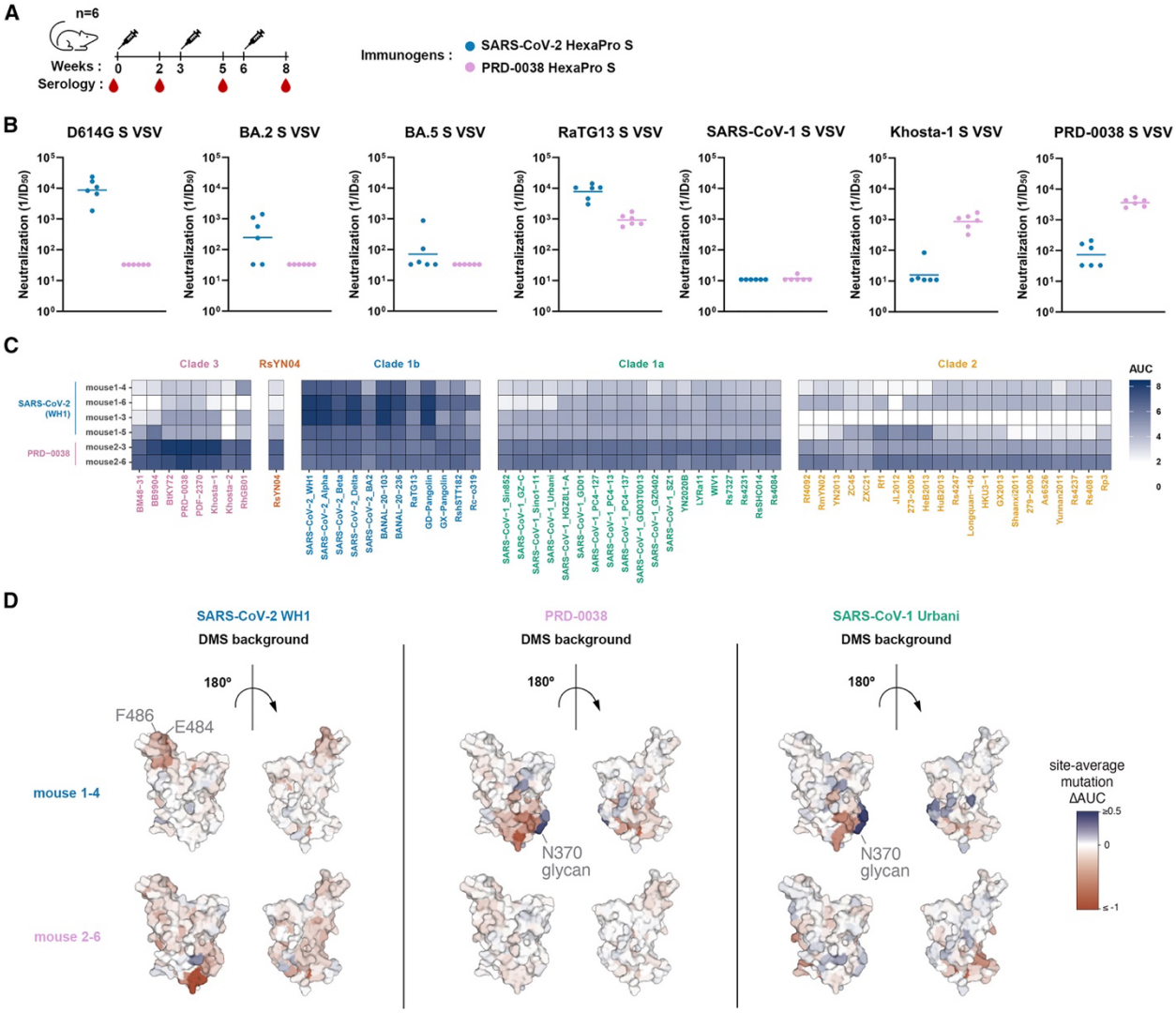
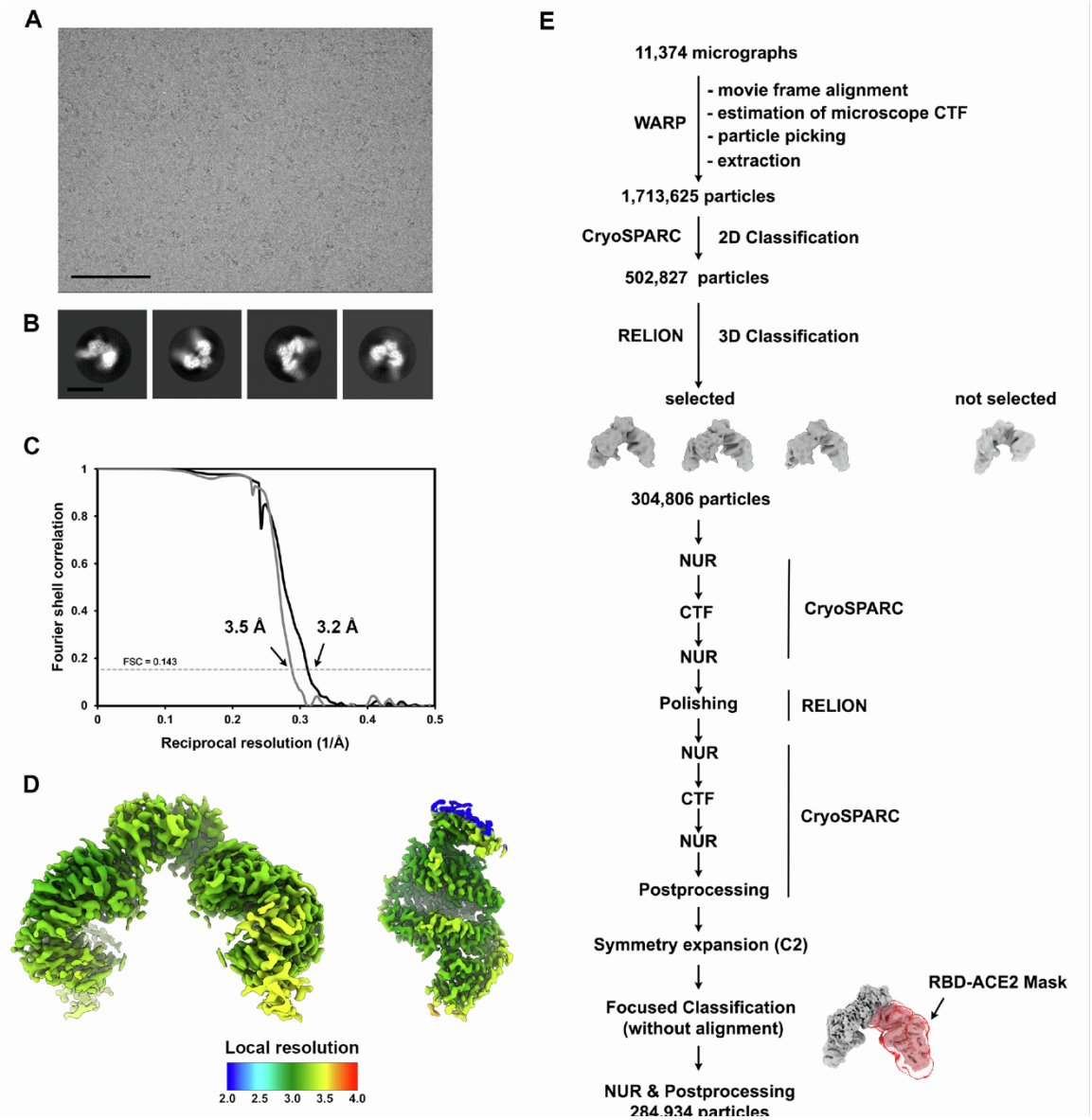


Figure 4.4: A clade 3 sarbecovirus S trimer elicits broadly reactive antibody responses.

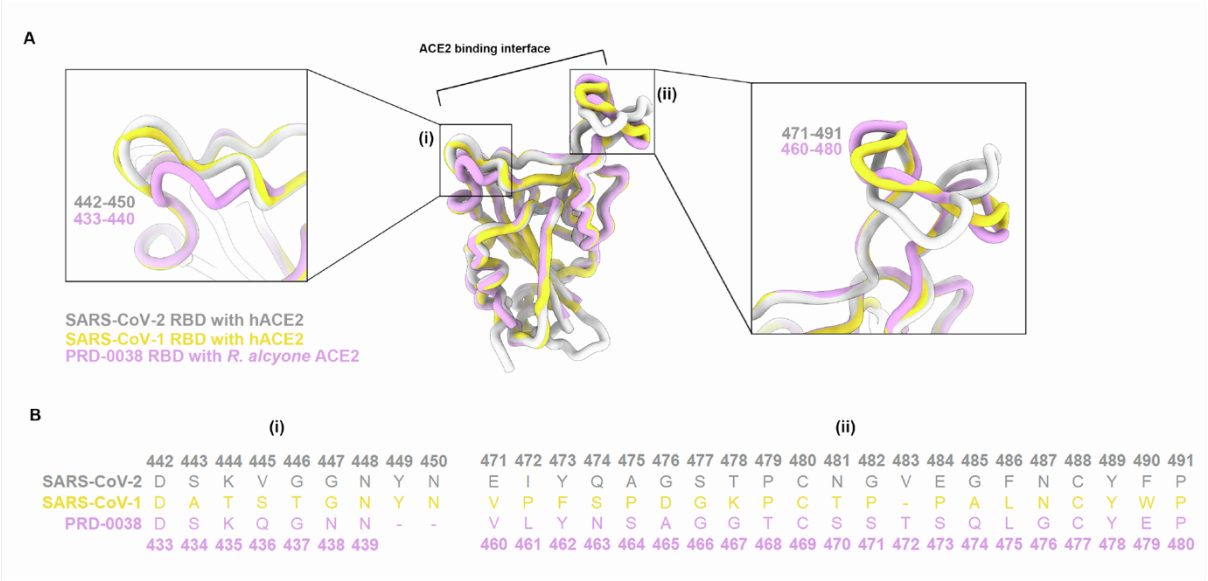
(A) Unsharpened cryo-EM map of the closed PRD-0038 S trimer at 2.8 Å resolution. (B and C) Ribbon diagram of the PRD-0038 S trimer atomic model viewed along (B, side) and normal (C, top) to the viral membrane. N-linked glycans are rendered as blue spheres. In panels (A–C), S protomers are colored blue, pink, and gold.

(D) Superimposition of the PRD-0038 S structure described here to the apo (gray, PDB: 6VXX[5]) and linoleic acid-bound (blue, PDB: 6ZB5[44]) SARS-CoV-2 S structures with a close-up view of the linoleic acid-binding pocket. N-linked glycans at positions N333 and N360 are rendered as blue spheres.
(E) Superimposition of the PRD-0038 RBDs from the apo S structure (pink) and from the *R. alcyone* ACE2-bound RBD structure (gray).
(F) Conservation of the SARS-CoV-2 linoleic acid-binding residues in PRD-0038 S. The PRD-0038 S Y355 side chain rotamer (Y365 in SARS-CoV-2 numbering) would sterically hinder linoleic acid binding (semi-transparent red circle denoted with an arrow).
(G) Evaluation of binding of a panel of monoclonal antibodies to PRD-0038 S Hexapro S and SARS-CoV-2 Hexapro S measured by ELISA. Each point represents the mean of technical triplicates. Standard deviations shown as error bars.
(H) Monoclonal antibody-mediated neutralization of PRD-0038 K482Y/T487W S pseudotyped VSV entry into HEK293T cells stably expressing human ACE2. Each point represents the mean of technical duplicates. Standard deviations shown as error bars.



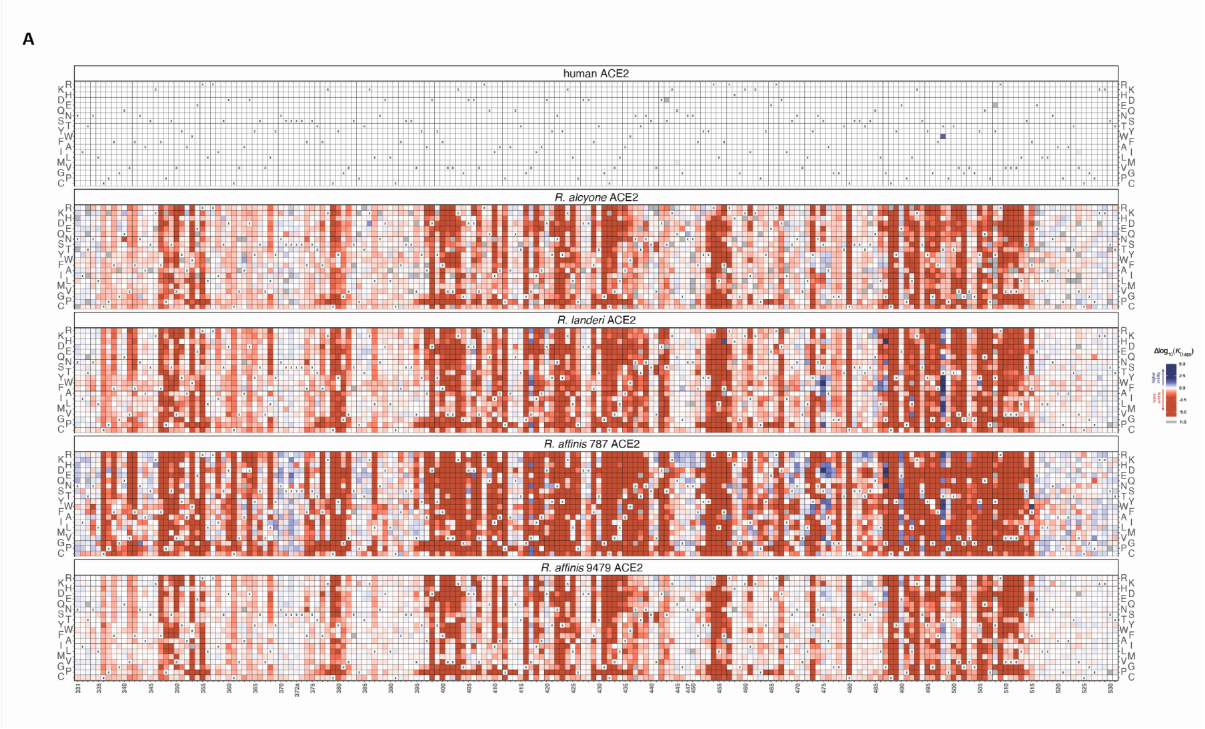
Supplemental Figure 4.1: CryoEM data collection and refinement of the dimeric R. alcyone ACE2-bound PRD-0038 RBD complex (related to **Fig. 4.1**).

(A and B) Representative electron micrograph (A) and 2D class averages (B) of -the dimeric *R. alcyone* ACE2-bound PRD-0038 RBD embedded in vitreous ice. The scale bar represents 100 nm (A) or 100 Å (B). (C) Gold-standard Fourier shell correlation curves for the final cryoEM reconstructions of the dimeric - ACE2/RBD complex (solid gray line) and locally refined ACE2/RBD (solid black line) shown in (D). The 0.143 cutoff is indicated with a gray dashed line.
(D) Local resolution map calculated using CryoSPARC and plotted onto the sharpened cryoEM map. (E) Data processing flowchart. CTF: contrast transfer function; NUR: non-uniform refinement.



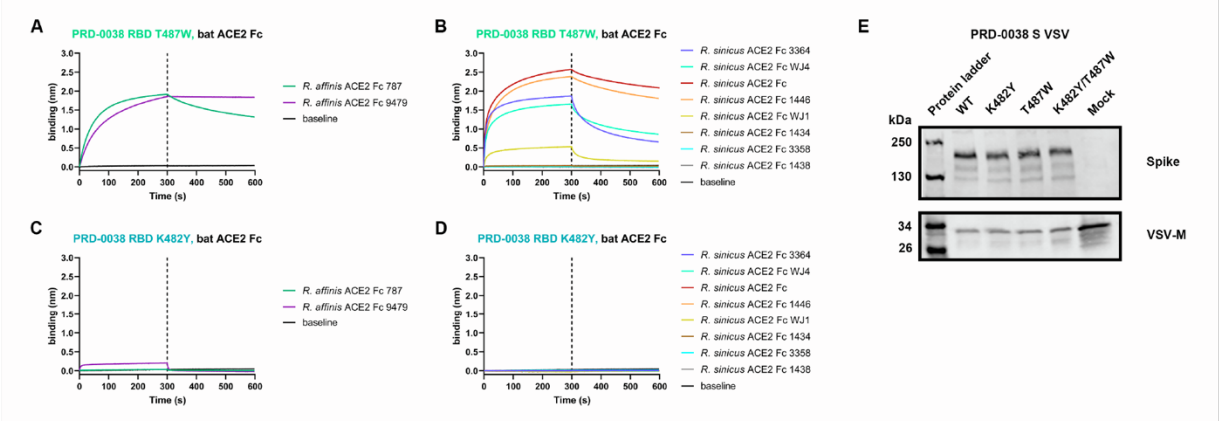
Supplemental Figure 4.2: Structural distinctions between the PRD-0038, SARS-CoV-2 and SARS-CoV-1 RBDs near the ACE2-binding interface (RBM) (related to **Fig. 4.1**).

(A) Ribbon diagrams showing a superimposition of the *R. alcyone* ACE2-bound PRD-0038 RBD (pink) superimposed to the human ACE2-bound SARS-CoV-2 RBD (gray, PDB 6M0J33) and SARS-CoV-1 RBD (gold, PDB 2AJF34) structures (ACE2s not shown for clarity). Insets: close-up views of two RBM regions. (B) Amino acid sequence alignment of the SARS-CoV-2, SARS-CoV-1, and PRD-0038 RBD regions highlighted in the insets shown in panel (A). - indicate deletions.



Supplemental Figure 4.3: Heatmaps of change in ACE2-binding avidity resulting from RBD mutations determined by DMS (related to **Fig. 4.2**).

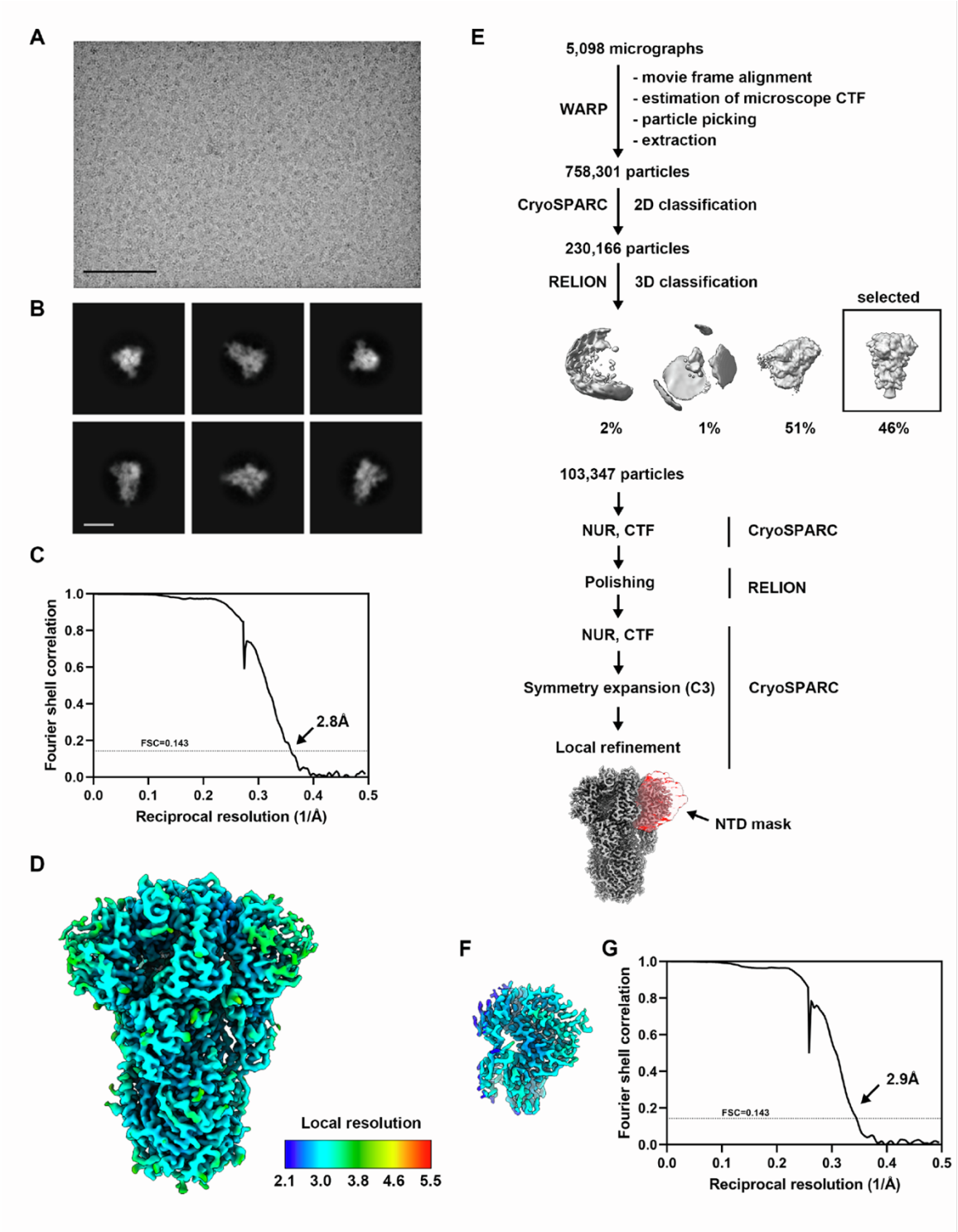
(A) Change in ACE2-binding avidity for hACE2, R. alcyone ACE2, and R. landeri ACE2 (delta- log10KD,apparent). Residue numbering corresponds to SARS-CoV-2.



*Supplemental Figure 4.4: Binding data and Western blot analysis (related to **Fig. 4.2**).*

(A and B) BLI binding analysis of 1 μM dimeric *R. affinis* (A) and *R. sinicus* (B) ACE2-Fc alleles to the biotinylated T487W PRD-0038 RBD immobilized on streptavidin biosensors.
(C and D) BLI binding analysis of 1 μM dimeric *R. affinis* (C) and *R. sinicus* (D) ACE2-

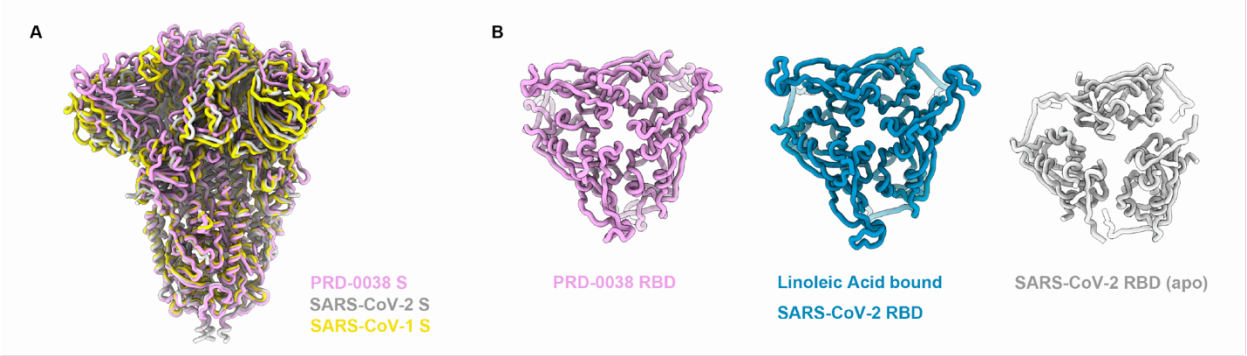
Fc alleles to the biotinylated K482Y PRD-0038 RBD immobilized on streptavidin biosensors.
(E) Representative Western Blot of wildtype (WT) and mutant PRD-0038 S VSV pseudoviruses normalized based on the amount of incorporated S and VSV-M. Anti-VSV-M Antibody (Kerafast) and Monoclonal ANTI- FLAG® M2 antibody produced in mouse (Sigma) were used as the primary antibody against VSV backbone and S, respectively. Alexa Fluor® 680 AffiniPure Goat Anti-Mouse IgG (Jackson ImmunoResearch) was used as the secondary antibody. iBrightTM Prestained Protein Ladder (Invitrogen, LC5615) was used as protein molecular weight markers.



Supplemental Figure 4.5: CryoEM data collection and refinement of PRD-0038 S (related to **Fig. 4.3**).

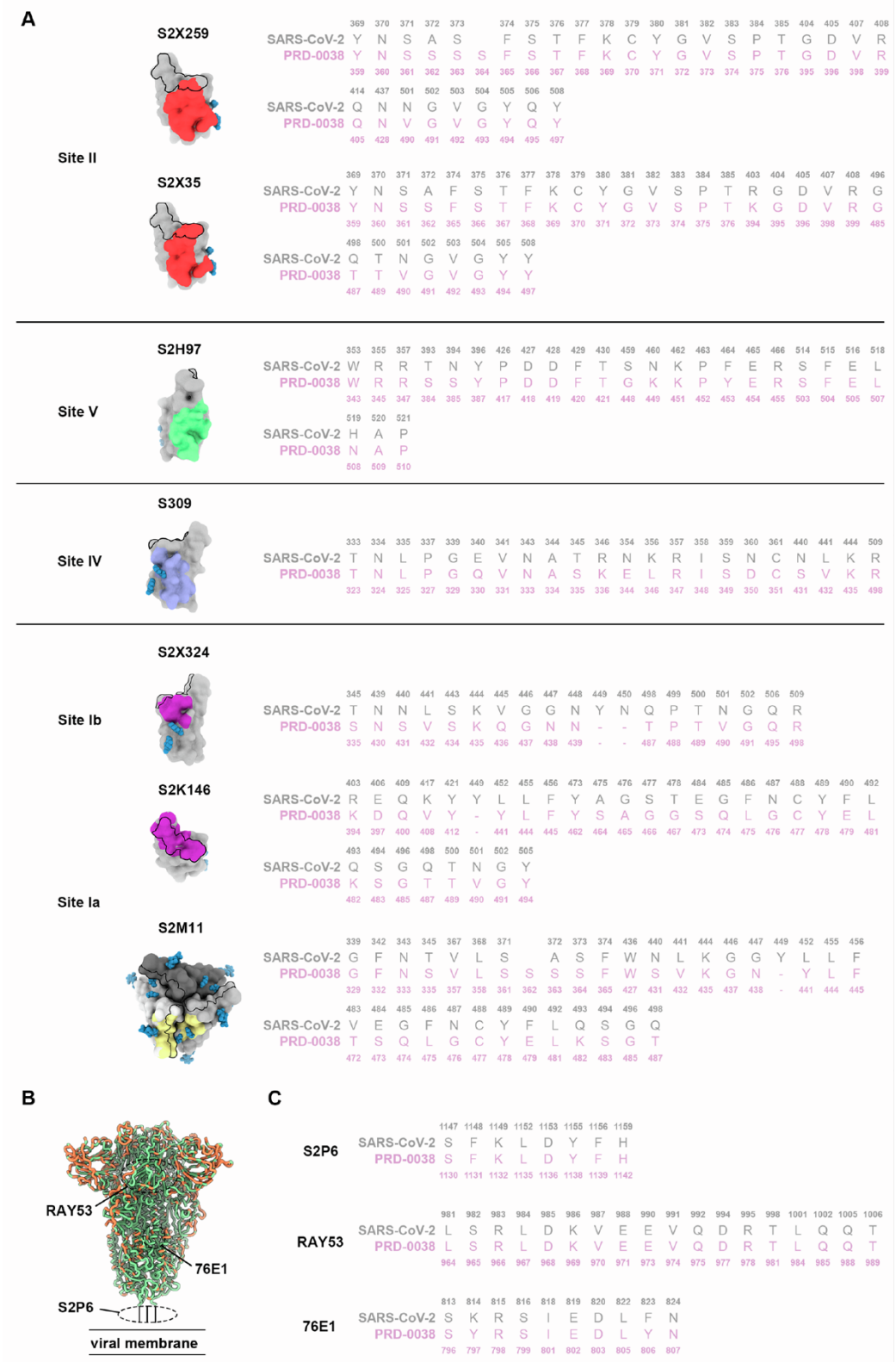
(A and B) Representative electron micrograph (A) and 2D class averages (B) of PRD-0038 PentaPro S embedded in vitreous ice. The scale bar represents 100 nm (A) or 100Å (B).

(C) Gold-standard Fourier shell correlation curve for the cryoEM reconstruction. The 0.143 cutoff is indicated with a gray dashed line.
(D) 3D reconstruction of PRD-0038 PentaPro S colored by local resolution as determined using cryoSPARC.
(E) Data processing flowchart. CTF: contrast transfer function; NUR: non-uniform refinement.
(F) 3D reconstruction obtained by local refinement of the PRD-0038 S NTD colored by local resolution as determined using cryoSPARC.
(G) Gold-standard Fourier shell correlation curve. The 0.143 cutoff is indicated with a gray dashed line.



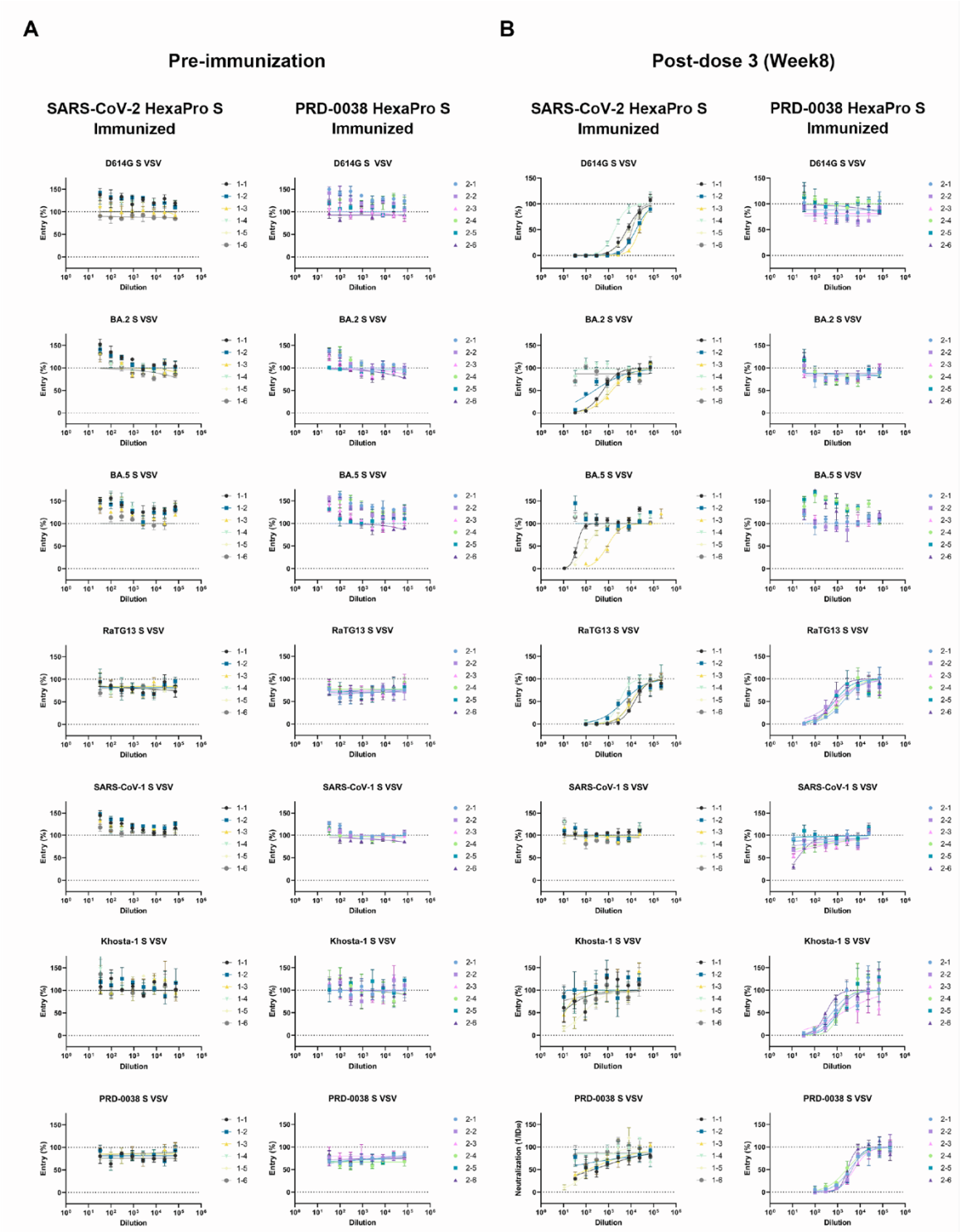
Supplemental Figure 4.6: Structural resemblance of PRD-0038 S to SARS-CoV-2 and SARS-CoV-1 (related to **Fig. 4.3**).

(A) Ribbon diagram of the PRD-0038 S trimer superimposed to SARS-CoV-2 S (PDB 6VXX5) and SARS- CoV-1 S (PDB 5X5B43) underscoring the similarity of the three trimers which mostly differ due to differences in domain orientation.
(B) Contact among RBDs within S trimers viewed from the apex along the 3-fold molecular axis of PRD- 0038 S (pink), linoleic acid-bound SARS-CoV-2 S (blue, PDB 6ZB544), and apo SARS-CoV-2 S (gray, PDB 6VXX5).



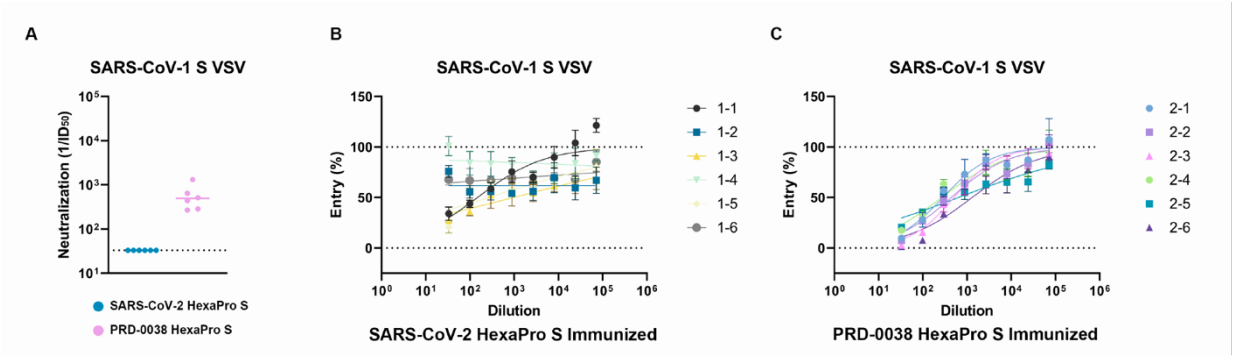
Supplemental Figure 4.7: Conservation analysis of epitopes targeted by monoclonal antibodies between PRD-0038 S and SARS-CoV-2 S (related to **Fig. 4.3**).

(A) Epitope mapped onto PRD-0038 RBD structure with sequence alignment of key residues at the interface. The PRD-0038 RBD is shown in gray (three RBDs are shown with distinct shades of gray for S2M11 which recognizes a quaternary epitope) and N-linked glycans are rendered as blue spheres. SARS-CoV-2 residue numbering is shown in gray and PRD-0038 residue numbering is shown in pink. The RBM is depicted with a black outline and the monoclonal antibody epitopes are colored according to their antigenic sites (I, purple; II, red; IV, violet; V, green). (B) PRD-0038 S sequence conservation with SARS-CoV-2 S at key S2 fusion machinery epitopes (indicated with dashed lines). Residues are colored according to sequence identity (orange : not conserved, green : conserved).
(C) Sequence alignment of corresponding S2 epitopes. SARS-CoV-2 residue numbering is shown in gray and PRD-0038 residue numbering is shown in pink.



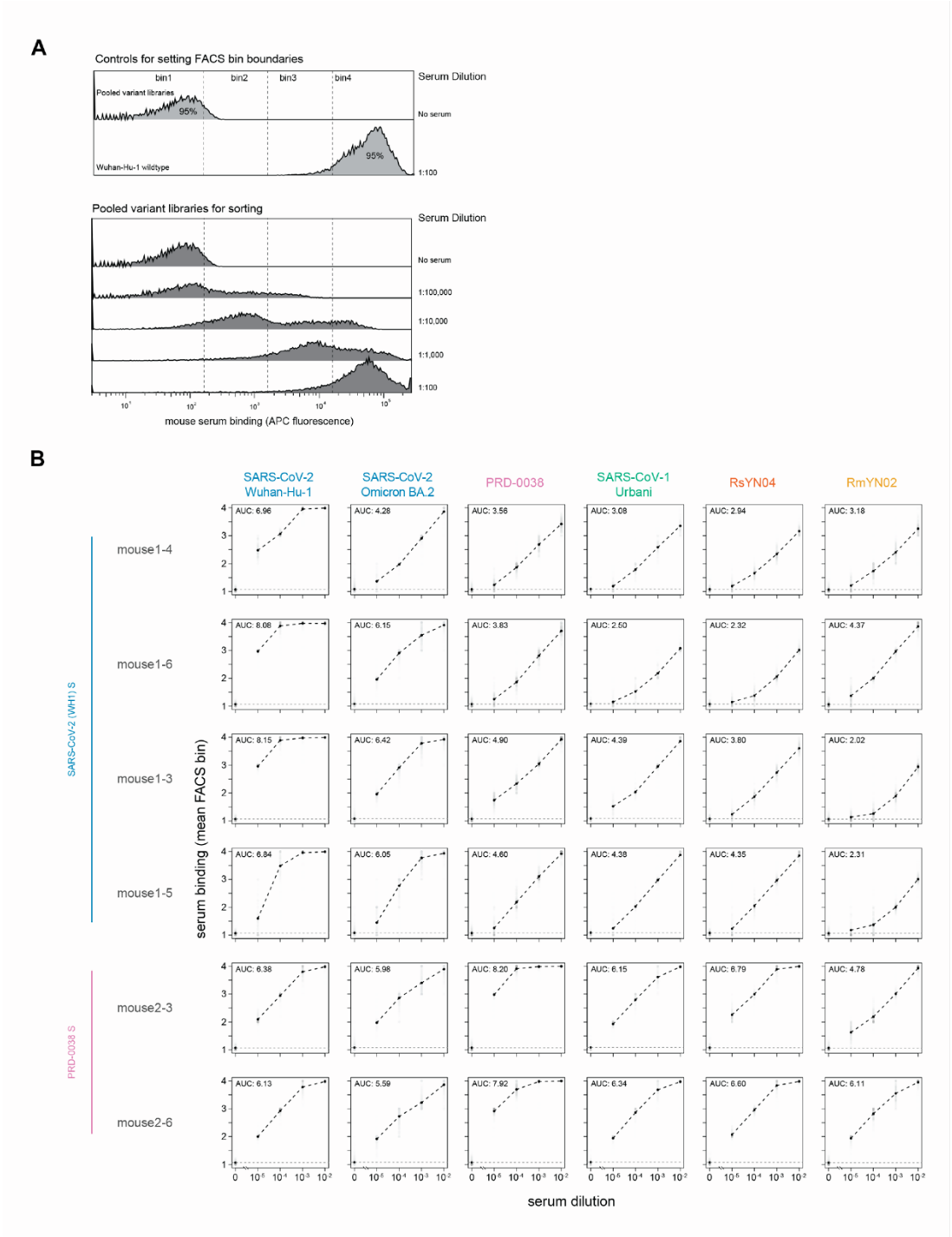
Supplemental Figure 4.8: Dose-response curves of mouse serum neutralization before and after three immunizations with SARS-CoV-2 HexaPro S or PRD-0038 HexaPro S (related to **Fig. 4.4**).

(A and B) Neutralization of the indicated sarbecovirus S VSV pseudotypes was assessed for each animal serum before immunization (A) and two weeks post dose 3 (week 8) (B), as indicated by the color key.



Supplemental Figure 4.9: Neutralization of SARS-CoV-1 S VSV by vaccine-elicited mouse sera using a highly diluted pseudovirus input (related to **Fig. 4.4**).

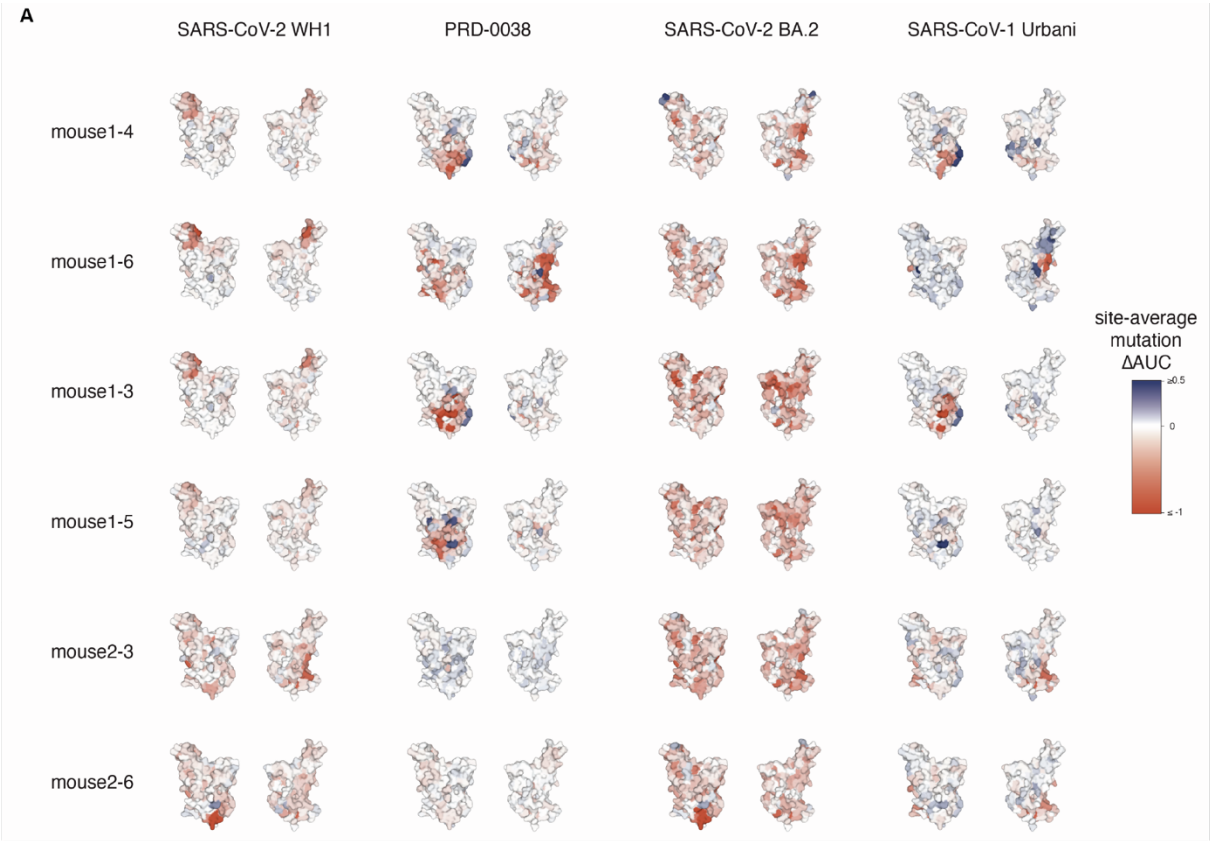
(A)Reciprocal ID50 value. Bar represents the geometric mean of each group. Limit of detection shown by a dotted line.
(B and C) Dose-response curves for neutralization of SARS-CoV-1 S VSV by SARS-CoV-2 HexaPro S- elicited sera (B) and PRD-0038 HexaPro S-elicited sera (C).



Supplemental Figure 4.10: Binding of vaccine-elicited mouse sera to yeast-displayed sarbecovirus RBDs (related to **Fig. 4.4**).

(A) Representative flow cytometry traces illustrating FACS binning scheme. Upstream FACS isolate single RBD+ cells via FSC/SSC and FITC(RBD)/FSC gates. Cells were then partitioned into bins based on APC fluorescence (serum binding). Bin 1 captures

95% of zero-serum cells, bin 4 captures 95% of 1:100 serum- labeled wildtype (SARS-CoV-2 Wuhan-Hu-1) control, and bins 2 and 3 split the remaining log-MFI scale evenly. Library samples labeled across a serum dilution series (mouse 1-5 serum traces shown for reference) were sorted into bins of serum-binding and deep sequenced to determine variant counts in each bin, enabling calculation of the “mean bin” into which each barcoded library genotype was sorted at each
sample concentration.
(B) Mean bin values determined from FACS-seq were plotted against serum dilution to generate an Area Under the Curve (AUC) metric (**Fig. 4.4C**). Curves underlying AUC calculation for select RBD genotypes are shown. Binding curves for all variants can be accessed at https://github.com/tstarrlab/SARSr- CoV_MAP_PRD0038-vaccine/blob/main/results/summary/compute_AUC.md#plot-all-curves). For each plot, transparent gray dots show the mean bin metric for each internal-replicate library barcode linked to the given RBD, and black points show the mean serum-binding for an RBD averaged across all barcodes. AUC is calculated as the area between the dashed black line and the baseline determined from the zero-
serum sample sort (dashed gray line).



Supplemental Figure 4.11: DMS epitope mapping of serum antibodies (related to Fig. 4.4).

(A) Evaluation of epitopes targeted by serum antibodies elicited by SARS-CoV-2 HexaPro S (mouse 1-3, 1-4, 1-5, and 1-6) or PRD-0038 HexaPro S (mouse 2-3 and 2-6) vaccination using yeast-displayed DMS of vaccine-matched and mismatched RBDs (indicated above each column). The average effect of mutations at each site are mapped to the SARS-CoV-2 structure, where blue and red indicate positions where mutations increase or decrease serum binding, respectively.

Table S4.1: Rhinolophus ACE2 alleles and contact residues (related to Fig. 4.1).

			ACE2-RBD Contact Residues													
Species	**Isolate #**	**Genbank**	**24**	**27**	**30**	**31**	**34**	**35**	**38**	**41**	**42**	**45**	**83**	**330**	**353**	**355**
R. affinis	9479	MT394208.1	R	I	D	N	H	E	D	Y	Q	L	Y	N	K	D
R. affinis	787	MT394203.1	R	I	D	N	R	E	E	Y	Q	L	Y	N	K	D
R. sinicus	3364	MT394200.1	R	I	D	E	S	E	D	Y	K	L	Y	N	K	D
R. sinicus	WJ4	MT394181.1	L	I	D	E	F	E	N	Y	Q	L	Y	N	K	D
R. sinicus	-	GQ262791.1	L	I	D	E	S	E	N	Y	Q	L	Y	N	K	D
R. sinicus	1446	MT394194.1	R	T	D	E	S	E	N	Y	Q	L	Y	N	K	D
R. sinicus	WJ1	MT394187.1	R	I	D	T	S	E	D	Y	Q	L	Y	N	K	D
R. sinicus	1434	MT394197.1	R	M	D	T	S	E	D	Y	Q	L	Y	N	K	D
R. sinicus	3358	MT394193.1	E	M	D	K	T	K	D	H	Q	L	Y	N	K	D
R. sinicus	1438	MT394184.1	E	I	D	K	T	K	D	H	Q	L	Y	N	K	D
R. alcyone	-	KR559016.1	L	I	D	N	S	E	N	H	Q	L	F	N	K	D
R. landeri	-	KR559015.1	L	T	D	D	S	A	N	Y	Q	L	F	N	K	D
R. ferrumequinum	-	XM_033107295.1	L	K	D	F	S	E	N	H	Q	L	F	N	K	D
R. ferrumequinum	-	FJ598617.1	L	T	E	K	T	E	D	Y	Q	L	Y	K	K	D
Human	-	BAB40370.1	Q	T	D	K	H	E	D	Y	Q	L	Y	N	K	D

Table S 4.2: CryoEM data collection and refinement statistics (related to Fig. 4.2 and 4.3).

Table S2 CryoEM data collection and refinement statistics (related to Figure 2 and 3).

	PRD-0038 RBD - *R. alcyone* ACE2 EMD-41786	PRD-0038 RBD - *R. alcyone* ACE2 (Local refinement) PDB 8U0T EMD-41784	PRD-0038 S PDB 8U29 EMD-41842	PRD-0038 S (NTD local refinement) EMD-41843
Data collection and processing				
Magnification	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300
Electron exposure (e^{-}/Å^2)	60	60	60	60
Defocus range (μm)	0.2 - 3.0	0.2 - 3.0	0.1 - 3.0	0.1 - 3.0
Pixel size (Å)	0.843	0.843	0.843	0.843
Symmetry imposed	C2	C1	C3	C1
Final particle images (no.)	304,806	284,934	103,347	310,041
Map resolution (Å)	3.5	3.2	2.8	2.9
FSC threshold	0.143	0.143	0.143	0.143
Map sharpening *B*factor (Å^2)	-133	-118	-82.6	-89.7
Validation				
MolProbity score		1.08	1.28	
Clashscore		1.62	1.69	
Poor rotamers (%)		0.47	1.13	
Ramachandran plot				
Favored (%)		97.02	95.44	
Allowed (%)		2.86	4.38	
Disallowed (%)		0.12	0.19	

Table S 4.3: Representative KD values determined from BLI binding analysis (related to Fig. 4.2).

	KD (M)	KD Error	ka (1/Ms)	ka Error	kdis (1/s)	kdis Error
PRD-0038 WT RBD : monomeric *R. alcyone* ACE2	5.07E-08	1.64E-10	1.64E+05	5.10E+02	8.33E-03	7.80E-06
PRD-0038 T487W RBD : monomeric *R. alcyone* ACE2	2.12E-08	3.96E-11	1.24E+05	2.18E+02	2.62E-03	1.63E-06
PRD-0038 WT RBD : monomeric *R. landeri* ACE2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
PRD-0038 T487W RBD : monomeric *R. landeri* ACE2	1.88E-07	4.22E-10	4.49E+04	9.64E+01	8.43E-03	5.62E-06

* N.D. - not determined due to weak binding

CHAPTER 5. MOLECULAR BASIS OF CONVERGENT EVOLUTION OF ACE2 RECEPTOR UTILIZATION AMONG HKU5 CORONAVIRUSES

DPP4 was considered a canonical receptor for merbecoviruses until the recent discovery of African bat-borne MERS-related coronaviruses using ACE2. The extent and diversity of ACE2 utilization among merbecoviruses and their receptor species tropism remain unknown. In this story, I detail our work in understanding the basis of HKU5 cellular entry utilizing *Pipistrellus abramus* (P.abr) and several non-bat mammalian ACE2s. I reveal the molecular mechanisms of this association through a binding mode distinct from that of any other known ACE2-using coronaviruses. I describe, our finding of a single amino acid mutation enabling HKU5 to utilize human ACE2, providing proof-of-principle for machine-learning assisted outbreak preparedness. Furthermore we show that MERS-CoV and HKU5 have markedly distinct antigenicity and identified several HKU5 inhibitors, including two clinical compounds. Our findings profoundly alter our understanding of coronavirus evolution, as several merbecovirus clades independently evolved ACE2 utilization, and pave the way for developing countermeasures against viruses poised for human emergence.

Adapted from: Park et al. (2024). Molecular Basis of Convergent Evolution of ACE2 Receptor Utilization among HKU5 Coronaviruses. BioRxiv. (**Preprint**)

5.1 Chapter Introduction

Coronaviruses comprise a large group of pathogens circulating in mammalian and avian hosts. 229E, NL63, HKU1, and OC43 are endemic in the human population. SARS-CoV-1 caused an epidemic in 2002-2004 with an approximately 8% fatality rate(Drosten, 2003; Ksiazek, 2003). MERS-CoV is highly lethal, with a 36% case fatality

rate, and has been circulating sporadically since its discovery in 2012(Zaki, 2012; Mok, 2020; Ngere, 2022) whereas SARS-CoV-2 is responsible for the COVID-19 pandemic(Zhu, 2020; Huang, 2020) with recurrent waves sweeping through the population(Cameroni, 2022; Viana, 2022). SARS-CoV-1 and SARS-CoV-2 (*Sarbecovirus* subgenus) originated in bats (Ge, 2012; Hu, 2017; Li, 2005; Yang, 2015; Zhou, 2020) and might have been transmitted to humans through intermediate hosts (Wang, 2005; Consortium, 2004; Kan, 2005). MERS-CoV (*Merbecovirus* subgenus) possibly originated in bats although spillovers occurred repeatedly from camels(Ngere, 2022; Haagmans, 2014; Sabir, 2016; Corman, 2014; Lau, 2019; Anthony, 2017), emphasizing the zoonotic threat of these coronaviruses.

The spike glycoprotein (S) is anchored in the viral membrane and required for cellular entry. S comprises an N-terminal S_1 subunit (receptor-binding) and a C-terminal S_2 subunit (membrane fusion)(Walls, 2016;Tortorici, 2019; Walls, 2020) and is a main target of neutralizing antibodies, a correlate of protection against coronaviruses (Piccoli, 2020; Arunachalam, 2021; Corbett, 2021; McCallum, 2021; Bowen, 2022; Barnes, 2020). Host protease-mediated S cleavage occurs at the S_1/S_2 junction for many coronaviruses, yielding non-covalently linked S_1 and S_2 subunits in the prefusion S trimer, and at the S_2' site (adjacent to the fusion peptide) for all coronaviruses, leading to large-scale fusogenic conformational changes(Millet, 2015; Walls, 2017). Coronavirus invasion of target cells thus requires the concerted action of receptor-binding and proteolytic S processing to lead to productive infection and both factors modulate the ability of coronaviruses to cross species barrier.

Our ability to prepare for zoonotic spillovers is undermined by the knowledge gap between discovery of a large number of coronaviruses and our understanding of the likelihood of their spillover to the human population. This is compounded by the ability of coronaviruses to undergo extensive antigenic changes and recombinations which can alter host receptor tropism through ortholog adaptation or receptor switch (Tolentino,2024). For instance, N501Y-harboring SARS-CoV-2 variants acquired the ability to use mouse ACE2(Niu, 2021) which was further enhanced with the addition of the Q493R mutation(Cameroni, 2022; McCallum, 2022). Furthermore, a recombination event in the merbecovirus S glycoprotein led to an unexpected receptor switch with MERS-CoV utilizing DPP4 (Tolentino, 2024; Raj, 2013; Chen, 2023; Wang, 2014) whereas a few African bat-borne merbecoviruses use ACE2(Xiong, 2022; Ma, 2023). However, the extent of and diversity with which merbecoviruses engage ACE2 and their host receptor tropism remain poorly understood.

Here, we show that the HKU5 clade of merbecoviruses, which was first described in 2006 in *Pipistrellus abramus* (Woo, 2006) (P.abr), can unexpectedly use ACE2 from P.abr and several other mammalian species, especially from the Artiodactyl order, for cell entry. We determined cryo-electron microscopy structures of the HKU5 receptor-binding domain (RBD) complexed with P.abr and with Bos taurus (B. tau) ACE2 orthologs along with that of the prefusion HKU5 S trimer. Our data reveal a receptor-binding mode distinct from that of any other known ACE2-using coronaviruses and enabled identification of the molecular determinants of receptor species tropism. We show that although MERS-CoV infection-elicited polyclonal plasma antibodies exhibit very weak to no HKU5 cross-neutralization, several broad-spectrum coronavirus

inhibitors (including two clinical compounds) block HKU5 propagation in human cells. Our findings unveil that the HKU5 clade of merbecovirus independently evolved ACE2 utilization, underscoring the extraordinary diversity of ACE2 recognition modes among coronaviruses, and paving the way for developing HKU5 countermeasures against a clade of pathogen poised for human spillover.

5.2 HKU5 utilizes ACE2 as entry receptor

Merbecoviruses comprise a large diversity of pathogens found in humans, bats, camels, pangolins and hedgehogs(Drosten, 2003; Ksiazek, 2003; Ngere, 2022; Haagmans, 2014; Sabir, 2016; Corman, 2014; Lau, 2019; Anthony, 2017; Chen, 2023; Ithete, 2013). Phylogenetic classification based on their S receptor-binding domain (RBD) amino acid sequences delineates at least six merbecovirus clades **(Fig. 5.1A)**, many of which do not have a confirmed host receptor. To investigate the receptor usage of the HKU5 merbecovirus clade, we evaluated binding of dimeric HKU5-19s and HKU5-1 RBDs N-terminally fused to a human IgG Fc fragment (RBD-hFc) to the surface of HEK293T cells transiently transfected with a panel of 117 mammalian ACE2 orthologs across the phylogeny (**Fig. 5.1B and Fig. S5.1A-B**). Although we only detected binding to P.abr ACE2 among the panel of 62 bat orthologs tested, both RBDs also bound to several non-bat mammalian ACE2s, including B.tau, *Ovis aries* (O.ari), *Bos mutus* (B.mut), *Odocoileus virginianus texanus* (O.vir), and *Bubalus bubalis* (B.bub) ACE2s (**Fig. 5.1C-D and Fig. S5.1C-D and Fig. S5.2A**). Furthermore, the HKU5-19s (but not the HKU5-1) RBD-hFc interacted with *Mustela erminea* (M.erm) ACE2. Accordingly, we observed efficient entry of vesicular stomatitis virus (VSV) pseudotyped

with HKU5-19s and HKU5-1 S into HEK293T cells transfected with P.abr, B.tau, O.ari, B.mut, O.vir, B.bub ACE2s and *Rattus norvegicus* (R.nor) ACE2 as well as with M.erm ACE2 for HKU5-19s S VSV (**Fig. 5.1C-D and Fig. S5.2B-E**). The observation of pseudovirus entry mediated by R.nor ACE2 for which no RBD binding was detected likely results from avidity due to multivalent presentation of S trimers on pseudoviruses along with the use of trypsin to enhance pseudovirus entry[40,44,45]. Moreover, efficient HKU5-19s S-mediated cell-cell membrane fusion was detected in Caco-2 cells expressing these ACE2 orthologs (**Fig. S5.2F**). Assessment of RBD-hFc binding by four additional HKU5 strains and PSV entry using S glycoprotein for three additional ones showed that they all use P.abr ACE2 but exhibited different receptor species tropism (**Fig. S5.2G-J**). These data suggest that most HKU5 strains use the host P.abr ACE2 as receptor, concurring with a recent report[46] and with the fact that currently available sequences were obtained from P.abr bat samples. Although all HKU5 strains harbor a minimal furin cleavage site at the S_1/S_2 junction, they exhibit varying degrees of S proteolytic cleavage during biogenesis (**Fig. S5.2J**). HKU5-19s S has a suboptimal furin cleavage site (RLAR) compared with HKU5-1 (RVRR) whereas some other HKU5 strains, such as HKU5-JX2020-Q279, harbor two consecutive suboptimal furin cleavage sites (RLARSAR) (**Fig. S5.3**).

Caco-2 cells stably overexpressing P.abr ACE2 (Caco-2/P.abr ACE2) enabled robust spread of propagation-competent recombinant VSVs encoding HKU5-19s S or HKU5-1 S instead of VSV-G (psVSV-HKU5-19s or psVSV-HKU5-1) and authentic HKU5-1, lending further support to the P.abr ACE2 receptor function (**Fig. 5.1E-F**). Furthermore, addition of exogenous trypsin promoted detectable (albeit weak)

replication of psVSV-HKU5-1 and authentic HKU5-1[47] in Caco-2 cells with or without stable hACE2 overexpression (this cell line endogenously expresses hACE2), in line with prior data[48]. These findings indicate that HKU5 utilizes its native bat host ACE2 as entry receptor along with several non-bat mammalian ACE2s from the Artiodactyl, Carnivore and Rodent orders.

5.3 Molecular basis of HKU5 recognition of ACE2

To understand HKU5 engagement of its host receptor, we characterized the HKU5-19s RBD in complex with a P.abr ACE2 ectodomain construct comprising the peptidase and dimerization domains using cryo-EM. Data processing revealed a predominance of a monomeric form of ACE2 (accounting for 99% of particle images in selected 2D class averages despite the presence of the native dimerization domain) which led us to determine an asymmetric 3D reconstruction focused on the ACE2 peptidase domain and bound RBD at 3.1 Å resolution (**Fig. 5.2, Fig. S5.4A-E and Table 1**). The HKU5 RBD binds to the P.abr ACE2 peptidase domain and buries a surface of ~950 $Å^2$ from each of the two binding partners through interactions dominated by hydrogen bonding, salt bridging and shape complementarity.

The HKU5 receptor-binding motif (RBM) folds as a compact four-stranded antiparallel β-sheet flanked by short α-helices on both sides and forms two consecutive insertions in the five-stranded antiparallel RBD core β-sheet (**Fig. 5.2A**). Specifically, HKU5 residues D460, M461, Y464, E472, I473, Y508, T510, E511, Y513, T515, S516, A517, Y518, G519, K520, Y522, K544, Y545, Q546, S547, G551, T556, Y558 and Y560 form the RBM which interacts with ACE2 residues R26, L29, V30, N33, H34, E37, N38, H41, D90, I92, I93, Q96, M322, T323, P324, G325, W327, R328, D329, K352, N353,

D354, R356, A385, N386, Q387, S388 and R392. The P.abr ACE2 α-helix 323-330 docks against the HKU5 RBM and forms a constellation of interactions, including $R328_{P.abrACE2}$ forming polar interactions with the $Y508_{HKU5}$, $Y558_{HKU5}$ and $D460_{HKU5}$ side chains along with pi-stacking interactions with the $Y560_{HKU5}$ side chain and $D329_{P.abrACE2}$ forming a salt bridge and a hydrogen bond with $K544_{HKU5}$ and $Y558_{HKU5}$, respectively (**Fig. 5.2B**). Furthermore, $K352_{P.abrACE2}$ is hydrogen-bonded to the $Y545_{HKU5}$ backbone carbonyl and the $N353_{P.abrACE2}$ side chain is hydrogen-bonded to the $T510_{HKU5}$ backbone carbonyl (**Fig. 5.2C**). At the tip of the RBM, $Y518_{HKU5}$ is hydrogen-bonded to $R26_{P.abrACE2}$ and $D90_{P.abrACE2}$ whereas $K520_{HKU5}$ forms polar interactions with the $N386_{P.abrACE2}$ side chain and backbone carbonyl along with the $A385_{P.abrACE2}$ backbone carbonyl (**Fig. 5.2C**).

Analysis of amino acid residue conservation reveals a marked RBM sequence diversity among HKU5 strains (**Fig. 5.2D**), possibly modulating receptor engagement. To understand the impact of RBM residue polymorphism on receptor recognition, we assessed binding of the aforementioned P.abr ACE2 ectodomain construct to three distinct HKU5 RBD strains (HKU5-19s, HKU5-1 and HKU5-33s) using biolayer interferometry (BLI). The HKU5-19s and HKU5-33s RBMs differ at positions $D_{HKU5\text{-}19s}460S_{HKU5\text{-}33s}$, $Y_{HKU5\text{-}19s}508H_{HKU5\text{-}33s}$, $T_{HKU5\text{-}19s}515F_{HKU5\text{-}33s}$, $S_{HKU5\text{-}19s}516T_{HKU5\text{-}33s}$, $K_{HKU5\text{-}19s}544D_{HKU5\text{-}33s}$, $Y_{HKU5\text{-}19s}545R_{HKU5\text{-}33s}$ and $Y_{HKU5\text{-}19s}558S_{HKU5\text{-}33s}$. HKU5-33s harbors two short insertions between residues 549 and 550 (QELP) and between residues 552 and 553 (F) of the HKU5-19s RBD (**Fig. 5.2E**). The HKU5-19s and HKU5-1 RBMs are much more similar to each other and only differ due to the $E_{HKU5\text{-}19s}472A_{HKU5\text{-}1}$ substitution. We observed that P.abr ACE2 bound with a slightly higher apparent

binding affinity to the HKU5-1 RBD ($K_{D,app}$=21.5 nM) than to the HKU5-19s RBD ($K_{D,app}$=28.0 nM) whereas binding was weakest for the HKU5-33s RBD ($K_{D,app}$=50.2nM) (**Fig. 5.2F-H and Fig. S5.5**). These findings indicate that the extensive RBM polymorphism among HKU5 strains modulates receptor binding affinity and breadth **(Fig. S5.2A)** and possibly antigenicity, as observed with SARS-CoV-2 variants[36,49–55].

Comparison of the structures of the P.abr ACE2-bound HKU5 RBD, *Pipistrellus pipistrellus* (P.pip) ACE2-bound NeoCoV RBD[40], hDPP4-bound MERS-CoV RBD[56,57], and hDPP4-bound HKU4[39] RBD shows that HKU5, MERS-CoV and HKU4 utilize the same side of the RBM β-sheet to engage their cognate receptors (**Fig. 5.2I-L**). Conversely, NeoCoV recognizes ACE2 via the tip of the RBM, including the distal β-hairpins, which are shorter for NeoCoV and HKU5 relative to MERS-CoV and HKU4. In spite of a conserved overall architecture of the RBMs of these four viruses, their fine molecular differences mediate the three observed binding modes: HKU5 and ACE2, NeoCoV and ACE2 as well as MERS-CoV/HKU4 and DPP4.

5.4 Determinants of ACE2 species tropism

To functionally validate the role of the ACE2-interacting residues identified in our cryo-EM structure, we evaluated binding of P.abr ACE2 to HKU5-19s RBD point mutants each harboring a residue substitution designed to interfere with receptor engagement. Using BLI, we observed that the HKU5-19s Y464A, A517R, G519W and Y558A substitutions abrogated binding entirely whereas the Y518A substitution did so almost completely (**Fig. 5.3A**). Furthermore, the K520G and Y545G RBD mutations also disrupted P.abr ACE2 binding in the HKU5-19s RBD background (**Fig. 5.3A**), possibly reducing receptor binding affinity for the HKU5 isolates in which these changes naturally occur although epistasis may modulate this effect[50]. Consistently, the HKU5-19s Y464A,

A517R, Y518G, G519W, K520G, Y545G, and Y558G RBD-hFc mutants were impaired to varying degrees in their ability to bind P.abr ACE2 transiently expressed at the surface of HEK293T cells (**Fig. 5.3B**). Moreover, HKU5-19s S VSV pseudovirus harboring the Y464A, A517R, K520G, or Y558G substitutions exhibited the most significant reduction of entry into HEK293T cells transiently expressing P.abr ACE2 (**Fig. 5.3C**).

Assessment of the impact of mutations of key P.abr ACE2 residues contacting the HKU5 RBD further supported these findings: the P324G/Q, R328N (glycan knockin), D329A and R328A/D329A substitutions, but not R26A/H34G/N38G or DPI90-92NLT substitutions, dampened HKU5-19s binding to and pseudovirus entry in HEK293T cells transiently transfected with these P.abr ACE2 mutants (**Fig. 5.3D,F**). Given that several HKU5-interacting P.abr ACE2 residues are conserved or conservatively substituted in human ACE2 (hACE2, **Table 5.2**), we set out to determine the molecular determinants of the lack of efficient hACE2 binding and utilization (**Fig. 5.1E-F**). Concurring with the central role of the P.abr ACE2 R328 and D329 residues, we found that a hACE2 double mutant harboring these residues (E329R/N330D) promoted HKU5-19s S pseudovirus entry into cells (**Fig. 5.3E,G**). A hACE2 triple mutant harboring the E329R/N330D substitutions along with F40S, N322Y, Q325P or G354R promoted more robust HKU5-19s S pseudovirus entry along with detectable binding of the RBD-hFc constructs except for the E329R/N330D/G354R (**Fig. 5.3E,G**). Although the $N321_{P.abrACE2}$ glycosylation sequon is conserved relative to hACE2 (equivalent to $N322_{hACE2}$), no evidence of glycosylation could be detected in the HKU5 RBD/P.abr ACE2 complex cryo-EM map (as is the case for P.abr ACE2 N38). The positioning of the $N322_{hACE2}$

oligosaccharide in SARS-CoV-2 RBD/hACE2 structures[36,58] suggests that it could dampen HKU5 RBD binding, thereby explaining the improved hACE2 usage resulting from the N322Y substitution in an E329R/N330D background (**Fig. 5.3E,G**). This is further demonstrated by a loss of detectable HKU5 RBD binding after introducing the H322N mutation in O.vir ACE2 (**Fig. 5.3H**).

Unlike HKU5-19s, the HKU5-1 RBD-hFc did not bind M.erm ACE2, likely due to the $E472_{HKU5\text{-}19s}A_{HKU5\text{-}1}$ substitution as the $E472_{HKU5\text{-}19s}$ side chain might form a salt bridge with $R354_{M.ermACE2}$ whereas $A472_{HKU5\text{-}1}$ would not (**Fig. 5.1C-D, Fig. S5.2F and Fig. 5.3H**). Accordingly, the M.erm ACE2 R354N or R354G mutation reduced HKU5-19s RBD-hFc binding, highlighting the importance of the residue at position $354_{hACE2/M.ermACE2}$ (equivalent to residue $353_{P.abrACE2}$) for receptor tropism (**Fig. 5.3H)**. Similar to HKU5-19s, a recently discovered HKU5 sequence obtained from deceased farmed minks[59] (Neogale vison, N.vis, Carnivora order) harbors E472, likely explaining the binding of HKU5-19s to N.vis ACE2 through the possible formation of a salt bridge with $H353_{N.visACE2}$ (**Fig. S5.5D-F**). No N.vis ACE2 binding was detected to the HKU5-1 RBD, likely resulting from loss of this salt bridge due to the $E472_{HKU5\text{-}19s}A472_{HKU5\text{-}1}$ substitution, or to the HKU5-33s RBD which has an RBM markedly different from that of HKU5-19s. Overall, we identified and validated the functional impact of four determinants of receptor species tropism, all of which are situated in the HKU5-binding footprint, along with preferred amino acid residue for each of them (**Fig. 5.3I-J**). These include P.abr ACE2 positions (i) 321 where tyrosine or histidine are preferred (disrupting an N-linked glycosylation sequon), (ii) 324 where a proline is preferred, (iii) 328-329 where arginine and aspartate residues are preferred, respectively, and (iv) 353 where

an asparagine or arginine is preferred by HKU5-19s. Other determinants of species tropism may exist within the HKU5-binding footprint but their contribution is likely less important than that of the aforementioned determinants.

5.5 HKU5 utilization of artiodactyl and human ACE2

To understand the molecular basis of binding between the HKU5 RBD and a large number of artiodactyl ACE2 orthologs **(Fig. 5.1D)**, we determined a cryo-EM structure at 2.4 Å resolution of the HKU5-19s RBD in complex with a B.tau ACE2 ectodomain construct comprising the peptidase and dimerization domains **(Fig. 5.4A and FigS5.4)**. In contrast to the P.abr ACE2 complex, most particle images comprised dimeric ACE2 with each peptidase domain bound to an RBD enabling the use of C2 symmetry.

Binding of B.tau ACE2 to the HKU5-19s RBD buries an average surface of approximately 900Å^2 from each of the two binding partners with a binding mode virtually identical to that observed for the P.abr ACE2-bound HKU5-19s RBD **(Fig. 5.4A-C).** Although the RBD contacts the same regions of the two ACE2 orthologs, several key interactions are distinct and likely explain the stronger binding affinity and receptor activity of P.abr ACE2 relative to B.tau ACE2. We observed remodeling of the contacts between the B. tau ACE2 α-helix 325-330 and the RBD, due to the $P_{P.abr}324Q_{B.tau}$ and $R_{P.abr}328D_{B.tau}/D_{P.abr}329N_{B.tau}$ receptor determinant substitutions, which are similar to the equivalent $Q325_{hACE2}$, $E329_{hACE2}$ and $N330_{hACE2}$ residues **(Fig. 5.4B)**. Specifically, the $Q324_{B.tau}$ side chain interacts through van der Waals interactions with $M461_{HKU5}$ and is hydrogen-bonded to $Y464_{HKU5}$ whereas $D328_{B.tau}$ is hydrogen-bonded to $Y558_{HKU5}$ and $N329_{B.tau}$ is hydrogen-bonded to to $K544_{HKU5}$ and to $Y558_{HKU5}$. Moreover, the $N_{P.abr}353G_{B.tau}$ receptor determinant substitution further reinforces the similarity with the

equivalent $G354_{hACE2}$ residue, leading to reduced interactions with the RBD **(Fig. 5.4C)**. The $D_{P.abr}90N_{B.tau}$ (similar to $N90_{hACE2}$) and $N_{P.abr}386A_{B.tau}$ mutations respectively introduce and remove an N-linked glycan near the tip of the RBM, likely modulating binding **(Fig. 5.4A)**.

We found that B.tau ACE2 bound more strongly to the HKU5-19s RBD ($K_{D,app}$= 131.9 nM) than to the HKU5-1 RBD ($K_{D,app}$= 535 nM) whereas no binding to the HKU5-33s could be detected **(Fig. 5.4E-G)**. These findings are likely resulting from formation of water-mediated hydrogen bonds between the $E472_{HKU5\text{-}19s}$ side chain and the ACE2 Q324 and G325 backbone amides (**Fig. 5.4C**, not possible for $A472_{HKU5\text{-}1}$) and from the extensive RBM polymorphism of HKU5-33s relative to HKU5-19s.

Based on the resemblance at several key receptor specificity determinants between B.tau ACE2 and hACE2 along with the observation of weak but detectable hACE2-mediated HKU5-1 propagation **(Fig. 5.1E-F)**, we set out to assess if HKU5 was poised to acquire more efficient hACE2 utilization for outbreak preparedness. Given the identification of the hACE2 helix spanning residues 325-330 as a major tropism barrier for HKU5 (**Fig. 5.3E,G**), we computationally screened for mutations of RBD residues that interface with this receptor region using proteinMPNN (Dauparas, 2022) and evaluated their compatibility with hACE2 binding using AlphaFold3(Abramson, 2024). We identified the Y560K residue substitution as a likely candidate for promoting hACE2 utilization **(Fig. 5.4D)** and evaluated its effect on receptor binding and pseudovirus entry. Biolayer interferometry analysis revealed that the HKU5-19s Y560K RBD mutant, but not the wildtype HKU5-19s or HKU5-1 RBDs, promoted detectable (albeit weak) binding to the hACE2 ectodomain dimer **(Fig. 5.4H)**. Accordingly, we observed entry of

the HKU5-19s Y560K S, but not wildtype HKU5-19s and HKU5-33s S, VSV pseudovirus into HEK293T cells transiently transfected with hACE2 upon pseudovirus treatment with trypsin. Furthermore, HKU5-19s Y560K S retained its ability to enter HEK293T cells expressing P.abr ACE2 although at a reduced level compared to wildtype HKU5-19s S. These results suggest that HKU5 could acquire more efficient hACE2 usage through a single amino acid change and should be prioritized for the development of countermeasures **(Fig. 5.4I-J)**.

5.6 Architecture of the prefusion HKU5 S glycoprotein trimer

Prior studies showed that coronavirus RBDs can transition from a closed to open state, enabling exposure of the RBM(Walls, 2019; Pallesen, 2017; Yuan, 2017; Pronker, 2023; Wrobel, 2020). These conformational changes along with glycan masking enable to balance immune evasion and the need to gain access to target host cells(Walls, 2016; Tortorici, 2022; Tortorici, 2024). As a result, coronavirus S trimers have evolved to rely on a conformational ratchet mechanism which allosterically couples receptor binding to exposure of the fusion peptide and membrane fusion(Walls, 2017; Liu, 2024; Walls, 2019; Low, 2022). To unveil the architecture and conformational landscape of HKU5-19s S, we used cryo-EM to characterize an ectodomain trimer construct stabilized in the prefusion state. 3D classification of the data revealed the presence of two populations in our dataset, mostly differing by the position of the C and D domains, from which we determined two structures at 2.0Å resolution **(Fig. 5.5A and Fig. S5.6)**.

The structures reveal that S adopts a prefusion closed conformation with all three RBDs tightly packed at the trimer apex, and resolve the entire fusion peptide region, a linoleic acid-like density in each RBD, and several molecules bound to the NTD, reminiscent of the locked SARS-CoV-2 S state(Wrobel, 2020; Xiong, 2020; Toelzer,

2020) **(Fig. 5.5A and Fig. S5.6)**. In this conformation, the RBM is partially masked by quaternary contacts and glycan N419 from a neighboring protomer and thus inaccessible for ACE2 binding **(Fig. 5.5A-B)**. Conformational masking and glycan shielding of the RBM occurs in a very similar way for betacoronaviruses MERS-CoV (N410 glycan)(Park, 2019), PDF-2180 (N414 glycan)(Xiong, 2022), BtCoV-422 (N414 glycan)(Tse, 2023) **(Fig. 5.5B)** and HKU1(McCallum, 2024), pointing to a conserved mechanism of immune evasion, as well as in the context of the distinct alphacoronavirus S architecture for NL63(Walls, 2016), CCoV-HuPn-2018(Tortorici, 2022) and FCoV-23(Tortorici, 2024).

Mapping the S glycoprotein conservation among merbecoviruses from the six aforementioned clades shows that the RBD and the NTD are the most variable regions **(Fig. 5.5C)**. This observation concurs with the fact that these two domains are involved in recognition of various receptors(Raj, 2013; Wang, 2014; Xiong, 2022; Ma, 2024; Yang, 2014) and that they are the main targets of neutralizing antibodies(Addetia, 2024). Although the S_1/S_2 site is disordered in our structure, the S_2' polybasic $R_{882}KYR_{885}$ cleavage site and adjacent fusion peptide regions are entirely resolved including glycans at positions N761 and N868 which participate in masking the S_2' R885 residue **(Fig. 5.5D)**. The conservation of this region, along with that of the stem helix, relative to other coronaviruses, suggests that HKU5 will be sensitive to antibodies targeting these fusion machinery epitopes(Sauer, 2021; Pinto, 2021; Zhou, 2023;Li, 2022) **(Fig. 5.5E)**. Overall, these data reveal the architecture of a coronavirus S trimer in unprecedented detail and provide a blueprint for guiding the selection and design of inhibitors and vaccines against this virus.

5.7 Identification of HKU5 countermeasures

To understand antigenic relationships among merbecoviruses, we assessed the ability of MERS-CoV infection-elicited human antibodies to inhibit HKU5-19s S VSV pseudovirus entry into HEK293T cells transiently transfected with P.abr ACE2. We used plasma collected from individuals who were hospitalized with MERS-CoV infection (prior to the COVID-19 pandemic) and selected samples at time points corresponding to peak neutralization titers against MERS-CoV EMC/2012(Addetia, 2024) **(Table 5.3)**. Neutralizing activity (reaching at least 50% inhibition of entry at 1/10 plasma dilution) was detected for three (15, 16 and 24) out of 28 plasma samples evaluated **(Fig. 5.6A and Fig. S5.7)**. The rare and very weak HKU5 cross-neutralization mediated by MERS-CoV-elicited polyclonal plasma antibodies underscores the genetic and antigenic distance of the S and RBD of these two viruses, sharing approximately 65% and 50% amino acid sequence identity, respectively. In contrast, none of the plasma samples evaluated obtained from individuals infected once by SARS-CoV-2 had activity reaching 50% inhibition of entry **(Fig. 5.6A and Fig. S5.7)**. However, broadly neutralizing monoclonal antibodies targeting the fusion machinery (S_2 subunit) stem helix(Sauer, 2021; Pinto, 2021) (S2P6 and B6) or the fusion peptide/S_2' cleavage site(Low, 2022; Sun, 2022) (76E1) inhibited authentic HKU5-1 propagation in Caco-2/P.abr ACE2 cells **(Fig. 5.6B)** and HKU5-19s S VSV pseudovirus **(Fig. S5.7)**, consistent with the high degree of conservation of the targeted epitopes.

We subsequently screened a panel of small molecule inhibitors to evaluate their ability to block propagation of authentic HKU5-1 in Caco-2/P.abr ACE2 cells. Camostat, which inhibits TMPRSS2 and related serine proteases(Hoffmann, 2020; Fraser, 2022),

but not the E64d cysteine protease (cathepsin B/L) inhibitor, abrogated HKU5-1 propagation at a concentration of 10 μM **(Fig. 5.6C)**. Bafilomycin A1, which hinders endosomal acidification(Icho, 2022), also inhibited HKU5-1 at a concentration of 1 μM **(Fig. 5.6C)**. These findings suggest that HKU5-1 can infect Caco-2/P.abr ACE2 cells via TMPRSS2-mediated plasma membrane fusion and via endosomal fusion, the latter process relying on non-cysteine proteases, possibly aspartyl or serine proteases. The involvement of TMPRSS2 and related TMPRSS proteases in HKU5-1 entry into Caco-2/P.abr ACE2 cells is consistent with the presence of a polybasic cleavage site at the S_1/S_2 junction ($R_{742}VRR_{745}$, HKU5-1 residue numbering), resulting in cleavage of the S trimers incorporated in pseudovirus during biogenesis **(Fig. S5.2J and Fig. S5.3)**, which was previously described to lead to plasma membrane fusion for MERS-CoV(Park, 2016; Millet, 2014). Conversely, the inability of E64d to reduce HKU5-1 infection of Caco-2 cells is likely explained by the suboptimal endosomal cysteine protease site ($N_{760}FTS_{763}$), which introduces an N-linked glycosylation sequon (Menachery, 2020; Biniossek, 2011;Yang, 2015). We note that the S_1/S_2 junction polymorphism among HKU5 isolates impacts S processing during biogenesis (**Fig. S5.3**), which could modulate the relative potency of each of these inhibitors. We also found that the OC43 HR2-derived EK1C4 lipopeptide fusion inhibitor(Xia, 2019; Xia, 2020), which interferes with S fusogenic conformational changes, and remdesivir(de Wit, 2020; Sheahan, 2020), the inhibitor of coronavirus RNA-dependent RNA polymerase(Koki, 2021), inhibited HKU5-1 propagation at a concentration of 1 μM and 10 μM, respectively **(Fig. 5.6C)**. These findings indicate that inhibitors targeting host proteases mediating S proteolytic activation, such as TMPRSS2, or conserved coronavirus target sites, such

as the S fusion machinery or viral polymerase, are promising candidates for pandemic preparedness against HKU5 and other coronaviruses.

5.8 Chapter Discussion

Phylogenetic classification of merbecovirus RBDs underscores the exceptional diversity of these pathogens harboring RBDs sharing as little as 30% amino acid sequence identity which cluster in at least six distinct clades. The identification of DPP4 as the host receptor for members of the MERS-CoV clade upon its emergence in 2012(Raj, 2013) and subsequently for members of the HKU4 clade(Chen, 2023; Wang, 2014; Yang, 2014), led to the assumption that DPP4 was a universal receptor for merbecoviruses. The discovery of bat-borne merbecoviruses in Africa utilizing a broad spectrum of ACE2 orthologs(Xiong, 2022; Ma, 2023) indicated that receptor diversity was greater than previously appreciated for these viruses. The discovery that viruses from the HKU5 clade and the MOW15-22 clade (cf. companion manuscript(Ma, 2024)) use the ACE2 receptor profoundly alters our understanding of merbecovirus evolution and shows that ACE2 usage is a broadly shared property among several clades. Furthermore, the structural specialization of the RBM within each clade along with the entirely distinct ACE2 binding modes lead to recognition of receptor regions located up to 50Å apart. These findings indicate that ACE2 utilization has been independently acquired at least three times among merbecoviruses (NeoCoV/PDF-2180, MOW15-22/PnNL2018B, HKU5) and at least twice among other coronaviruses (SARS-CoV-1/SARS-CoV-2(McCallum, 2022; Lan, 2020; Li, 2005) and NL63(Wu, 2009), **Fig. 5.7)**. SARS-CoV-1/SARS-CoV-2, NL63 and HKU5 recognize distinct ACE2 regions but with sizable overlap relative to the other two binding modes identified. The regions targeted by these different viruses are likely determined based on accessibility (as receptor

dimerization and N-linked glycosylation can modulate binding) and geometrical constraints(Liu, 2024). We note that utilization of amino-peptidase N among a large set of coronaviruses also results from at least three convergent evolution events for TGEV/CCoV-HuPn-2018/FCoV-23(Tortorici, 2022;Tortorici, 2024; Reguera, 2012), PDCoV(Ji, 2022) and 229E(Wong, 2017). Recurrent coronavirus convergence to use a small number of host transmembrane protease receptors, for which the proteolytic activity does not participate in entry, might result from specific size and geometrical requirements for infection or of the clustering of receptors in specialized membrane patches(Liu, 2024).

Given that the RBD is the main target of coronavirus neutralizing antibodies(Piccoli, 2020;Bowen, 2022; Addetia, 2024; Greaney, 2021), the rare and weak cross-neutralization of HKU5 observed with MERS-CoV infection-elicited polyclonal plasma antibodies concurs with the extensive genetic divergence among merbecovirus RBD clades. These findings are reminiscent of the weak cross-clade neutralization observed for sarbecoviruses(Lee, 2023; Walls, 2021) which harbor RBDs much more similar to each other than merbecovirus RBDs. As a result, the development of vaccines with pan-merbecovirus neutralizing activity may prove to be exceptionally challenging(Addetia, 2024; Chao, 2024). However, our data suggest that rare monoclonal antibodies with broadly neutralizing merbecovirus activity were present in the plasma of a few subjects in the cohort analyzed, explaining the detectable (albeit weak) cross-neutralization and motivating future isolation and characterization of these antibodies. We hypothesize that these rare antibodies may target conserved antigenic sites in the RBD core, the NTD or

fusion machinery epitopes(Low, 2022; Sauer, 2021; Pinto, 2021; Zhou, 2023; Song, 2021; Tortorici, 2021; Starr, 2021; Wang, 2022).

Although we could not detect hACE2-mediated wildtype HKU5 RBD binding or S VSV pseudovirus entry, exogenous trypsin addition enabled weak propagation of pcVSV-HKU5-1 and authentic HKU5-1 in Caco-2 human cells, in line with a previous study(Menachery, 2020). These results further concur with the R.nor ACE2-mediated entry of HKU5-19s pseudovirus in the absence of detectable RBD binding and with observations made for MOW15-22 and PnNL2018B in our companion manuscript(Ma, 2024). HKU5-19s entry likely resulted from avidity due to multivalent presentation of S trimers on pseudoviruses along with the use of trypsin to enhance pseudovirus entry(Xiong, 2022; Yan, 2021; Ma, 2024). Given the extensive RBM polymorphism observed in the HKU5 clade, possibly resulting from immune pressure in the reservoir host or modulating receptor species tropism, and our limited understanding of the genetic diversity of these viruses (due to sparse sampling of merbecoviruses in bats), HKU5 may acquire more efficient hACE2 usage and constitute a zoonotic threat due to the lack of merbecovirus immunity in the human population. This is further supported by the recent detection of HKU5-like viruses in two deceased farmed minks, which had pneumonia and high viral titers in their lungs(Zhao, 2024), and our identification of a single HKU5 RBD amino acid mutation (Y560K) enabling more effective use of hACE2. These results provide proof-of-principle for using machine-learning to expedite the identification of zoonotic viruses poised for emergence and to identify countermeasures to be prepared for possible spillovers. Furthermore, our data provide a blueprint for monitoring HKU5 residue substitutions at the interface with the molecular determinants

of ACE2 species tropism delineated here to detect possible future HKU5 adaptations to the human receptor. The identification of several inhibitors of HKU5-1 propagation in human cells, including camostat (a clinical inhibitor of host serine proteases) and remdesivir (an FDA-approved inhibitor of coronavirus RNA-dependent RNA polymerase), suggests that these compounds could be used in case of possible future outbreaks.

5.9 Methods

Cells

Cell lines used in this study were HEK293T (Female cell line, a derivative of HEK293T, ATCC, CRL-11268), HEK293T (Female cell line, CRL-3216), Caco-2 (Male cell line, HTB-37), BHK21 (Unsexed, CCL-10) and I1-Hybridoma (CRL-2700). HEK293T(CRL-11268, and CRL-3216), BHK21(CCL-10) and Caco-2(HTB-37) cells were cultured in 10% FBS (Fisher Scientific-Cytiva), 1% penicillin-streptomycin (Thermo Fisher Scientific) DMEM at 37℃, 5% CO_2. I1-Hybridoma cell line producing a neutralizing antibody targeting the VSV glycoprotein (VSV-G) was cultured in Minimum Essential Medium (MEM) with Earles's balanced salts, 2.0 mM of L-glutamine (Gibico), and 10% FBS. HEK293T or Caco-2 stable cell lines overexpressing various receptors were generated using lentivirus transduction and selected and maintained in the growth medium with puromycin (1 μg /ml).

Construct design

The wildtype P.abr (Pipistrellus abramus) ACE2 ectodomain encoding residues 18-724 (ACT66266.1) was inserted after the N-terminal signal peptide MPMGSLQPLATLYLLGMLVASVLA and C-terminally fused to a thrombin cleavage site

followed by a SSGGS linker, an avi tag, a GGS short linker and an octa-histidine tag for affinity purification. The wildtype B.tau (Bos Taurus) ACE2 ectodomain with native signal peptide encoding residues 1-739 (Q58DDO) was inserted in the CMVR vector with C-terminal octa-histidine tag and AVI tag for affinity purification. The wildtype N.vis (Neogale vison) ACE2 ectodomain with native signal peptide encoding residues 1-740 (XP_044091952.1) was inserted in the CMVR vector with C-terminal octa-histidine and AVI tag for affinity purification. The dimeric Human ACE2 ectodomain construct encodes residues 19-740 and harbors a N-terminal signal peptide MGILPSPGMPALLSLVSLLSVLLMGCVAQ and a C-terminal octa-histidine tag followed by a short GGSSG linker and an avi tag. The monomeric Human ACE2 ectodomain construct encodes residues 19-615 and harbors a N-terminal signal peptide MGSLQPLATLYLLGMLVASVLA and a C-terminal 10x-histidine tag. The HKU5-19s wildtype (AGP04932.1) RBD encoding residues 390-587 and the Y560K, Y558A, K520G, Y518A, Y464A, Y545G, G519W, A517R receptor-interface mutants, contain the N-terminal signal peptide MGILPSPGMPALLSLVSLLSVLLMGCVAETGT, followed by a thrombin cleavage site, a SSGGS flexible linker, an avi tag, a GGS flexible linker and a C-terminal octa-histidine tag. HKU5-33s wildtype (AGP04943.1) spike RBD encoding residues 390-592 contain the N-terminal signal peptide MGILPSPGMPALLSLVSLLSVLLMGCVAETGT, followed by a thrombin cleavage site, an avi tag, a short GGS flexible linker and a C-terminal octa-histidine tag. HKU5-19s wildtype (AGP04932.1) spike trimer ectodomain encodes residues 22-1296 with a mu-phosphatase N-terminal signal peptide (MGILPSPGMPALLSLVSLLSVLLMGCVAETGT), 2P spike prefusion-stabilizing

mutations[63], and C-terminal Foldon trimerization domain followed by an avi tag, a short GGS flexible linker and an octa-histidine tag. All the above genes were codon-optimized for expression in mammalian cells, synthesized, and inserted in the pcDNA3.1(+) by Genscript. Plasmids expressing wild-type (WT) or mutated bat and non-bat mammalian ACE2 orthologs were constructed by inserting human codon-optimized sequences with/without specific mutations into a lentiviral transfer vector (pLVX-EF1a-Puro, Genewiz) with C-terminus 3×FLAG tags (DYKDHD-G-DYKDHD-I-DYKDDDDK) and single FLAG tag (DYKDDDDK) for non-bat mammalian ACE2 orthologs. For the production of pseudovirus used for entry assays (Fig.3), human codon-optimized spike sequences of HKU5-19s (AGP04932.1), HKU5-19s with Y560K mutation, HKU5-1 (YP_001039962), HKU5-JX2020-Q257 (WCC63324.1), HKU5-GD2017-Q262 (WCC63369.1) and HKU5-JX2020-Q279 (WCC63520.1) fused C-terminal HA tag (YPYDVPDYA) were cloned into the pCAGGS vector with C-terminal truncation (Δ13, residues 1341-1353)(Schwegmann-Weßels, 2009). HKU5-19s S full length (AGP04932.1), HKU5-19s S full length with Y560K mutation, and HKU5-33s S full length (AGP04943.1) were also cloned into the pcDNA3.1(+) with C-terminal truncation (Δ21) without any tag for VSV pseudovirus production that were used for hACE2 and P. abramus ACE2 entry assays (Fig. 4) and sera neutralization assay (Fig. 6A). The DNA fragments for cloning ACE2 chimera or RBD mutants were generated by overlap extension PCR or gene synthesis and verified by commercial DNA sequencing. For the expression of recombinant CoVs RBD-hFc fusion proteins, plasmids were constructed by inserting coding sequences of HKU5-19s RBD (residues 386-587), HKU5-1 RBD (residues 385-586), HKU5-JX2020-Q279 RBD (residues 385-586), HKU5-JX2020-Q257

RBD (residues 385-592), HKU5-GD2017-Q262 RBD (residues 385-590) HKU5-BY140535 RBD (residues 386-592) or their mutants into the pCAGGS vector containing an N-terminal CD5 secretion signal peptide (MPMGSLQPLATLYLLGMLVASVLA) and C-terminal hFc-twin-strep-3xFLAG tag (WSHPQFEKGGGSGGGSGGSAWSHPQFEKGGGRSDYKDHDGDYKDHDIDYKDDDDK) for purification and detection. Heavy chain and light chain genes encoding 76E1, S2P6 and B6 IgG were inserted into pcDNA3.1(+) plasmid, respectively or pCAGGS plasmid with N-terminal fused CD5 signal peptide.

Recombinant protein production

The P. abr, B. tau, N. vis, dimeric Human, and monomeric Human ACE2 ectodomains, the wildtype HKU5-19s, HKU5-19s Y560K, HKU5-1, and HKU5-33s RBDs along with HKU5-19s interface point mutants were expressed in Expi293F cells (Thermo) maintained at 37°C and 8% CO_2. Cells were transfected using Expifectamine293 (Thermo) following the manufacturer's protocol. Four to five days post-transfection, Expi293 cell supernatant was clarified by centrifugation at 4,121g for 30 minutes. Supernatants were supplemented to a final concentration of 100 mM Tris-HCl pH 8.0, and 300 mM NaCl before binding to Ni excel resin (Cytiva) previously equilibrated in 100 mM Tris pH 8.0 or 100 mM Tris pH 7.4 and 300 mM NaCl. Nickel resin was washed with 100mM Tris pH 8.0 or 100mM Tris pH 7.4, 300 mM NaCl, and 10-20mM imidazole prior to elution with 100mM Tris-HCl pH 8.0 or 100mM Tris-HCl pH 7.4, 300 mM NaCl, and 300mM imidazole. For monomeric human ACE2 purification, supernatant was 0.22μm vacuum filtered and supplemented to a final concentration of 300mM NaCl, 25mM phosphate pH 8.0, and 0.5mM PMSF prior to binding to a Histrap Excel column

(Cytiva) for 24 hours. Affinity column was then washed with buffer containing 300mM NaCl, 25mM phosphate pH 8.0, and 10mM imidazole and then protein was eluted with buffer containing 300mM NaCl, 25mM phosphate pH 8.0, and 300mM imidazole. ACE2 ectodomains were concentrated using centrifugal filter devices with a MWCO of 30kDa and eluted on a Superdex200 increase 10/300 size-exclusion column (Cytiva) equilibrated in 50mM Tris pH 7.4 and 150 mM NaCl or 100mM NaCl, 20mM phosphate pH 8.0 for only monomeric human ACE2 ectodomain. Monodisperse fractions were pooled and used immediately or flash frozen and stored at -80℃ until use. For structural characterization, purified HKU5-19s RBD was concentrated using a centrifugal filter device with a MWCO of 10kDa and eluted on a Superdex200 increase 10/300 size-exclusion column (Cytiva) equilibrated in 50mM Tris pH 7.4 and 150 mM NaCl and fractions containing monomeric RBD were used immediate or flash frozen and stored at -80C until use. For BLI, purified RBDs were biotinylated using the BirA biotin-protein ligase reaction kit (Avidity) after buffer exchanging into 50 mM Tris-HCl pH 7.4 and 150 mM NaCl using a centrifugal filter device with a MWCO of 10kDa. Affinity purification as above-described was carried out a second time to get rid of BirA and free biotin. Purified biotinylated HKU5-19s interface point mutant RBDs were buffer exchanged into 50 mM Tris pH 7.4 and 150 mM NaCl and used immediately or flash frozen. Purified biotinylated HKU5-19s, HKU5-19s Y560K, HKU5-1, and HKU5-33s RBD were concentrated using a centrifugal filter device with a MWCO of 10kDa and run on a Superdex200 increase 10/300 size-exclusion column (Cytiva) equilibrated in 50mM Tris pH 7.4 and 150 mM NaCl and fractions containing monomeric RBD were flash frozen and stored at -80℃ until use.

HEK293T cells were also transfected with HKU5-19s WT or mutant RBD plasmids using GeneTwin reagent (Biomed, TG101-01). Subsequently, the culture medium of the transfected cells was replenished with the SMM 293-TII Expression Medium (Sino Biological, M293TII) 4-6 hours post-transfection, and the protein-containing supernatant was collected every two days for 2-3 batches. Antibodies and recombinant RBD-hFc proteins were purified using Pierce Protein A/G Plus Agarose (Thermo Scientific, 20424). In general, Fc-containing proteins were enriched by the agarose, washed with wash buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA), eluted using 100 mM Glycine pH 3.0, and immediately neutralized with 1/10 volume of 1M Tris-HCl, pH 8.0 (15568025, Thermo Scientific). Proteins with twin-strep tag were purified using Strep-Tactin XT 4Flow high-capacity resin (IBA, 2-5030-002), washed by wash buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA), and then eluted with buffer BXT (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM biotin). All eluted proteins were concentrated using Ultrafiltration tubes, buffer-changed to PBS, and stored at -80℃. Protein concentrations were determined by the Omni-Easy Instant BCA Protein Assay Kit (Epizyme, ZJ102) or BLI assays using Octet RED96 instrument (Molecular Devices).

To express the 76E1 monoclonal antibody, Expi293F cells were transiently transfected with plasmids encoding for the IgG heavy and light chains (with an equal ratio of each plasmid) using Expifectamine 293 following the manufacturer's protocol. Four days after transfection, Expi293F cell supernatant was clarified by centrifugation at 3,500xg for 30 minutes and diluted with 20mM phosphate pH 8.0 and 0.1 mM PMSF. Supernatant was passed through a Protein A affinity column (Cytiva) twice before being

washed with 20 mM phosphate pH 8.0. 76E1 was eluted with 100 mM citric acid pH 3.0 into tubes containing 1M Tris-HCl pH 9.0 for immediate neutralization of the low pH used for elution. Eluted 76E1 was buffer-exchanged into 20 mM phosphate buffer pH 8.0 and 100 mM NaCl and the purity was confirmed by SDS-PAGE.

To express the B6 monoclonal antibody, ExpiCHO cells were transiently transfected with plasmids encoding for the IgG heavy and light chains (with an equal ratio of each plasmid) following the manufacturer's protocol using Expifectamine CHO. Eight days after transfection, ExpiCHO cell supernatant was clarified by centrifugation at 3,500xg for 30 minutes then filtered via vacuum through a 0.2 μm aPES membrane, supplemented with a buffer containing 500 mM sodium phosphate pH 8.0 to a final concentration of 20 mM and PMSF added to a final concentration of 0.1 mM. Supernatant was passed over a Protein A affinity column (Cytiva) twice before being washed with 20 mM sodium phosphate pH 8.0. B6 was eluted with 100 mM citric acid pH 3.0 into tubes containing 1M Tris-HCl pH 9.0 for immediate neutralization to a final pH of 8.0. The eluted B6 antibody was buffer-exchanged into 25 mM phosphate pH 8.0, 100 mM NaCl and purity was confirmed by SDS-PAGE in both reducing and non-reducing conditions.

Cell-cell fusion assays

Dual-split proteins (DSPs) were utilized in Caco-2 cells stably expressing the indicated ACE2 receptors. For the HKU5-19s S glycoprotein-receptor interaction-mediated membrane fusion between cells, group A cells expressed S glycoprotein and rLucN(1-155)-sfGFP1-7(1-157), while group B cells expressed S glycoprotein and sfGFP8-11(158-231)-rLuc(156-311). 24 hours post-transfection, group A and group B cells were

trypsinized, mixed, and seeded into a 96-well plate at 8×10^4 cells per well. Then the cells were washed once with DMEM and refreshed by new DMEM containing 2% FBS with or without indicated concentrations of TPCK-treated trypsin (Sigma-Aldrich, T8802) at room temperature at 36 h post-transfection. 24 hours later, nuclei were stained with Hoechst 33342 (1:5,000 dilution in HBSS) for 30 min at 37 °C, and images of syncytia formation with green fluorescence were subsequently captured using a fluorescence microscope (MI52-N; Mshot).

RBD-hFc live-cell binding assay

RBD-hFc live-cell binding assays were conducted following a previously described protocol. The coronavirus RBD-hFc recombinant proteins were diluted in DMEM at 4 μg/mL and incubated with HEK293T cells transiently expressing different ACE2 for 30 minutes at 37℃ at 36 hours post-transfection. Subsequently, cells were washed once with Hanks' Balanced Salt Solution (HBSS) and incubated with 1 μg/mL of Alexa Fluor 488-conjugated goat anti-human IgG (Thermo Fisher Scientific; A11013) and Hoechst 33342 (1:10,000 dilution in HBSS) diluted in HBSS/1% BSA for 1 hour at 37℃. After another round of washing with HBSS, the cells were incubated with fresh HBSS. The images were captured using a fluorescence microscope (MI52-N). The relative fluorescence intensities (RFUs) of the stained cells were determined by a Varioskan LUX Multi-well Luminometer (Thermo Scientific).

Pseudovirus production and entry assays

VSV-dG-based pseudovirus (PSV) carrying trans-complementary S glycoproteins from various coronaviruses were produced following a modified protocol as previously described. Briefly, HEK293T cells were transfected with coronaviruses S glycoproteins

expression plasmids. At 24 hours post-transfection, cells were infected with 1.5×10^6 TCID50 VSV-G glycoprotein- deficient VSV expressing GFP and firefly luciferase (VSV-dG-fLuc-GFP, constructed and produced in-house) diluted in DMEM with 8 μg/mL polybrene and for 4-6 hours at 37 ℃. After three PBS washes, the culture medium was replenished with fresh DMEM or SMM 293-TII Expression Medium (Sino Biological, M293TII), along with the presence of the neutralizing antibody (from I1-mouse hybridoma) targeting the VSV-G to eliminate the background produced by the remaining VSV-dG-fLuc-GFP. Twenty-four hours later, the pseudovirus containing supernatant was clarified through centrifugation at 12,000 rpm for 5 minutes at 4℃, aliquoted, and stored at -80℃. The $TCID_{50}$ of the PSV was calculated using the Reed-Muench method(Reed, 1938).

HEK293T or Caco-2 cells transiently or stably expressing the indicated ACE2 orthologues were infected by the single-round pseudovirus. Approximately 3×10^4 trypsinized cells were incubated with pseudovirus (2×10^5 TCID50/100 μL) in a 96-well plate to facilitate attachment and viral entry simultaneously. Before inoculation, pseudoviruses were typically treated with 100 μg/mL TPCK-trypsin (Sigma-Aldrich, T8802). Specifically, pseudoviruses produced in serum-free SMM 293-TII Expression Medium were incubated with TPCK-treated trypsin for 10 minutes at room temperature, and the proteolytic activity was neutralized by FBS in the culture medium. I1 Intracellular luciferase activity (Relative light units, RLU) was measured using the Bright-Glo Luciferase Assay Kit (Promega, E2620) and detected with a GloMax 20/20 Luminometer (Promega) or Varioskan LUX Multi-well Luminometer (Thermo Fisher Scientific) at 18 hours post-infection.

For pseudovirus loss-of-function assay, the indicated VSV-pseudotyped HKU5-19s mutants were produced using HEK293T cells transfected with S expressing plasmids using GeneTwin (Biomed, TG101-01) in DMEM. After 6 hours post-transfection, cells were supplemented with DMEM containing 10% FBS. 24 hours post-transfection, cells were infected with 1.5×10 6 TCID50 VSV-G glycoprotein-deficient VSV expressing GFP and firefly luciferase (VSV-dG-fLuc-GFP, constructed and produced in-house) with 8 μg/mL polybrene for 5h and than I1-mouse hybridoma supernatant (ATCC CRL-2700) was added to neutralize the VSV-dG-fLuc-GFP for lower background. After 24h post-infection, pseudoviruses were collected and cell debris was removed by centrifugation at 12,000g for 5 min, aliquoted, and stored at -80℃.

Pseudovirus production for entry and neutralization assays

HKU5-19s, HKU5-19s Y560K, HKU5-33s S VSV pseudoviruses were produced using HEK293T cells seeded on BioCoat Cell Culture Dish: poly-D-Lysine 100 mm (Corning). Cells were transfected with respective S constructs using Lipofectamine 2000 (Life Technologies) in Opti-MEM transfection medium. After 5h of incubation at 37 °C with 5% CO_2, cells were supplemented with DMEM containing 10% of FBS. On the next day, cells were infected with VSV (G*ΔG-luciferase) for 2h, followed by five time wash with DMEM medium before addition of anti-VSV G antibody (I1-mouse hybridoma supernatant diluted 1:40, ATCC CRL-2700) and medium. After 18-24 h of incubation at 37 °C with 5% CO_2, pseudoviruses were collected and cell debris was removed by centrifugation at 3,000g for 10 min. Pseudoviruses were further filtered using a 0.45 μm syringe filter and concentrated 10x for neutralization assay and 25x for entry assay prior to storage at -80°C.

For HKU5-19s, HKU5-19s Y560K, HKU5-33s S VSV pseudovirus entry assays (Fig. 4), HEK293T cells were transfected with full-length hACE2 and P. abramus ACE2 using Lipofectamine 2000 (Life Technologies) in Opti-MEM transfection medium. After 5h of incubation at 37 °C with 5% CO_2, cells were seeded into 96-well plates precoated with poly-L-lysine solution (Sigma, P4707). The following day, a half-area 96-well plate (Greiner) was prepared with each VSV (22μL in total per well). The pseudovirus was normalized by the amount of VSV (HKU5 VSVs) and incorporated spikes (HKU5-19s VSV) using western blot. An equal volume of DMEM was added to each well for the non-trypsin treated group. For the trypsin-treated group, pseudoviruses were incubated with 100 μg/mL trypsin for 15 min at room temperature before addition of trypsin soybean inhibitor to stop the reaction. 40 μL from each well of the half-area 96-well plate containing pseudovirus were transferred to the 96-well plate seeded with receptor-transfected cells and incubated for 1h. After 1h, an additional 40 μL of DMEM supplemented with 20% FBS and 2% PenStrep was added to the cells. After 18–20h of incubation at 37 °C with 5% CO_2, 40 μL of One-Glo-EX substrate (Promega) was added to each well and incubated on a plate shaker in the dark for 5 min before reading the luciferase signal using a BioTek Neo2 plate reader. hACE2 and P. abramus ACE2 groups were each autogained and read at separate gains. RLUs were plotted in Prism (GraphPad).

For plasma and monoclonal antibody neutralization assay (Fig. 6 and Fig. S7, HEK293T cells were transfected with full-length P. abramus ACE2 using Lipofectamine 2000 (Life Technologies) in Opti-MEM transfection medium. After 5h of incubation at 37 °C with 5% CO_2, cells were seeded into 96-well plates precoated with poly-L-lysine solution

(Sigma, P4707). The following day, a half-area 96-well plate (Greiner) was prepared with 3-fold serial dilutions of each plasma (starting dilutions of 1:10, 22µL in total per well) or of the B6 and 76E1 monoclonal antibodies (starting dilutions of 500 µg/mL). An equal volume of DMEM with diluted pseudoviruses was added to each well. The pseudovirus was diluted to 1:15 to reach a target entry of ~10^6 relative luciferase units (RLUs). The pseudoviruses were not treated with trypsin in these assays. The mixture was incubated at room temperature for 45-60 minutes. 40 µL from each well of the half-area 96-well plate containing sera and pseudovirus were transferred to the 96-well plate seeded with cells and incubated for 1h. After 1h, an additional 40 µL of DMEM supplemented with 20% FBS and 2% PenStrep was added to the cells. After 18–20h of incubation at 37 °C with 5% CO_2, 40 µL of One-Glo-EX substrate (Promega) was added to each well and incubated on a plate shaker in the dark for 5 min before reading the luciferase signal using a BioTek Neo2 plate reader. RLUs were plotted and normalized in Prism (GraphPad): 100% neutralization being cells in the absence of pseudovirus and 0% neutralization being pseudovirus entry into cells without plasma. Prism (GraphPad) nonlinear regression with "log[inhibitor] versus normalized response with a variable slope" was used to fit the curve. The mean percent pseudovirus entries at 1:10 sera dilution point are shown from the three biological replicates per sample-pseudovirus pair. Demographics of the sera samples were previously described(Addetia, 2024).

Western blot for VSV entry assay

Western blot assays were conducted to examine the S glycoprotein proteolytic processing and incorporation efficiency (Fig.3), the PSV-containing supernatant was

concentrated using a 30% sucrose cushion (30% sucrose, 15 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA) at 20,000×g for 1 hour at 4℃. The concentrated virus pellet was resuspended in 1×SDS loading buffer and incubated at 95℃ for 30 minutes, followed by western blot detecting the S glycoproteins by S2-targeting S2P6 or C-terminal HA tags and with the VSV-M [23H12] serving as a loading control. The blots were then washed three times by PBST and then visualized using an Omni-ECL Femto Light Chemiluminescence Kit (EpiZyme, SQ201) by a ChemiDoc MP Imaging System (Bio-Rad).

For normalized VSV entry assay (Fig. 4), pseudoviruses were blotted first with either 76E1 IgG or anti-VSV-M [23H12] antibody (Kerafast), followed by secondary antibody IRDye® 680RD Goat anti-Human IgG Secondary Antibody (LICORbio) and IRDye® 680RD Donkey anti-Mouse IgG Secondary Antibody (LICORbio), respectively. Intensity of the bands were quantified using ImageJ(Schneider, 2012). iBright™ Prestained Protein Ladder (Invitrogen) was used for protein sizing.

Immunofluorescence assay

The expression levels of ACE2 orthologs with C-terminal fused FLAG tags were determined through immunofluorescence assays. Specifically, the transfected cells were fixed and permeabilized by incubation with 100% methanol for 10 minutes at room temperature. Subsequently, the cells were incubated with a mouse antibody M2 (Sigma-Aldrich, F1804) diluted in PBS/1% BSA for one hour at 37℃. After once PBS wash, the cells were incubated with Alexa Fluor 594-conjugated goat anti-mouse IgG (Thermo Fisher Scientific, A32742) secondary antibody diluted in 1% BSA/PBS for one hour at 37℃. The images were captured and merged with a fluorescence microscope (Mshot,

MI52-N) after the nucleus was stained blue with Hoechst 33342 reagent (1:5,000 dilution in PBS).

Cryo-electron microscopy data collection, processing and model building

The P.abr ACE2 ectodomain-bound HKU5-19s (non-biotinylated) RBD complex was prepared by mixing at 1:1.2 molar ratio followed by a 1 hour incubation at room temperature. Specimen vitrification followed two methods. 3μL of 4mg/ml complex with 6 mM 3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO) were applied onto freshly glow discharged R 2/2 UltrAuFoil grids(Russo, 2014) prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 5.5 sec blot time at 100% humidity and 22°C. 3μL of 0.2 mg/mL complex without detergent was added to the glow discharged side of R 2/2 UltrAuFoil grids and 1μL was added to the back side before plunging into liquid ethane using a GP2 (Leica) with 6 sec blot time. The data were acquired using an FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K3 direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon (Suloway, 2005) at a nominal magnification of 105,000× with a pixel size of 0.843 Å. The dose rate was adjusted to 9 counts/pixel/s, and each movie was acquired in counting mode fractionated in 100 frames of 40 ms. A total 11,223 micrographs were collected with a defocus range between -0.2 and -3 μm and stage tilt angle of 0˚, 20˚, 25˚ and 30˚. Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking, and extraction were carried out using Warp (Tegunov, 2019). Particles were extracted with a box size of 120 pixels with a pixel size of 1.686Å.

Two rounds of reference-free 2D classification were performed using cryoSPARC (Punajani, 2017) to select well-defined particle images. Initial model generation was carried out using ab-initio reconstruction in cryoSPARC and the resulting maps were used as references for heterogeneous 3D refinement. Particles belonging to classes with the best resolved HKU5 RBD and ACE2 density were selected. To further improve the data, the Topaz model (Bepler, 2020) was trained on Warp-picked particle sets belonging to the best classes after 2D classification and particles picked using Topaz were extracted and subjected to 2D-classification and heterogenous 3D refinements. The two different particle sets from the Warp and Topaz picking strategies were merged and duplicates were removed using a minimum distance cutoff of 95Å. After two rounds of ab-initio reconstruction-heterogeneous refinements, 3D refinement was carried out using non-uniform refinement in cryoSPARC (Punjani, 2020). The dataset was transferred from cryoSPARC to Relion(Zivanov, 2018) using the pyem program package(Asarnow, 2019) and particle images were subjected to the Bayesian polishing procedure implemented in Relion(Zivanov, 2019) during which particles were re-extracted with a box size of 320 pixels and a pixel size of 1.0 Å. To further improve the map quality, ab-initio reconstruction in cryoSPARC was used to classify the data in three bins and the generated models were used as references for heterogeneous 3D refinement. The final 3D refinements were carried out using non-uniform refinement along with per-particle defocus refinement in cryoSPARC to yield the final reconstruction at 3.1 Å resolution comprising 550,684 particles. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-

resolution noise substitution(Rosenthal, 2003; Chen, 2013). Local resolution estimation, filtering, and sharpening were carried out using cryoSPARC. UCSF Chimera(Pettersen, 2004), Coot(Emsley, 2010), AlphaFold3[61], Rosetta(Frenz, 2019; Wang, 2016) and Phenix(Liebschner, 2019), were used to fit, build, and refine the model using the sharpened and unsharpened cryo-EM maps. Validation used Phenix(Liebschner, 2019), Molprobity(Chen, 2010), EMRinger(Barad, 2015) and Privateer(Agirre, 2015).

For the B.tau ACE2 ectodomain-bound HKU5-19s RBD structure, the complex formation was performed by mixing a 1:1.5 molar ratio of the B.tau ACE2 ectodomain (residue 1-739) with the HKU5-19s RBD (residues 390-587) before incubation for 1 hour at room temperature. 3 μL of 5 mg/ml complex with 6 mM CHAPSO were applied onto freshly glow discharged R 2/2 UltrAuFoil grids prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 5.5 sec blot time at 100% humidity and 22°C. 12,185 movies were collected with a defocus range comprised between -0.2 and -3.5 μm. The overall data processing methods were the same as that for the P.abr ACE2 ectodomain-bound HKU5-19s RBD complex structure except that non-uniform refinement used C2 symmetry. To further improve the density of the RBD-ACE2 interface, the particles were symmetry expanded and subjected to focus 3D classification without refining angles and shifts using a soft mask encompassing the RBD using a tau value of 60 in Relion. Particles belonging to classes with the best resolved RBD density were selected and then subjected to local refinement using CryoSPARC. The final dataset contained 1,598,132 asymmetric units used for the final local refinement with a soft mask comprising one ACE2 peptidase domain and the

bound RBD resulting in a 2.4 Å resolution reconstruction. More details are shown in **Fig. S5.3**.

For the prefusion HKU5-19s spike ectodomain trimer structure, 1 μM spike ectodomain (residues 22-1296) trimer was incubated overnight with 100 mM 5-N-acetyl neuraminic acid at 4ºC. Cryo-EM grids were prepared using two separate methods and data were combined during data processing. The first dataset was collected from the UltrAuFoil grids. 3 μL of 1.5 mg/mL spike were loaded onto freshly glow discharged R 2/2 UltrAuFoil grids prior to plunge freezing using a vitrobot MarkIV with a blot force of 0 and 6 sec blot time at 100% humidity and 22°C. The second dataset was collected from a lacey carbon grid with a home-made continuous thin carbon layer. 3 μL of 0.15 mg/mL spike were loaded onto the glow discharged (10s at 20mA) lacey carbon grid followed by plunge freezing using a vitrobot MarkIV. The lacey carbon grid was blotted with a blot force of -1, 3 second blot time, and 10 second wait time before the plunge freeze at 100% humidity and 22 °C. Data were acquired using an FEI Titan Krios transmission electron microscope operated at 300 kV. A total of 10,456 and 5,045 micrographs were collected from UltrAuFoil grids and thin home-made continuous carbon grids, respectively. Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking and extraction were carried out using Warp. Particles were extracted with a box size of 512 pixels with a pixel size of 1.686 Å. The overall data processing methods were the same as that for the P.abr ACE2 ectodomain-bound HKU5-19s RBD complex structure except that non-uniform refinement used C3 symmetry. During Bayesian polishing procedure in Relion, particles were re-extracted with a box size of 608 pixels and a pixel size of 0.843 Å. After polishing, particles were

subjected to heterogeneous refinements in cryoSPARC to select particles belonging to the two different classes with the best-resolved density. NUR with per-particle defocus refinement yielded final two reconstructions of prefusion HKU5-19s spike ectodomain trimer with the structural variations of C and D domains at 2.0 Å resolution enabling atomic model building. More details are shown in **Fig. S5.6**.

Biolayer interferometry

P. abr, B. tau, and N. Vis ACE2 binding kinetics to the HKU5-19s, HKU5-1 and HKU5-33s RBDs was assessed by BLI using an Octet RED96 instrument (Sartorius) and the Octet Data acquisition software. All BLI measurements were performed at 30°C and shaking at 1,000 rpm. Biotinylated HKU5 RBDs were diluted to 10µg/mL in 10x Octet kinetics buffer (Sartorius) and loaded onto hydrated streptavidin (SA) biosensors to 1 nm shift, equilibrated in 10x Octet kinetics buffer for 150 seconds, and dipped into P. abramus ACE2 at 900nM, 300nM, 100nM, 33.3nM and 11.1nM for 900s to observe association. Dissociation was observed by dipping biosensors in a 10x Octet kinetics buffer for 600s. Baseline subtraction was done by subtracting the response from unloaded SA tips dipped into 300nM P. abramus ACE2. In addition, association phases were aligned to 0 seconds and 0 response in Octet Data Analysis HT software. Apparent dissociation constants were determined by plotting the concentration of ACE2 versus the average responses from 890-895s and fitting a one site–specific binding nonlinear fit in GraphPad Prism10. We anticipate that the numerical accuracy of the estimated $K_{D,app}$ values are affected by the combination of the oligomeric nature of the ACE2 dimers used, the equilibrium observed between monomeric and dimeric forms by

mass photometry (**Data S5.4**) and the fitting model used. Sensorgrams were plotted in GraphPad Prism10.

P. abramus ACE2 binding to wildtype and interface mutant HKU5-19s RBDs was assessed by by loading each RBD onto hydrated streptavidin biosensors to a 1nm shift, equilibrating in 10x Octet kinetics buffer for 60 seconds, before dipping into P. abramus ACE2 at 100nM for 300s to observe association. Dissociation was observed by dipping biosensors in 10x Octet kinetics buffer for 300s. Baseline subtraction was done by subtracting the response from unloaded SA tips dipped into 100nM P. abramus ACE2. In addition, association phases were aligned to 0 seconds and 0 response in Octet Data Analysis HT software. Sensorgrams were plotted in GraphPad Prism10.

Human ACE2 binding to wildtype HKU5-19s, HKU5-1, and HKU5-19s Y560K RBDs was assessed by loading each RBD diluted in 10µg/mL modified kinetics buffer (1x Cold Springer PBS, 1% w/v BSA, 0.06% v/v Tween 20) to a 1nm shift onto hydrated streptavidin biosensors followed by equilibration in modified kinetics buffer for 150s. Binding was carried out by dipping the biosensors in 10 µM hACE2 in modified kinetics buffer for 300s to observe association, and then dipping into modified kinetics buffer for 300s to observe dissociation. Baseline correction used the response of a tip loaded with HKU5-19s Y560K RBD and then dipped in modified kinetics buffer for the rest of the assay which was subtracted from all the data. Sensorgram traces were aligned to the start of the association and an interstep correction was made by aligning to 0.5 seconds into the dissociation step. Sensorgrams were plotted in Graph Pad Prism 10.

Mass Photometry

Mass photometry was performed on a Refeyn TwoMP system (Refeyn Ltd). Microscope cover slides were rinsed with isopropanol and Milli-Q water then dried under nitrogen flow. Sample chambers were assembled by placing microscope slides on top of clean silicon gaskets, instrument lens was coated with immersol and assembled sample chambers were placed on the MP sample stage. Samples were added to the sample chamber and the instrument was focused immediately prior to each data acquisition. Data was acquired over the course of 1 minute. Mass calibration was conducted using in-house protein standards and events were exported in DiscoverMP. Representative histograms were generated using a python script. Gaussian mixture models with 2 components and full covariance generated from the Sci Kit Learn Library were overlaid on histogram. In addition, the mean and weight for each gaussian component was determined. ACE2 oligomeric state was determined by identifying gaussian components with means between 50-150kDa as monomeric and 150-300kDa as dimeric. The portion of ACE2 that was observed as monomeric and dimeric was reported as the weight of the gaussian component corresponding to that oligomeric state as a percentage as previously described (Dadonaite, 2024).

Authentic HKU5-1 infection assay

In vitro experiments related to authentic HKU5-1 infection were authorized by the Biosafety Committee of the State Key Laboratory of Virology, Wuhan University. All experiments involving HKU5-1 authentic viruses rescue and infection assays were conducted in the certified negative-pressure Biosafety Level 2 laboratory at Wuhan University in accordance with standard operating procedures (SOPs) in a BSL-2

laboratory. The HKU5-1 authentic viruses were rescued by transfection of a full-length cDNA clone[47] (LMH03f isolate, GenBank: NC_009020). For transfection, 4 μg of plasmids were transfected into Caco-2/P.abr ACE2 cells (1×10^6 cells) using GeneTwin (Biomed, TG101). Progeny viruses collected from the supernatant at 72 hours post-transfection (P0) were utilized to generate stocks for subsequent analyses. For HKU5-1(isolate LMH03F) infection-related experiments, Caco-2 cells with or without human or P.abr ACE2 stable expression were initially seeded in 96-well plates and washed with DMEM before inoculation, either in the presence or absence of indicated concentration of trypsin. Following a one-hour incubation of the indicated MOI of HKU5-1 at 37℃, the cells were washed with DMEM and further incubated for the indicated hours at 37℃. For antibodies and EK1C4 neutralization assay, HKU5-1 was preincubated with the antibodies or peptides for 1 hour at 37℃ before adding to the cells. For other small molecule inhibitors, HKU5-1 was mixed with the inhibitors and immediately incubated with the cells. After 1-hour inoculation, cells were further incubated with culture medium containing indicated concentrations of inhibitors. For immunofluorescence assays, cells were washed once by DMEM and then fixed with 4% paraformaldehyde (PFA) for 20 minutes at room temperature, permeabilized by 0.1% Triton X-100 at 25℃ for 15 minutes, and blocked by 1% BSA at 37℃ for 30 minutes at indicated time points. The expression of HKU5 N proteins was detected by rabbit anti-HKU5 N protein serum (diluted at 1:4000, from Zheng-Li Shi's lab, Wuhan Institute of Virology), followed by DL594-conjugated goat anti-rabbit IgG (Thermo, 1:1000) staining. HKU5-1 titers were determined in Caco-2/P.abr ACE2 cells infected with serially-diluted inocula and calculated using the Reed-Muench method.

Merbecovirus RBD and mammalian ACE2 phylogenetic analysis

A dataset of merbecovirus S glycoproteins was gathered based on literature search(Scott, 2022; Ye, 2020; He, 2020; Hsueh, 2021; Lednicky, 2021) followed by additional BLAST searches of NCBI databases. Identical S amino acid sequences were pared via CD-HIT(Fu, 2012) and aligned via the structure-guided approach of MAFFT-DASH(2019), which utilized the PDB codes 3JCL, 5I08, 5KWB, 5X59, 6B3O, 6OHW, 6PZ8, 6Q04, and 6VSJ during alignment. RBD sequences (MERS-CoV EMC-2012 reference numbering Q377-A591, Genbank Accession JX869059) were parsed from the alignment and further pared via CD-HIT to eliminate redundant sequences with 100% RBD amino acid sequence identity or to remove sequences with ambiguous base calls, leaving 71 unique RBD sequences. The full set of sequences, along with accession numbers, are available from GitHub: https://github.com/tstarrlab/MERSr_phylo/blob/main/unaligned_sequences/GenBankAccessions.xlsx. The RBD sequence alignment was determined to be free of detectable recombination via the likelihood-based model selection program GARD(Kosakovsky Pond, 2006). The final RBD sequence set was aligned again via MAFFT-DASH (which utilized PDBs 4KQZ, 4L3N, 4QZV, 4XAK, 5XGR, 6C6Z, and 6PZ8) and combined with MHV (ACN89763), OC43 (KX344031) and HKU1 (KF686346) RBD sequences as embecovirus outgroups, with final RBD alignment available from GitHub: https://github.com/tstarrlab/MERSr_phylo/blob/main/merbeco_alignments/merbeco_RBD_aligned_v2.fasta. The maximum likelihood phylogenetic relationship among RBD sequences was inferred using RAxML(Stamatakis, 2014) with the WAG+G evolutionary model, and visualized based on rooting on the embecovirus outgroups. The maximum

likelihood tree originally placed a poorly supported paraphyletic relationship for the HKU25-related RBD clade (Fig. 1A). This paraphyletic relationship is unstable (i.e., has fluctuated as we iterated on the phylogeny when adding newly described taxa); disagrees with the most parsimonious placement of an indel (deletion of equivalent MERS-CoV residues 543-552 that is specific to the HKU25-related RBDs); and disagrees with monophyletic HKU25-related clade grouping from phylogenies based on whole genome and full S sequences(Tolentino, 2024; Tan, 2023) which are more confident due to longer alignment length. For our final phylogeny, we therefore enforced a constraint so that RBDs from HKU25-related viruses form a monophyletic clade. The full pipeline and intermediate analysis files are available from GitHub: https://github.com/tstarrlab/MERSr_phylo. Sequence alignments of different ACE2 orthologs or HKU5 S_1/S_2 regions were performed with the MUSCLE(Edgar, 2004) algorithm in MEGA-X(Kumar, 2018) (version 10.1.8). The logo plot of HKU5 S_1/S_2 sequences was generated by the Geneious Prime software (Kearse, 2012). Phylogenetic trees of ACE2 were produced using the maximal likelihood method with Q.plant+I+R3 model in IQ-TREE2(Minh, 2020) (1000 Bootstraps) and rendered with iTOL (Letunic, 2024)(v6).

Statistical analysis

Most experiments were conducted 2–3 times with 3 biological repeats unless otherwise specified. Representative results were shown. Data were presented by means ± SD as indicated in the figure legends. Unpaired two-tailed t-tests were conducted for all statistical analyses using GraphPad Prism 8 unless specified otherwise. $P < 0.05$ was considered significant. *$P < 0.05$,**$P < 0.01$, ***$P < 0.005$, and ****$P < 0.001$.

5.10 Figures & Tables

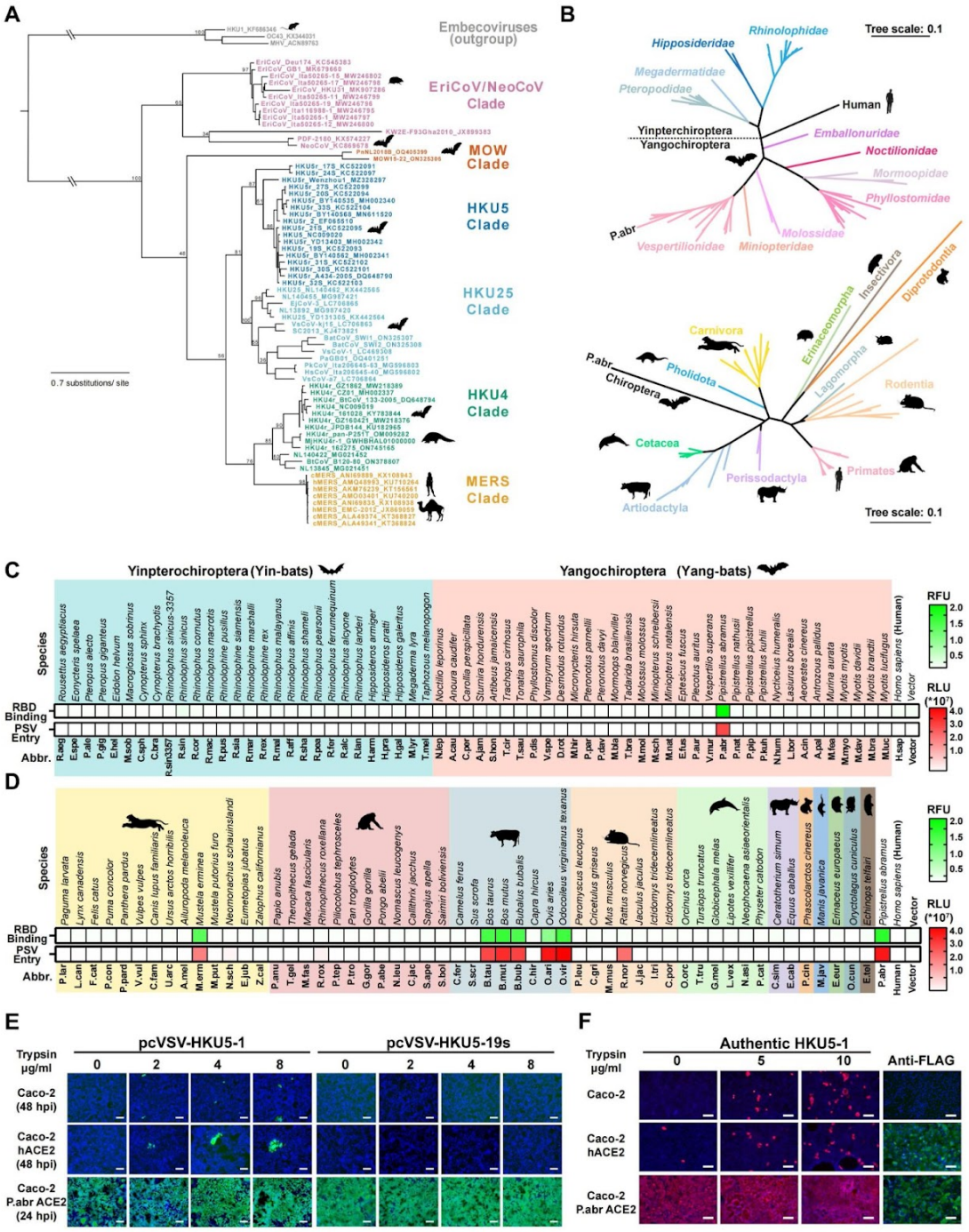


Figure 5.1: Identification of several mammalian ACE2s as functional HKU5 entry receptors.

(A) Merbecovirus RBD phylogenetic tree based on amino acid sequences defining 6 clades. c/hMERS: camel/human MERS-CoV isolates. Each merbecovirus is listed along with its GenBank ID. The animal symbols represent the hosts in which viruses have been detected. **(B)** Phylogenetic trees of bat (top) or non-bat (bottom) mammalian

ACE2 orthologs based on amino acid sequences, with genera and orders indicated for the bat and non-bat mammalian species, respectively. **(C-D)** Binding of the HKU5-19s RBD-hFc to and entry of HKU5-19s S VSV pseudovirus (pretreated with 100 μg/ml TPCK-treated trypsin) into HEK293T cells transiently transfected with the indicated bat (C) or non-bat (D) mammalian ACE2 orthologs. Abbr: abbreviations used for species names. Mean values are shown in C and D with n=3 biological replicates. **(E-F)** Propagation of pcVSV-HKU5-1 (E, left), pcVSV-HKU5-19s (E, right), and authentic HKU5-1 (F) in wildtype human Caco-2 cells (endogenously expressing hACE2) or Caco-2 cells with stable expression of either hACE2 or P.abr ACE2. The propagation of pcVSV-HKU5 recombinant viruses was detected by the expression of the GFP reporter gene at the indicated hours post-infection (hpi) whereas authentic HKU5-1 was detected by immunofluorescence with an anti-HKU5 N antibody at 24 hpi. The trypsin concentration used is indicated. ACE2 expression was confirmed by immunofluorescence using C-terminally-fused FLAG tags. Scale bars: 200 μm. See also **Fig. S5.1-S5.3**.

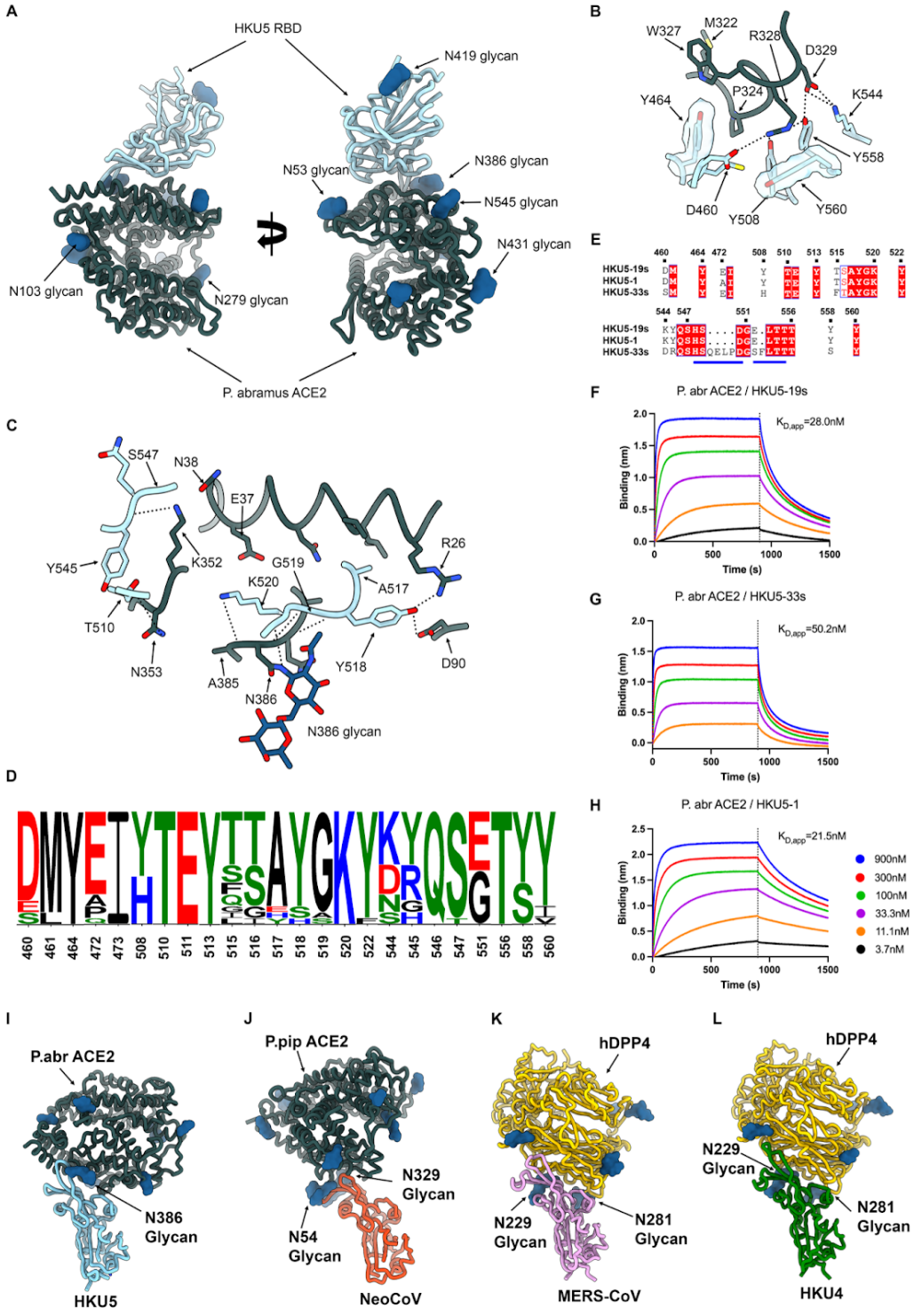


Figure 5.2: Molecular basis of HKU5 recognition of the P. abramus ACE2 receptor.

(A) Ribbon diagrams in two orthogonal orientations of the cryo-EM structure of the HKU5-19s RBD (light blue) bound to the P. abramus ACE2 peptidase domain (green). **(B-C)** Zoomed-in views of the interface highlighting key interactions between the HKU5-19s RBD and P. abramus ACE2. Selected polar interactions are shown as black dotted

lines. **(D)** Polymorphism of ACE2-interacting residues (receptor-binding motif, RBM) among HKU5 isolates (Genbank ID: HKU5-17s: AGP04930.1, HKU5-24s: AGP04936.1, HKU5-33s: AGP04943.1, HKU5-27s: AGP04938.1, HKU5-20s: AGP04933.1, HKU5-28s: AGP04939.1, HKU5-32s: AGP04942.1, HKU5-21s: AGP04934.1, HKU5-19s: AGP04932.1, HKU5-1-LMH03F: NC_009020.1, HKU5-31s: AGP04941.1, HKU5-22s: AGP04935.1, HKU5-30s: AGP04940.1, HKU5-2: ABN10884.1, HKU5-1: YP_001039962.1, HKU5-JX2020-Q257: WCC63324.1, HKU5-JX2020-Q279: WCC63520.1, HKU5-GD2017-Q262: WCC63369.1, HKU5r-BY140535: AWH65910.1) shown as logoplot. Amino acids are colored according to their chemical properties: hydrophobic (black), polar non-charged (green), positively charged (blue) and negatively charged. **(E)** Amino acid sequence alignment of the HKU5-19s, HKU5-33s and HKU5-1 RBMs. Residues indicated with a red background are strictly conserved whereas residues indicated with a white background show variability with red text indicating conserved residues and black indicating variable residues. The residue numbering corresponds to HKU5-19s. HKU5-33s insertions are shown as black dots. The blue lines indicate residues outside the RBM shown for visualization purposes around the HKU5-33s insertions. **(F-H)** Biolayer interferometry analysis of the P.abr ACE2 ectodomain binding to the HKU5-19s (F), HKU5-33s (G) or HKU5-1s (H) RBDs immobilized on streptavidin (SA) biosensors. Binding avidities were determined by steady state kinetics and are reported as apparent affinities ($K_{D,app}$). The vertical dotted lines indicate the transition between association and dissociation phases. One representative out of two biological replicates with independently produced batches of proteins are shown for F-H. **(I-L)** Comparison of the structures of the P.abr ACE2-bound HKU5 RBD (PDB 9D32, this study), the P.pip ACE2-bound NeoCoV RBD (PDB 7WPO[40]), hDPP4-bound MERS-CoV RBD (PDB 4KR0[56]) and hDPP4-bound HKU4 RBD (PDB 4QZV[39]). See also Fig. S4-S5 and Tables 1-2.

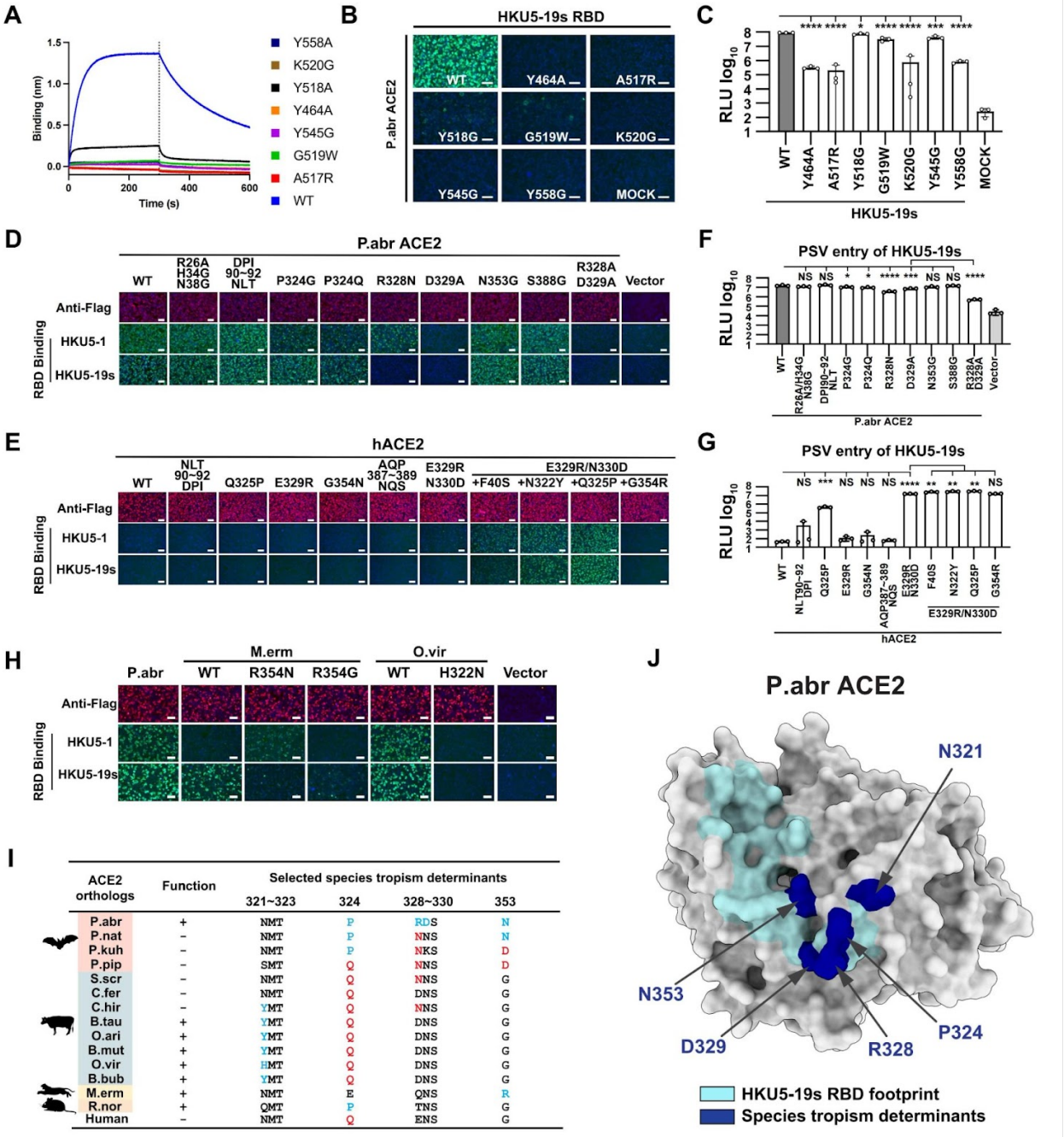


Figure 5.3: HKU5 molecular determinants of ACE2 host species tropism

A, Binding of the P.abr ACE2 construct comprising the peptidase and the dimerization domains (residues 20-724) to the wildtype (WT) HKU5-19s and to the listed RBD interface mutants immobilized on biolayer interferometry streptavidin (SA) biosensors. **B-C**, RBD-hFc binding (B) and pseudovirus entry (C) efficiencies of HKU5-19s mutants in HEK293T cells transiently expressing P.abr ACE2. **D-E,** HKU5-19s and HKU5-1 RBD-hFc binding to HEK293T cells transiently expressing the indicated P.abr ACE2 or hACE2 mutants assessed by immunofluorescence. **F-G,** HKU5-19s S VSV pseudovirus entry into HEK293T cells transiently expressing the indicated ACE2 mutants. **H**, HKU5-

19s and HKU5-1 RBD-hFc binding to HEK293T cells transiently expressing wildtype and mutated M.erm ACE2 and O.vir ACE2. **I,** Summary of ACE2 residues governing species tropism for HKU5-19s. Favorable and unfavorable residues in ACE2 orthologs are highlighted in blue and red, respectively, using P.abr ACE2 residue numbering. (+): functional; (-): non-functional. **J,** Critical determinants of HKU5-19s receptor species tropism are mapped in navy color on the P.abr ACE2 structure rendered as a gray surface. The rest of the HKU5-19s RBD footprint is shown in cyan. Data are shown as the MEAN ± SD of 3 biological replicates for C, F, and G. Statistical analyses used unpaired two-tailed t-tests: *: $p < 0.05$,**: $p < 0.01$, ***: $p < 0.005$, and ****: $p < 0.001$, NS: not-significant. One representative out of two biological replicates with independently produced batches of proteins are shown for A. Data representative of two independent experiments for C, F and G each performed with three biological replicates. Scale bars: 100 μm.

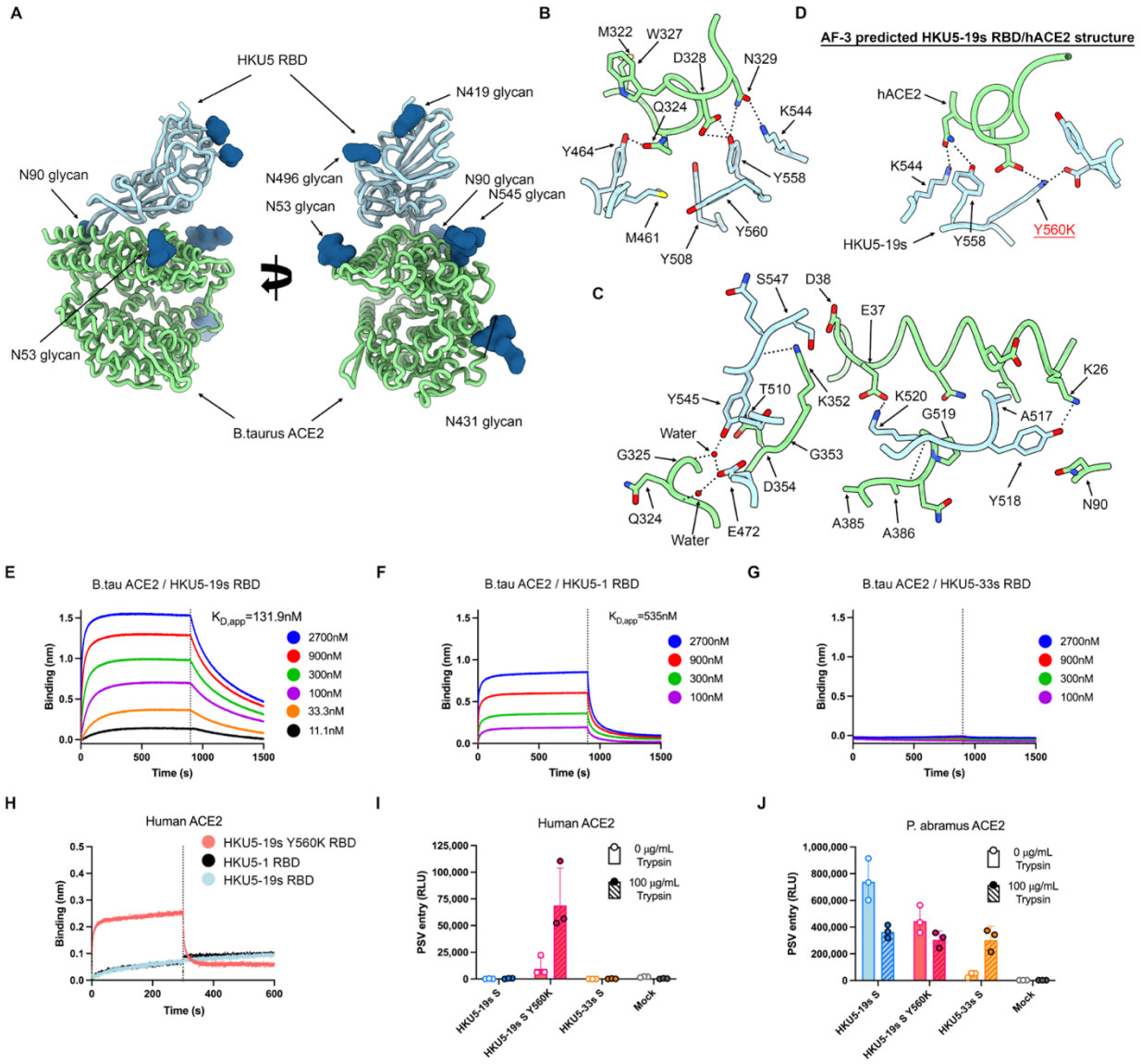


Figure 5.4: HKU5 utilization of artiodactyl and human ACE2 receptors

(A) Ribbon diagrams in two orthogonal orientations of the cryo-EM structure of the HKU5-19s RBD (light blue) bound to the B.tau ACE2 peptidase domain (light green). **(B-C)** Zoomed-in views of the interface highlighting key interactions between the HKU5-19s RBD and B.tau ACE2. Selected polar interactions are shown as black dotted lines. **(D)** AlphaFold3-predicted structure of the interface between the HKU5-19s Y560K RBD and hACE2. **(E-G)** Biolayer interferometry analysis of the B.tau ACE2 ectodomain binding to the HKU5-19s (E), HKU5-1 (F) and HKU5-33s (G) RBDs. Binding avidities were determined by steady state kinetics and are reported as apparent affinities ($K_{D,app}$). **(H)** Evaluation of binding of the hACE2 ectodomain dimer to the immobilized wildtype HKU5-19s and HKU5-1 RBDs and to the HKU5-19s Y560K mutant RBD. In panels E-H, the vertical dotted lines indicate the transition between association and dissociation phases. One representative out of

two biological replicates with independently produced batches of proteins are shown for E-H. **(I-J)** Pseudovirus entry efficiencies in HEK293T cells transiently expressing hACE2 (I) or P.abr ACE2 (J) of wildtype and Y560K HKU5-19s S VSV and wildtype HKU5-33s S VSV. Pseudoviruses were either used untreated or after incubation with 100 µg/mL trypsin for 15 min at room temperature before addition of trypsin soybean inhibitor to stop the reaction. Mock: VSV pseudovirus produced without transfecting an S construct. Each bar represents the geometric mean of three biological replicates with geometric SD and each dot represents the geometric mean of three technical replicates. See also Fig. S4-S5 and Table1.

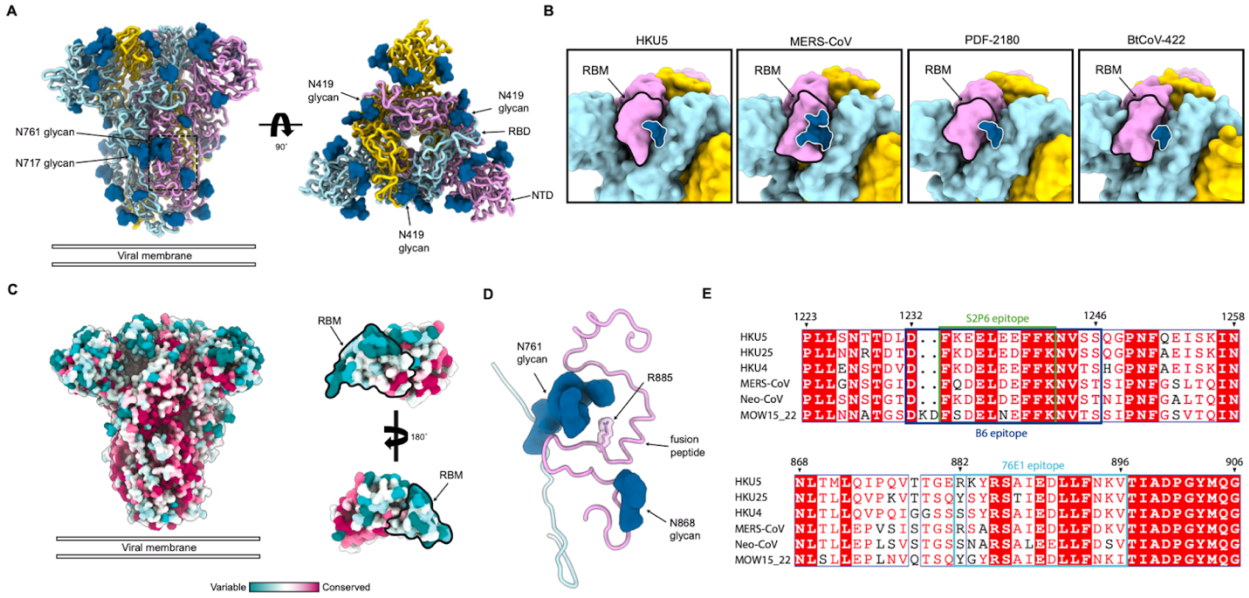


Figure 5.5: Architecture and conservation of the HKU5 S glycoprotein trimer.

(A) Ribbon diagrams in two orthogonal orientations of the cryo-EM structure of prefusion HKU5-19s S. The dotted black box outlines the S_2' cleavage site and fusion peptide region shown in (D). **(B)** Close-up view of the receptor-binding motif (RBM, outlined in black) and shielding glycan (outlined in white) in HKU5-19s, MERS-CoV (PDB 6Q04), PDF-2180 (PDB 7U6R), and BtCoV-422 (PDB 8SAK). **(C)** Conservation analysis using Consurf[82] of S residues across the merbecovirus phylogeny shown in Fig. 1A mapped on prefusion HKU5-19s S and the RBD. Residue conservation ranges from variable (teal), average (white), to conserved (red). **(D)** Zoomed-in view of the HKU5 S_2' cleavage site (including R885) and adjacent fusion peptide region highlighting the glycans at position N761 and N868 limiting access. **(E)** Amino acid sequence alignment of one representative S glycoprotein from each of the six merbecovirus clades focused on the epitopes recognized by the S2P6 (green) and B6 (navy) stem helix-directed antibodies as well as the 76E1 (cyan) fusion peptide-directed antibodies. Residues indicated with a red background are strictly conserved whereas residues indicated with a white background show variability with red text indicating conserved residues and black indicating variable residues. The residue numbering corresponds to

HKU5-19s. Alignment and visualization were done using ESPript[83]. See also Fig. S6 and Table 1.

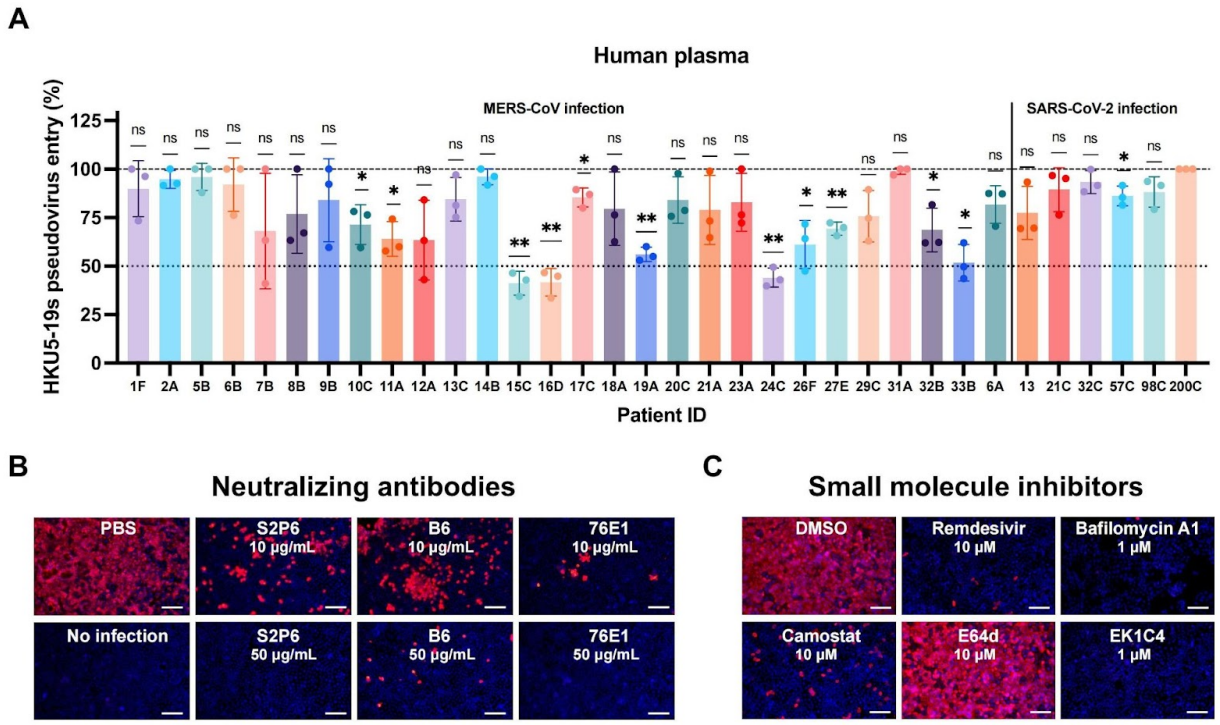


Figure 5.6: Identification of countermeasures against HKU5 merbecoviruses.

A, Neutralization of HKU5-19s S VSV pseudovirus mediated by a panel of MERS-CoV or SARS-CoV-2 infection-elicited polyclonal antibody-containing human plasma. Bars represent the mean of three biological replicates with SD and each data point corresponds to the mean of two technical replicates within each biological replicate carried out with distinct batches of pseudoviruses. Statistical significance was measured using the one sample t test against the maximum VSV entry value of 100%. ns:P > 0.05, *:P≤0.05, **:P≤ 0.01, ***:P≤ 0.001, ****:P≤ 0.0001. **B-C**, Evaluation of inhibition of authentic HKU5-1 propagation in Caco-2 cells stably expressing P.abr ACE2 by the indicated concentration of broadly neutralizing antibodies (B) or small molecule/peptide inhibitors (C). HKU5-1 was detected by immunofluorescence using an anti-HKU5 N protein antibody at 24 hpi. Scale bars: 200 μm. See also Fig. S7.

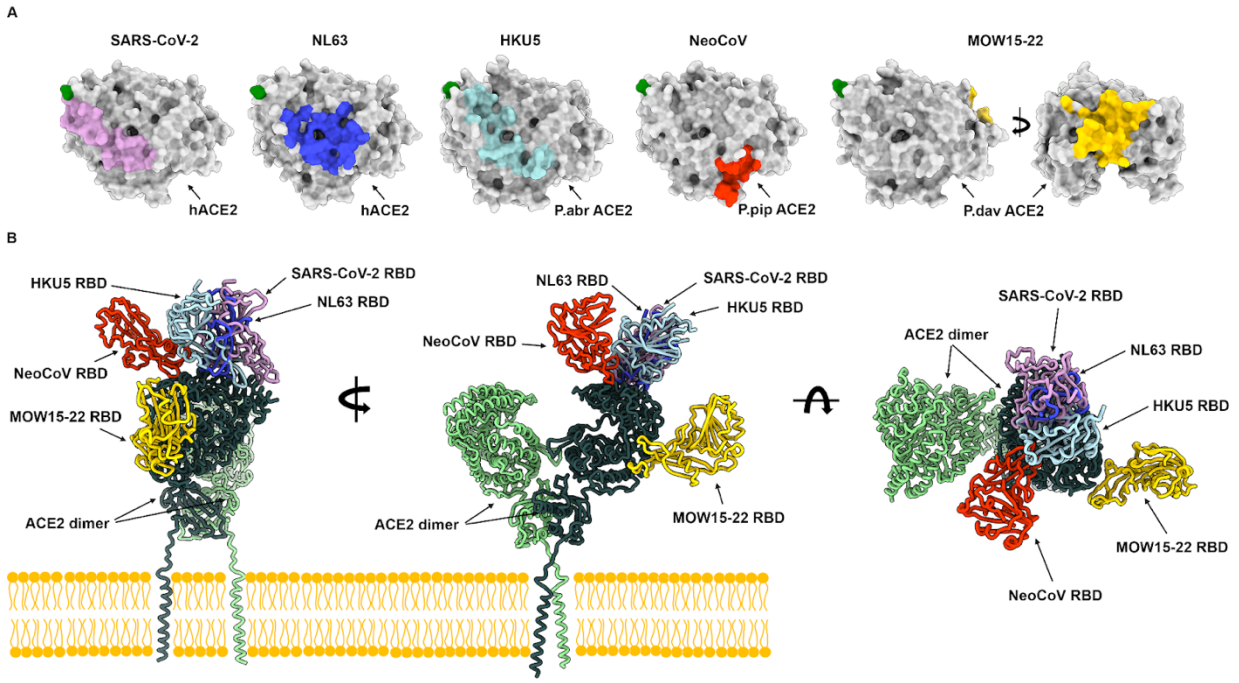
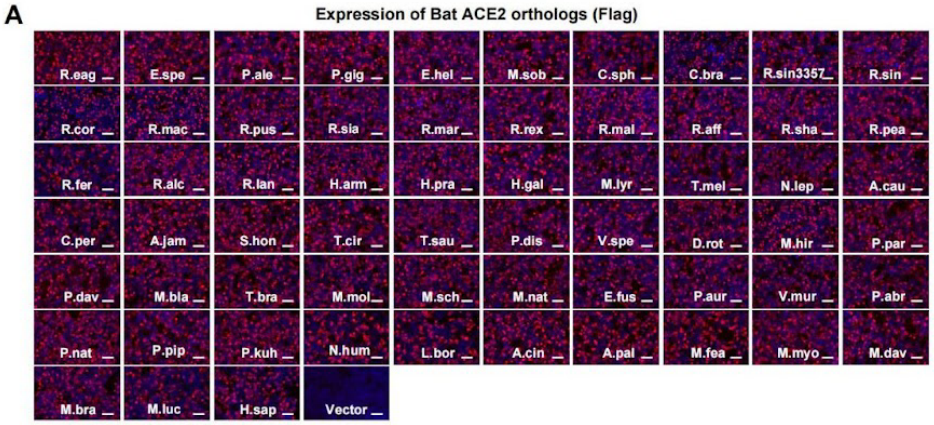
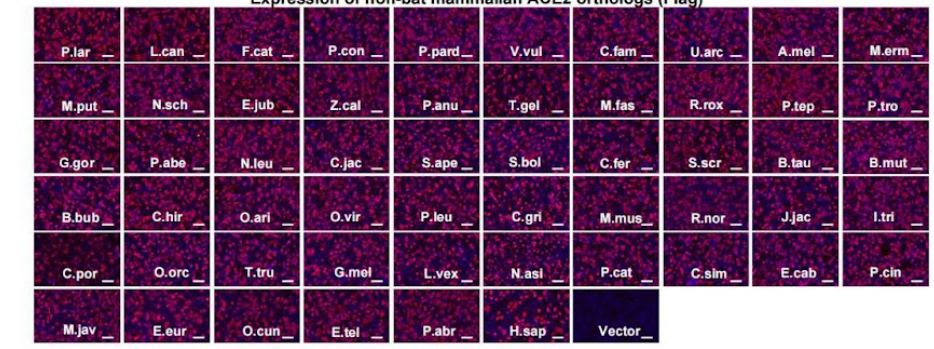
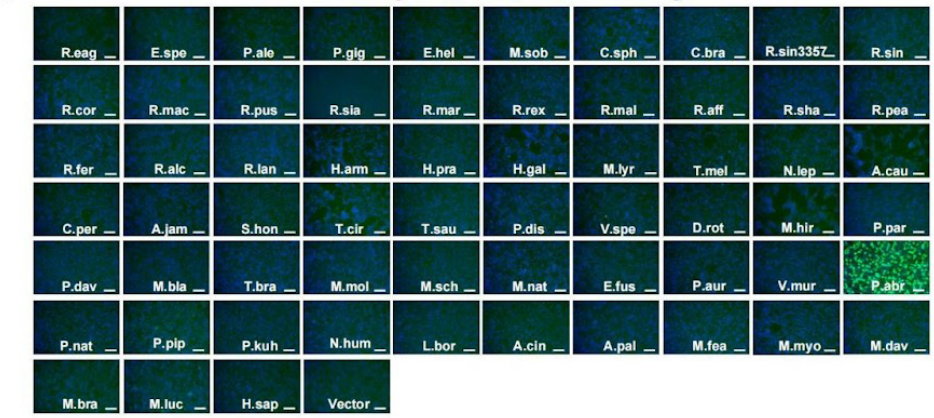


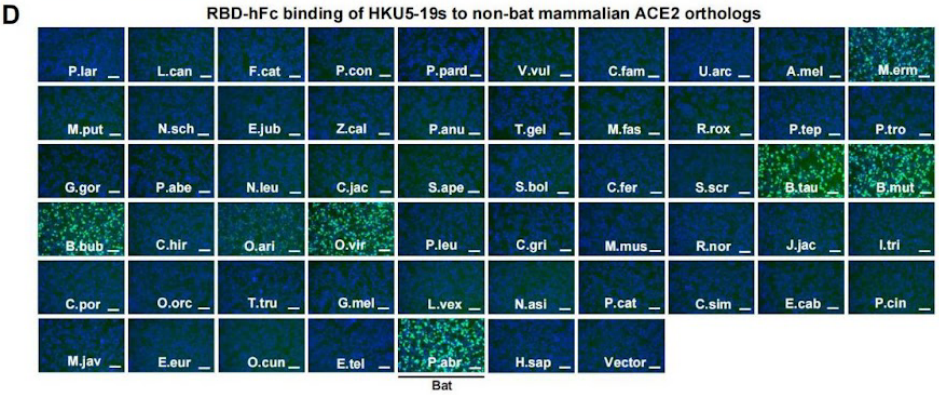
Figure 5.7: Coronaviruses have evolved ACE2 utilization at least five times independently.

(A) RBD footprints of ACE2-using coronaviruses on their cognate receptors. **(B)** Comparison of the binding modes of the SARS-CoV-2 (PDB 7TN0), NL63 (PDB 3KBH), HKU5 RBDs, NeoCoV (PDB 7WPO) and MOW15-22 (PDB 9C6O) to bat ACE2 (not shown for clarity) or hACE2 (PDB 6M1D, B0AT1 not shown for clarity).


A
Expression of Bat ACE2 orthologs (Flag)
R.eag
E.spe
P.ale
P.gig
E.hel
M.sob
C.sph
C.bra
R.sin3357
R.sin
R.cor
R.mac
R.pus
R.sia
R.mar
R.rex
R.mal
R.aff
R.sha
R.pea
R.fer
R.alc
R.lan
H.arm
H.pra
H.gal
M.lyr
T.mel
N.lep
A.cau
C.per
A.jam
S.hon
T.cir
T.sau
P.dis
V.spe
D.rot
M.hir
P.par
P.dav
M.bla
T.bra
M.mol
M.sch
M.nat
E.fus
P.aur
V.mur
P.abr
P.nat
P.pip
P.kuh
N.hum
L.bor
A.cin
A.pal
M.fea
M.myo
M.dav
M.bra
M.luc
H.sap
Vector

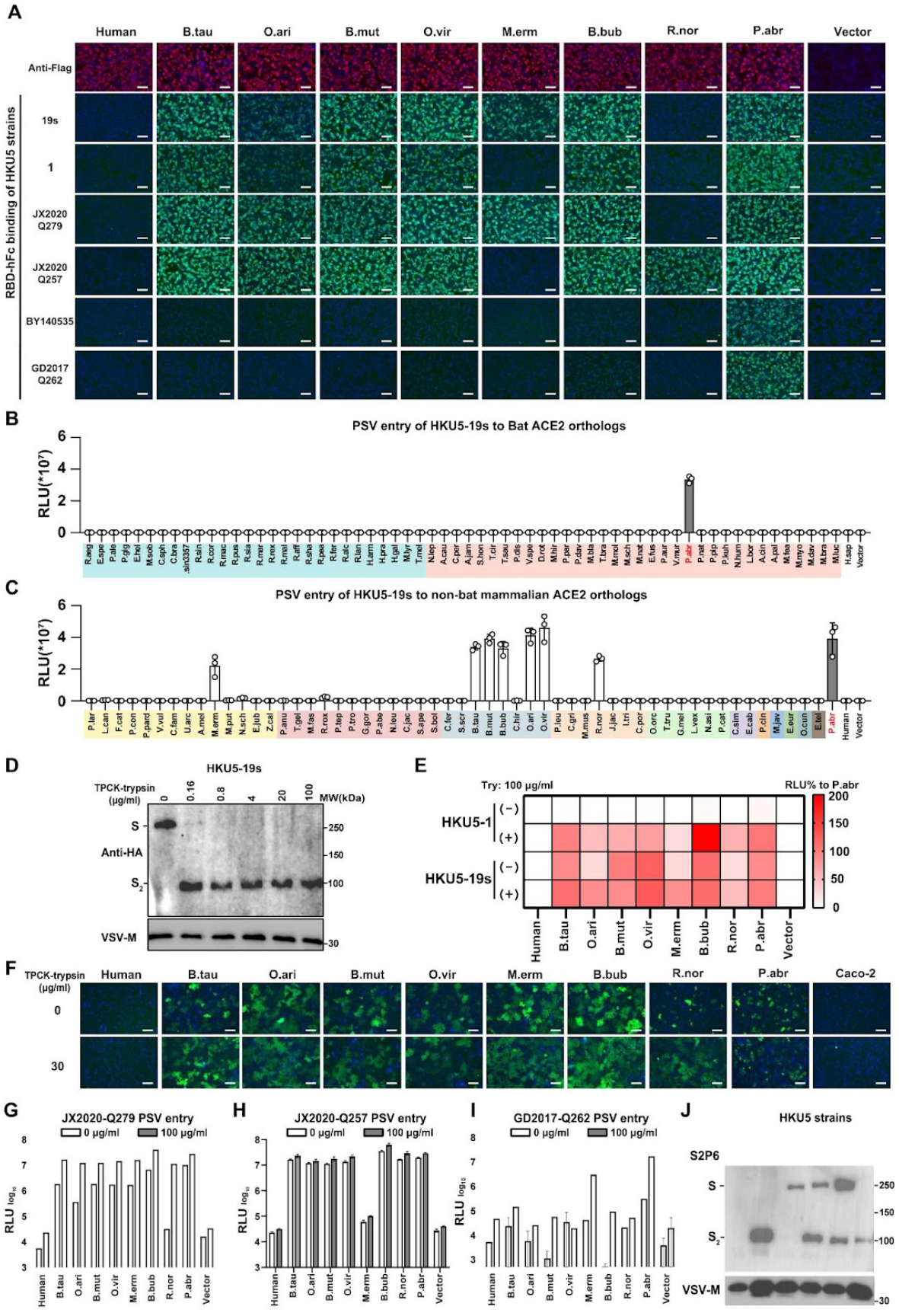

B
Expression of non-bat mammalian ACE2 orthologs (Flag)
P.lar
L.can
F.cat
P.con
P.pard
V.vul
C.fam
U.arc
A.mel
M.erm
M.put
N.sch
E.jub
Z.cal
P.anu
T.gel
M.fas
R.rox
P.tep
P.tro
G.gor
P.abe
N.leu
C.jac
S.ape
S.bol
C.fer
S.scr
B.tau
B.mut
B.bub
C.hir
O.ari
O.vir
P.leu
C.gri
M.mus
R.nor
J.jac
I.tri
C.por
O.orc
T.tru
G.mel
L.vex
N.asi
P.cat
C.sim
E.cab
P.cin
M.jav
E.eur
O.cun
E.tel
P.abr
H.sap
Vector


C
RBD-hFc binding of HKU5-19s to Bat ACE2 orthologs
R.eag
E.spe
P.ale
P.gig
E.hel
M.sob
C.sph
C.bra
R.sin3357
R.sin
R.cor
R.mac
R.pus
R.sia
R.mar
R.rex
R.mal
R.aff
R.sha
R.pea
R.fer
R.alc
R.lan
H.arm
H.pra
H.gal
M.lyr
T.mel
N.lep
A.cau
C.per
A.jam
S.hon
T.cir
T.sau
P.dis
V.spe
D.rot
M.hir
P.par
P.dav
M.bla
T.bra
M.mol
M.sch
M.nat
E.fus
P.aur
V.mur
P.abr
P.nat
P.pip
P.kuh
N.hum
L.bor
A.cin
A.pal
M.fea
M.myo
M.dav
M.bra
M.luc
H.sap
Vector


D
RBD-hFc binding of HKU5-19s to non-bat mammalian ACE2 orthologs
P.lar
L.can
F.cat
P.con
P.pard
V.vul
C.fam
U.arc
A.mel
M.erm
M.put
N.sch
E.jub
Z.cal
P.anu
T.gel
M.fas
R.rox
P.tep
P.tro
G.gor
P.abe
N.leu
C.jac
S.ape
S.bol
C.fer
S.scr
B.tau
B.mut
B.bub
C.hir
O.ari
O.vir
P.leu
C.gri
M.mus
R.nor
J.jac
I.tri
C.por
O.orc
T.tru
G.mel
L.vex
N.asi
P.cat
C.sim
E.cab
P.cin
M.jav
E.eur
O.cun
E.tel
P.abr
H.sap
Vector
Bat

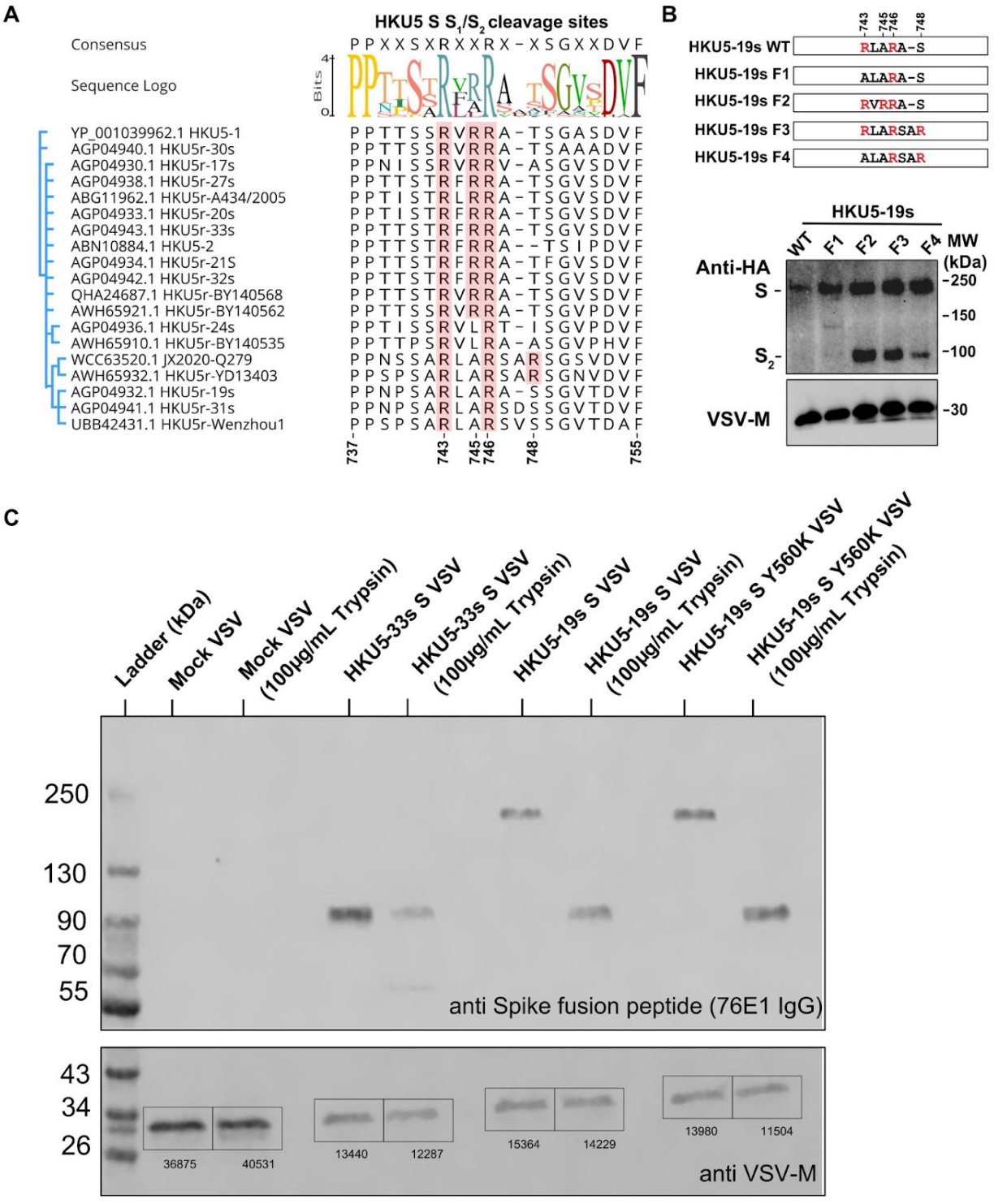
Supplemental Figure 5.1: related to Fig. 1. HKU5 utilizes several mammalian ACE2s as receptor.

(**A-B**) Immunofluorescence analysis of the expression of bat (A) or non-bat mammalian (B) ACE2 orthologs in HEK293T cells by detecting 3×Flag tags fused to the C-terminal of the constructs. (**C-D**) Immunofluorescence analysis of binding of the HKU5-19s RBD-hFc to HEK293T cells transiently transfected with the indicated bat (C) or non-bat (D) mammalian ACE2 orthologs. Scale bars: 100 μm. Data representative of two independent experiments for C-D. Scale bars: 100 μm.


A
Human
B.tau
O.ari
B.mut
O.vir
M.erm
B.bub
R.nor
P.abr
Vector
Anti-Flag
RBD-hFc binding of HKU5 strains
19s
1
JX2020 Q279
JX2020 Q257
BY140535
GD2017 Q262
B
PSV entry of HKU5-19s to Bat ACE2 orthologs
RLU(*10⁷)
C
PSV entry of HKU5-19s to non-bat mammalian ACE2 orthologs
RLU(*10⁷)
D
HKU5-19s
TPCK-trypsin (μg/ml)
0
0.16
0.8
4
20
100
MW(kDa)
S
Anti-HA
S2
VSV-M
250
150
100
30
E
Try: 100 μg/ml
RLU% to P.abr
HKU5-1
HKU5-19s
(-)
(+)
200
150
100
50
0
F
TPCK-trypsin (μg/ml)
Caco-2
G
JX2020-Q279 PSV entry
0 μg/ml
100 μg/ml
RLU log10
H
JX2020-Q257 PSV entry
I
GD2017-Q262 PSV entry
J
HKU5 strains
S2P6

Supplemental Figure 5.2, related to Fig. 1. Evaluation of receptor species tropism for several HKU5 strains.

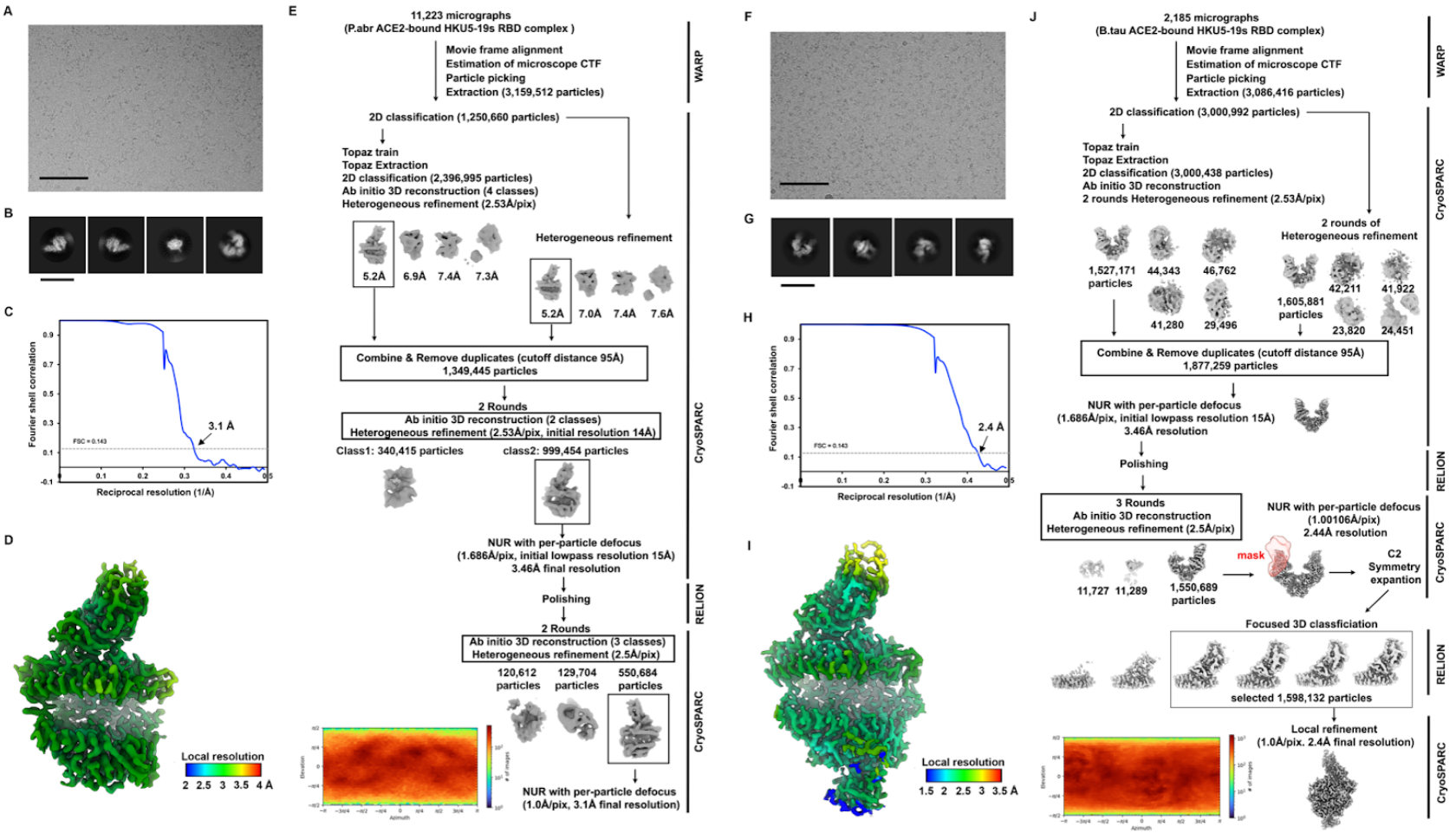
(A) Binding (green) of the HKU5-19s, HKU5-1, HKU5-JX2020-Q279, HKU5-JX2020-Q257, HKU5r-BY140535 and HKU5-GD2017-Q262 RBD-hFc constructs to HEK293T cells transiently transfected with the indicated mammalian ACE2 orthologs evaluated by immunofluorescence assessed with an anti-hFc antibody (binding) or anti-Flag antibody (ACE2 expression). **(B-C)** Entry of HKU5-19s S VSV pseudovirus into HEK293T cells transiently transfected with the indicated bat (A) or non-bat (B) mammalian ACE2 orthologs. **(D)** Western blot assessment of HKU5-19s S cleavage in VSV PSV with or without trypsin treatment at the indicated concentrations. (**E**) Entry efficiency of HKU5-19s S and HKU5-1 S VSV PSV (with or without pretreatment with 100 µg/ml TPCK-treated trypsin (Try)) in HEK293T cells transiently transfected with the indicated ACE2 orthologs . Data are shown as the MEAN for n=3 biological replicates. (**F**) HKU5-19s S-mediated cell-cell fusion using Caco-2 cells stably expressing the indicated ACE2 orthologs with or without the addition of 30 µg/ml TPCK-treated trypsin. **(G-I)** Entry of HKU5-JX2020-Q279, HKU5-JX2020-Q257 and HKU5-GD2017-Q262 S VSV pseudovirus (with or without pretreatment with 100 µg/ml TPCK-treated trypsin) into HEK293T cells transiently transfected with the indicated mammalian ACE2 orthologs. (**J**) VSV packaging efficiencies of different HKU5 S glycoproteins detected using the S2P6 stem helix antibody. VSV-M was used as a loading control. Data are shown as the MEAN ± SD for A-B and G-I. n=3 biological replicates. Data representative of two independent experiments for C-D and F-J, and a single experiment for A-B and E. Scale bars: 100 µm.



Supplemental Figure 5.3, related to Fig. 1 and 4. HKU5 S proteolytic processing during biogenesis.

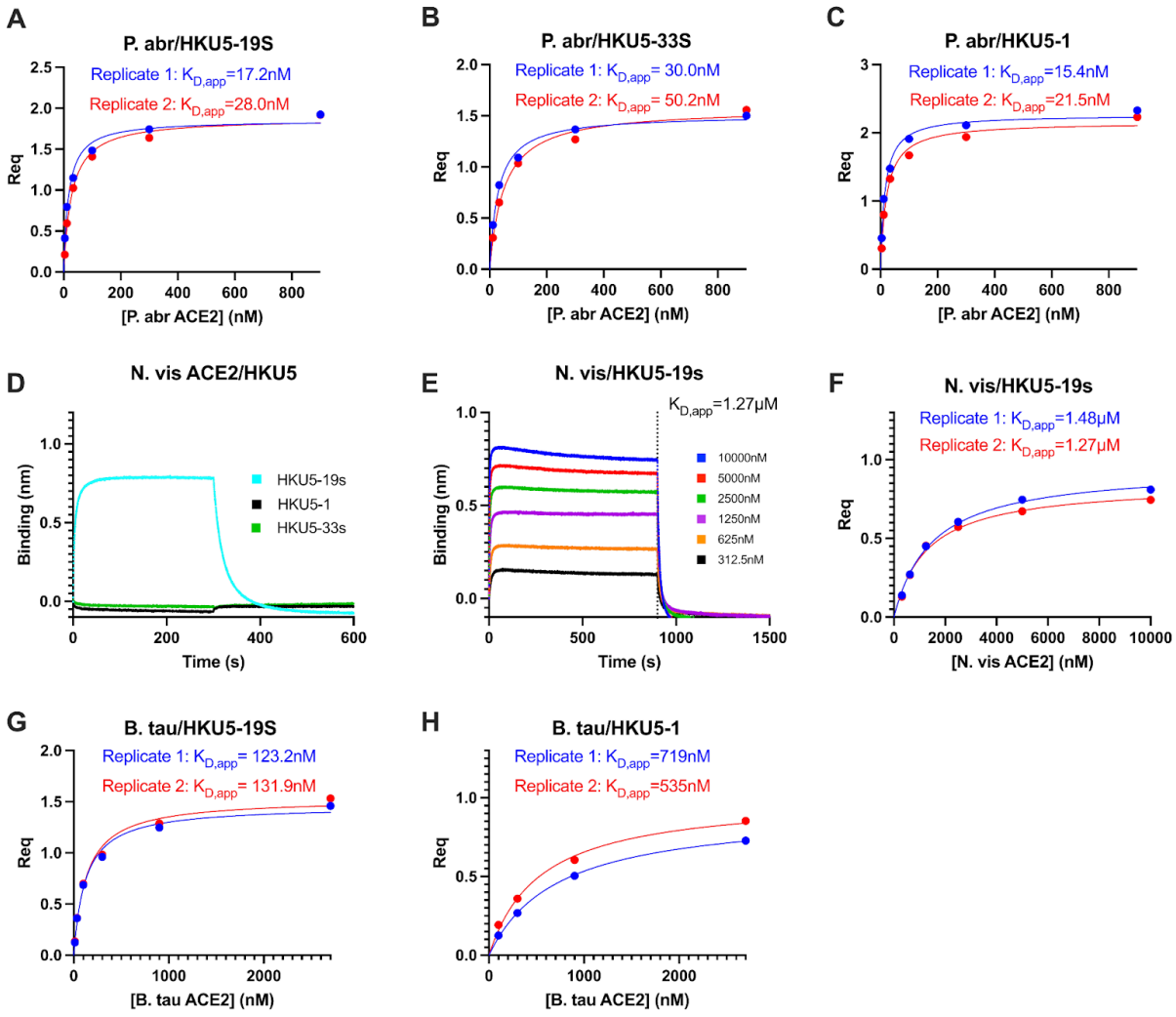
A, Sequence analysis of S glycoprotein S_1/S_2 junction from the indicated HKU5 strains, with arginine (R) residues highlighted with a red background. **B**, Quantification of S glycoprotein proteolytic processing and incorporation in VSV pseudoviruses for wildtype

and S_1/S_2 HKU5-19s mutants analyzed by Western blot detecting the C-terminal-fused HA tag**.** VSV-M was used as a loading control. HKU5-19s S residue numbering is shown. **C**, Quantification of S glycoprotein and VSV-M in VSV pseudoviruses for HKU5-19s S, HKU5-19s S Y560K, HKU5-33s S and mock VSV analyzed by Western blot detecting fusion peptide with the 76E1 antibody or VSV-M with anti-VSV-M [23H12] antibody (Kerafast).


A
B
C
Fourier shell correlation
FSC = 0.143
3.1 Å
Reciprocal resolution (1/Å)
D
Local resolution
2 2.5 3 3.5 4 Å
E
11,223 micrographs
(P.abr ACE2-bound HKU5-19s RBD complex)
Movie frame alignment
Estimation of microscope CTF
Particle picking
Extraction (3,159,512 particles)
WARP
2D classification (1,250,660 particles)
Topaz train
Topaz Extraction
2D classification (2,396,995 particles)
Ab initio 3D reconstruction (4 classes)
Heterogeneous refinement (2.53Å/pix)
5.2Å
6.9Å
7.4Å
7.3Å
Heterogeneous refinement
5.2Å
7.0Å
7.4Å
7.6Å
Combine & Remove duplicates (cutoff distance 95Å)
1,349,445 particles
2 Rounds
Ab initio 3D reconstruction (2 classes)
Heterogeneous refinement (2.53Å/pix, initial resolution 14Å)
Class1: 340,415 particles
class2: 999,454 particles
CryoSPARC
NUR with per-particle defocus
(1.686Å/pix, initial lowpass resolution 15Å)
3.46Å final resolution
Polishing
RELION
2 Rounds
Ab initio 3D reconstruction (3 classes)
Heterogeneous refinement (2.5Å/pix)
120,612 particles
129,704 particles
550,684 particles
CryoSPARC
NUR with per-particle defocus
(1.0Å/pix, 3.1Å final resolution)
F
G
H
Fourier shell correlation
FSC = 0.143
2.4 Å
Reciprocal resolution (1/Å)
I
Local resolution
1.5 2 2.5 3 3.5 Å
J
2,185 micrographs
(B.tau ACE2-bound HKU5-19s RBD complex)
Movie frame alignment
Estimation of microscope CTF
Particle picking
Extraction (3,086,416 particles)
WARP
2D classification (3,000,992 particles)
Topaz train
Topaz Extraction
2D classification (3,000,438 particles)
Ab initio 3D reconstruction
2 rounds Heterogeneous refinement (2.53Å/pix)
CryoSPARC
2 rounds of
Heterogeneous refinement
1,527,171 particles
44,343
46,762
41,280
29,496
1,605,881 particles
42,211
41,922
23,820
24,451
Combine & Remove duplicates (cutoff distance 95Å)
1,877,259 particles
NUR with per-particle defocus
(1.686Å/pix, initial lowpass resolution 15Å)
3.46Å resolution
Polishing
RELION
3 Rounds
Ab initio 3D reconstruction
Heterogeneous refinement (2.5Å/pix)
NUR with per-particle defocus
(1.00106Å/pix)
2.44Å resolution
CryoSPARC
11,727
11,289
1,550,689 particles
mask
C2 Symmetry expantion
Focused 3D classficiation
RELION
selected 1,598,132 particles
Local refinement
(1.0Å/pix. 2.4Å final resolution)
CryoSPARC

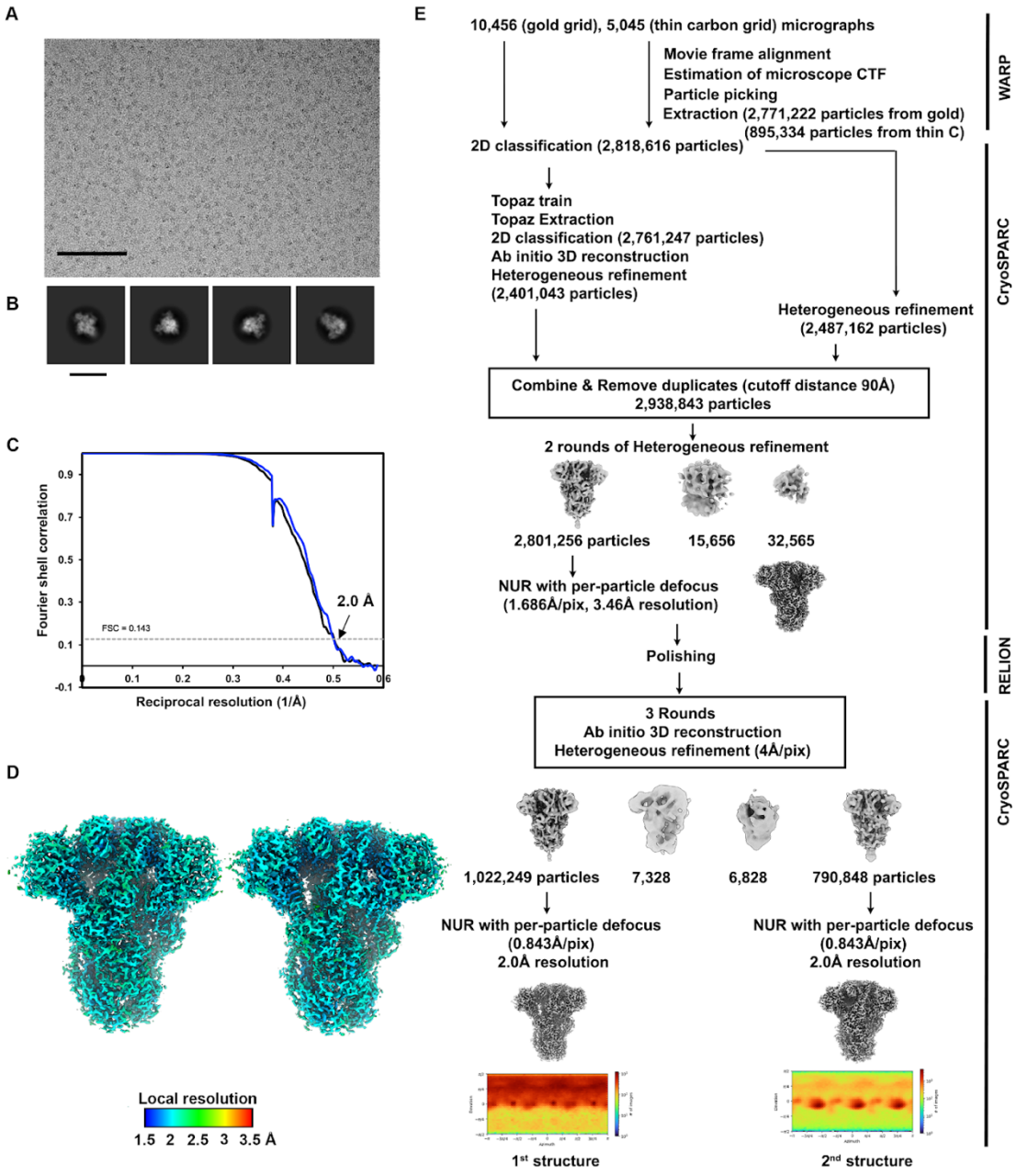
Supplemental Figure 5.4, related to Fig. 2 and 4. Cryo-EM data processing of the P.abr ACE2-bound HKU5-19s RBD and B.tau ACE2-bound HKU5-19s RBD datasets.

A, B, F and G, Representative electron micrographs (A, F) and 2D class averages (B, G) of the P.abr ACE2-bound HKU5-19s RBD complex (A, B) or B.tau ACE2-bound HKU5-19s RBD complex (F, G) embedded in vitreous ice. Scale bars: 100 nm (A, F) and 150 Å (B, G). **C,H,** Gold-standard Fourier shell correlation curve of the P.abr ACE2-bound HKU5-19s RBD reconstruction (C) or B.tau ACE2-bound HKU5-19s RBD reconstruction (H). The 0.143 cutoff is indicated by a horizontal dashed line. **D,I,** Local resolution estimation of the P.abr ACE2-bound HKU5-19s RBD reconstruction (D) or B.tau ACE2-bound HKU5-19s RBD reconstruction (I) calculated using cryoSPARC and plotted on the sharpened maps. **E,J,** Data processing flowchart. CTF: contrast transfer function; NUR: non-uniform refinement. The angular distribution of particle images calculated using cryoSPARC is shown as a heat map.



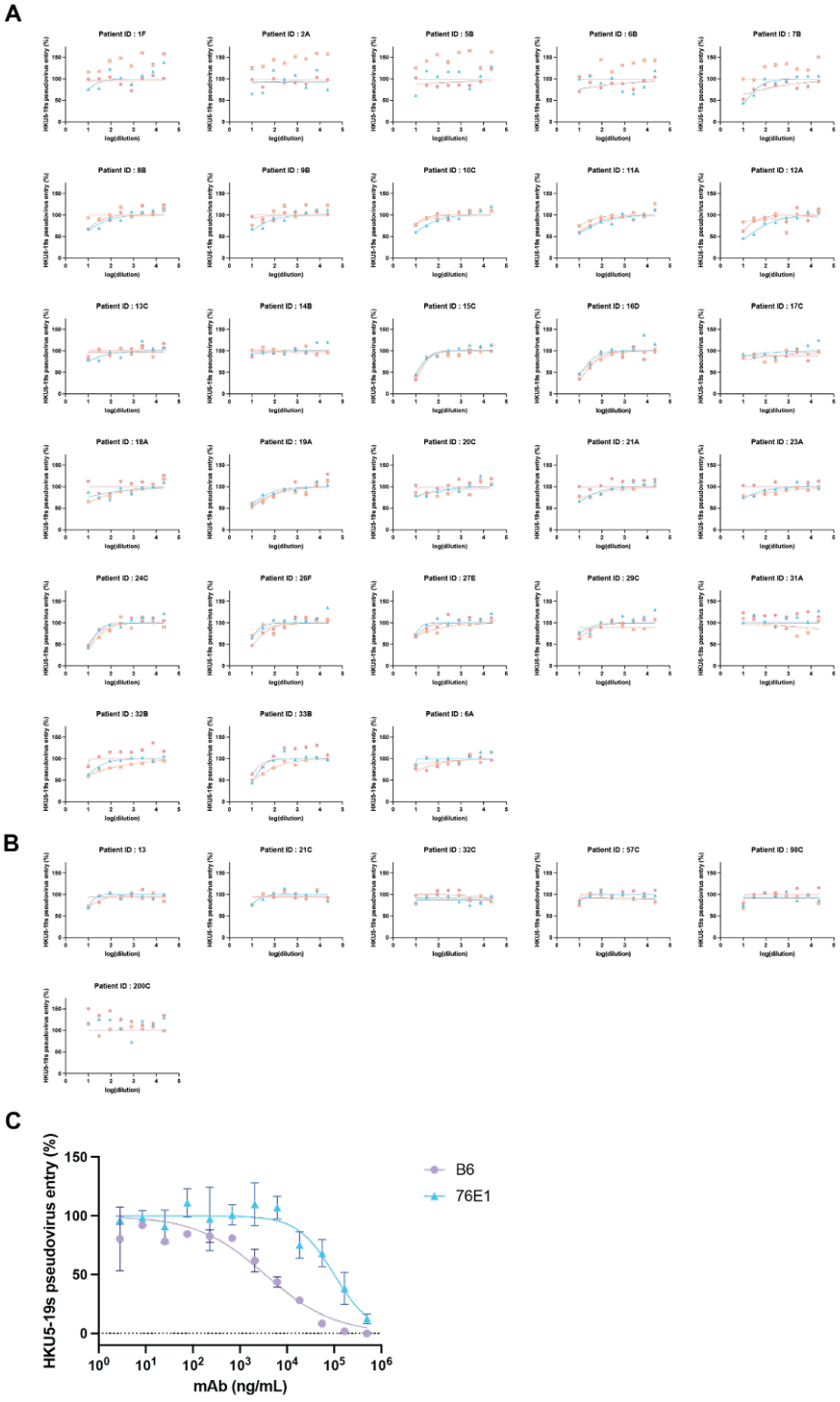
Supplemental Figure 5.5, related to Fig. 2 and 4. Biolayer interferometry analysis of the P.abr, B.tau or N.vis ACE2 ectodomain dimers binding to immobilized HKU5-19s, HKU5-33s or HKU5-1 RBDs.

Steady-state analysis was used to estimate apparent binding affinities ($K_{D, app}$). Two replicates from proteins produced independently are shown. **A-C** Steady state analysis of P.abr ACE2 binding to HKU5-19s (A), HKU5-33s (B), and HKU5-1 (C) RBDs. **D** BLI sensorgram of HKU5-19s, HKU5-1 and HKU5-33s RBDs binding to 5μM N.vis ACE2. **E** BLI sensorgram of N. vison ACE2 binding at various concentrations to immobilized HKU5-19s. One representative out of two biological replicates is shown. **F** Steady-state analysis of N. vison ACE2 binding to the HKU5-19s RBD. **G-H** Steady state-analysis of B.taurus ACE2 binding to the HKU5-19s (F) and HKU5-1 (G) RBDs respectively.



Supplemental Figure 5.6, , related to Fig. 5. Cryo-EM data processing of the prefusion HKU5-19s S trimer dataset.

(A-B) Representative electron micrograph and 2D class averages of the complex embedded in vitreous ice. Scale bars: 100 nm (A) and 200 Å (B). **(C)** Gold-standard Fourier shell correlation curves of the two structures (1st class structure in black, 2nd class structure in blue). The 0.143 cutoff is indicated by a horizontal dashed line. **(D)** Local resolution estimation of two structures calculated using cryoSPARC and plotted on the sharpened map. **(E)** Data processing flowchart. CTF: contrast transfer function; NUR: non-uniform refinement. The angular distribution of particle images calculated using cryoSPARC is shown as a heat map.



Supplemental Figure 5.7, related to Fig. 6. Dose-response curves of HKU5-19s VSV pseudovirus neutralization by MERS-CoV or SARS-CoV-2 infection-elicited plasma or monoclonal antibodies.

A-B, Dose-response curves of HKU5-19s VSV pseudovirus neutralization by MERS-CoV infection-elicited (**A**) or SARS-CoV-2 infection-elicited (**B**) plasma. Each data point represents the mean of two technical replicates and three biological replicates with three distinct batches of pseudoviruses are shown with different colors. C, Dose-response curves of HKU5-19s VSV pseudovirus neutralization by fusion machinery-directed monoclonal antibodies. Each data point represents the mean of two technical replicates. A representative curve from three biological replicates is shown.

Data 5.1 Tables 1-3

Table 5-1, related to Fig. 2, 4 and 5. Cryo-EM data collection and refinement statistics.

	HKU5 RBD – P.abr ACE2 PDB 9D32 EMD 46512	HKU5 RBD – B.tau ACE2 PDB 9E0I EMD 47358	HKU5 Spike PDB ### EMD ### Conformation 1	HKU5 Spike PDB ### EMD ### Conformation 2
Data collection and processing				
Magnification	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300
Electron exposure (e–/Å^2)	60	60	60	60
Defocus range (μm)	-0.2 - -3.5	-0.2 - -3.5	-0.2 - -3.5	-0.2 - -3.5
Pixel size (Å)	0.843	0.843	0.843	0.843
Symmetry imposed	C1	C2	C3	C3
Initial particle images (no.)	3,159,512	3,086,416	3,666,556	3,666,556
Final particle images (no.)	550,684	1,598,132	790,848	1,022,249
Map resolution (Å) **FSC threshold**	3.1 0.143	2.4 0.143	2.0 0.143	2.0 0.143
Refinement				
Model resolution (Å) **FSC threshold**	3.2 0.5	2.5 0.5	2.1 0.5	2.1 0.5
Map sharpening *B* factor (Å^2)	-135	-83.9	-59.5	-61.6

Model composition				
Non-hydrogen atoms	6140	6421	30222	29139
Protein residues	778	786	3561	3525
Ligands	9	14	129	129
Water	0	91	933	933
***B* factors (Å^2)**				
Protein	15.0	32.7	29.93	20.29
Ligand	18.5	89.7	62.58	40.58
Water		22.7	14.76	
R.m.s. deviations				
Bond lengths (Å)	0.010	0.003	0.006	0.005
Bond angles (°)	1.300	0.579	0.957	0.848
Validation				
MolProbity score	1.39	1.38	1.56	1.21
Clashscore	4.9	3.19	2.91	2.63
Poor rotamers (%)	0.88	2.30	1.95	1.15
Ramachandran plot				
Favored (%)	97.28	99.23	96.22	97.42
Allowed (%)	2.59	0.77	3.5	2.58
Disallowed (%)	0.13	0	0.28	0

Table 5-2, related to Fig 2.4. Conservation analysis of P.abr ACE2 residues recognized by HKU5 in hACE2. Non-conserved residues are italicized.

P. abr ACE2	**H. sap ACE2**
R26	K26
L29	L29
V30	*D30*
N33	N33
H34	H34
E37	E37
N38	*D38*
H41	*Y41*

D90	*N90*
I92	*T92*
I93	*V93*
Q96	Q96
M322	M323
T323	T324
P324	*Q325*
G325	G326
W327	W328
R328	*E329*
D329	*N330*
K352	K353
N353	*G354*
D354	*N355*
R356	R357
A385	A386
N386	*A387*
Q387	Q388
S388	*P389*
R392	R393

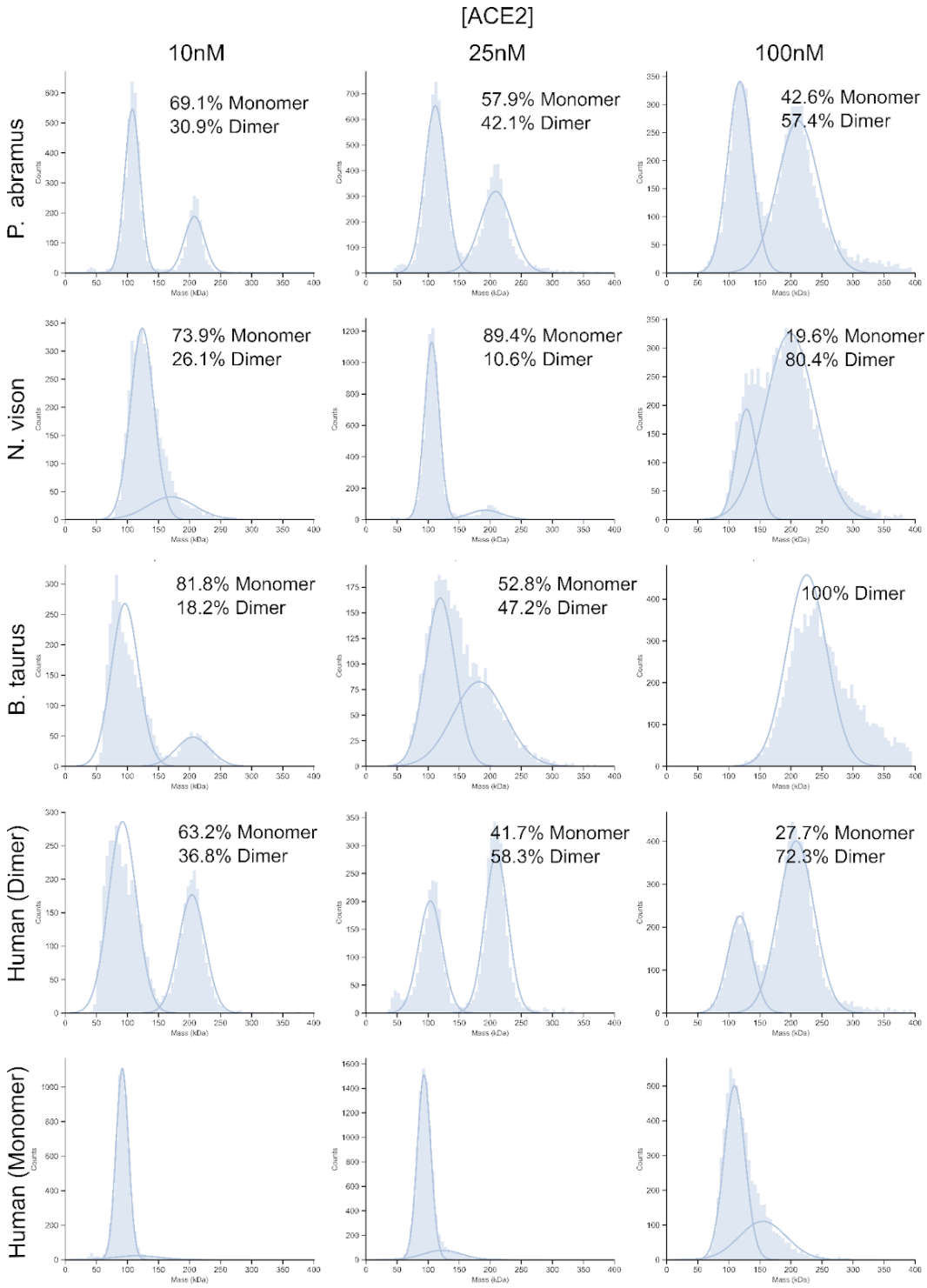
Table 5.3, related to Fig. 6. Demographic information of MERS-CoV-infected and SARS-CoV-2-infected individuals.

Patient ID	Sex	Age	Comorbidities	MERS-CoV ID_{50} at peak neutralizing time point	Mean % Inhibition of HKU5-19s at 1:10 sera dilution
1	M	42	DM, HTN	1352	89.87
2	M	80	ESRD, DM, HTN	10	94.72
5	M	27	DM, HTN	27.96	95.96
6	M	64	DM, HTN	31.99	92.06
7	M	44	DM	672.9	68.11
8	M	51	DM, HTN	214.2	76.82
9	M	70	DM,HTN, IHD	1014	83.96
10	M	35	psychiatric illness	2064	71.4
11	F	55	DM, hypothyroidism, AKI	252.1	64.02
12	M	49	DM	52.07	63.45
13	F	41	DM	112.9	84.5
14	M	65	DM, HTN, IHD	615.5	96.02
15	M	46	DM, HTN, COPD	656.2	41.18
16	M	53	DM, HTN	665.1	41.7
17	M	49	heart failure	265.1	85.44
18	M	53	HTN	891.2	79.64
19	M	32	–	797.5	56.06
20	M	75	DM, HTN, epilepsy, dyslipidemia	209	84.08

21	F	67	DM, HTN	768.4	78.98
23	M	63	DM, BA, HTN	838.4	82.93
24	M	50	DM	537.8	44.05
26	M	31	ESRD, DM, HTN	2120	61.12
27	F	70	ESRD	390.2	69.33
29	M	69	BPH, Afib	738	75.73
31	F	62	ARDS, Alzheimer, DM, HTN	29.33	99.02
32	M	42	DM	1353	68.61
33	M	58	–	1410	51.74

Patient ID	**Sex**	**Age**	**Days since symptoms onset**	**SARS-CoV-2 Wu/G614 ID_{50}**	**Mean % Inhibition of HKU5-19s at 1:10 sera dilution**
13	M	54	32	172	77.46
21C	F	36	26	4	89.38
32C	M	36	30	140	93.29
57C	F	56	35	44	86.18
98C	M	33	38	187	88.2
200C	M	61	42	90	100

DM, diabetes mellitus; HTN, hypertension; ARDS, acute respiratory distress syndrome; ESRD, end stage renal disease; IHD, ischemic heart disease; AKI, acute kidney injury; COPD, chronic obstructive pulmonary disease; BA, bronchial asthma; BPH, benign prostatic hypertrophy; Afib, atrial fibrillation. MERS-CoV ID_{50} values were adapted from Addetia et al[77]. SARS-CoV-2 ID_{50} values were adapted from Bowen et al[30].



Data 5.2: Mass spectrophotometry analysis of the oligomerization state of recombinant ACE2 ectodomains

Histograms of the mass distribution of events detected by mass photometry of purified P. abramus, N. vison, B. taurus, and Human ACE2 dimeric ectodomains at 10nM, 25nM and 100nM as well as the Human ACE2 monomeric ectodomain at 20nM, 50nM, and 200nM.

CHAPTER 6. N-TERMINAL DOMAIN ANTIGENIC MAPPING REVEALS A SITE OF VULNERABILITY FOR SARS-COV-2

Over the next several studies, I switch my focus to from characterizing specific CoV to pandemic response and zoonotic spillover preparedness describing over the next several sections broadly neutralizing antibodies. In this story, I will describe our efforts to understand antibody responses to the NTD. While the RBD has been identified at the immunodominant site targeted by most neutralizing antibodies (Abs) in COVID-19 patient plasma. When this story began, we knew very little about neutralizing Abs binding to epitopes outside the RBD and their contribution to protection. In the following chapter, we describe 41 human monoclonal Abs (mAbs) derived from memory B cells, which recognize the SARS-CoV-2 S N-terminal domain (NTD) and show that a subset of them neutralize SARS-CoV-2 ultrapotently. We define an antigenic map of the SARS-CoV-2 NTD and identify a supersite (designated site i) recognized by all known NTD-specific neutralizing mAbs. However, due to the high degree of divergence in the NTD among bat-borne sarbecoviruses, NTD directed responses demonstrate limited efficacy as a stockpile for pandemic preparedness in the event of a future spillover event.

Adapted from: McCallum M, De Marco A, Lempp FA, Tortorici MA, Pinto D, Walls AC, Beltramello M, Chen A, Liu Z, Zatta F, Zepeda S, di Iulio J, Bowen JE, Montiel-Ruiz M, Zhou J, Rosen LE, Bianchi S, Guarino B, Fregni CS, Abdelnabi R, Foo SC, Rothlauf PW, Bloyet LM, Benigni F, Cameroni E, Neyts J, Riva A, Snell G, Telenti A, Whelan SPJ, Virgin HW, Corti D, Pizzuto MS, Veesler D. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. Cell. 2021 Apr 29;184(9):2332-2347.e16.

doi: 10.1016/j.cell.2021.03.028. Epub 2021 Mar 16. PMID: 33761326; PMCID: PMC7962585.

6.1 Chapter Introduction

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) at the end of 2019 resulted in the ongoing coronavirus disease 2019 (COVID-19) pandemic, bringing the world to a standstill (Zhou et al., 2020; Zhu et al., 2020). The lack of pre-existing immunity to SARS-CoV-2 combined with its efficient human-to-human transmission has already resulted in more than 109 million infections and over 2.3 million fatalities as of February 2021. Although vaccines are being developed and deployed at an unprecedented pace, the timeline for large-scale manufacturing and distribution to a large enough population for achieving community protection remains uncertain. As a result, prophylactic and/or therapeutic anti-viral drugs are expected to play a role in controlling COVID-19 disease and the ongoing pandemic. Such drugs may be helpful for unvaccinated individuals or those who respond poorly to vaccination, as well as upon waning of immunity and emergence of antigenically distinct variants or strains.

SARS-CoV-2 infects host cells through attachment of the viral transmembrane spike (S) glycoprotein to angiotensin-converting enzyme 2 (ACE2) followed by fusion of the viral and host membranes (Letko et al., 2020; Walls et al., 2020b). SARS-CoV-2 S also engages cell-surface heparan-sulfates (Clausen et al., 2020), neuropilin-1 (Cantuti-Castelvetri et al., 2020; Daly et al., 2020) and L-SIGN/DC-SIGN (Chiodo et al., 2020; Gao et al., 2020; Soh et al., 2020; Thépaut et al., 2020), which were proposed to serve as co-receptors, auxiliary receptors, or adsorption factors. SARS-CoV-2 S is the

main target of neutralizing antibodies (Abs) in infected individuals and the focus of the many nucleic acid, vectored, and protein subunit vaccines currently deployed or in development (Corbett et al., 2020a, 2020b; Erasmus et al., 2020; Hassan et al., 2020; Keech et al., 2020; Mercado et al., 2020; Walls et al., 2020a). Besides blocking ACE2 attachment (Piccoli et al., 2020; Tortorici et al., 2020), some neutralizing Abs are alternatively expected to interfere with heparan-sulfate, neuropilin-1, or L-SIGN/DC-SIGN interactions.

The SARS-CoV-2 S protein comprises an N-terminal S_1 subunit responsible for virus-receptor binding and a C-terminal S_2 subunit that promotes virus-cell membrane fusion (Walls et al., 2020b; Wrapp et al., 2020). The S_1 subunit comprises an N-terminal domain (NTD) and a receptor-binding domain (RBD), also known as domain A and B, respectively (Tortorici and Veesler, 2019). Abs targeting the RBD account for ~90% of the neutralizing activity in COVID-19 convalescent sera (Greaney et al., 2021; Piccoli et al., 2020), and numerous monoclonal antibodies (mAbs) recognizing this domain have been isolated and characterized (Barnes et al., 2020a, 2020b; Baum et al., 2020b; Brouwer et al., 2020; Hansen et al., 2020; Ju et al., 2020; Piccoli et al., 2020; Pinto et al., 2020; Tortorici et al., 2020; Wang et al., 2020; Wu et al., 2020). Several RBD-specific mAbs capable of protecting small animals and non-human primates from SARS-CoV-2 challenge are able to neutralize viral infection by targeting multiple distinct antigenic sites (Baum et al., 2020a; Hansen et al., 2020; Jones et al., 2020; Pinto et al., 2020; Rogers et al., 2020; Tortorici et al., 2020; Zost et al., 2020). A subset of these mAbs is currently being evaluated in clinical trials or have recently received emergency use authorization from the FDA.

The limited immunogenicity of the SARS-CoV-2 NTD in COVID-19 patients (Piccoli et al., 2020; Rogers et al., 2020) has been hypothesized to result from its extensive N-linked glycan shielding (Walls et al., 2020b; Watanabe et al., 2020). However, recent studies have reported the isolation of NTD-targeted mAbs and their ability to neutralize SARS-CoV-2 infection *in vitro*, suggesting they could be useful for COVID-19 prophylaxis or treatment (Chi et al., 2020; Liu et al., 2020a). Although the NTD has been proposed to interact with auxiliary receptors in cell types that do not express ACE2 (e.g., DC-SIGN/L-SIGN), its role and the mechanism of action of NTD-targeted neutralizing mAbs remain unknown (Soh et al., 2020). Understanding the immunogenicity of different S domains and the function of mAbs targeting them, including the NTD, is critical to a full understanding of immunity during the pandemic.

Here, we analyze Ab responses in three COVID-19 convalescent individuals and describe 41 NTD-specific human mAbs. Integrating cryoelectron microscopy (cryo-EM), binding assays, and Ab escape mutant analysis, we define a SARS-CoV-2 NTD antigenic map and identify a supersite (designated site i) recognized by potent neutralizing mAbs. We show that these mAbs exhibit neutralization activities on par with best-in-class RBD-specific mAbs and efficiently activate Fc-mediated effector functions. We also identify immunologically important variation of the SARS-CoV-2 NTD suggesting that the S glycoprotein is under selective pressure from the host humoral immune response. Finally, we provide proof-of-principle that prophylactic administration of a highly potent NTD mAb confers protection of Syrian hamsters against SARS-CoV-2 challenge albeit selecting for escape mutants in some animals. Overall, these findings

highlight the importance of NTD-specific neutralizing mAbs for protective immunity, virus evolution, and vaccine design.

6.2 NTD-specific mAbs with potent neutralizing activity

To discover mAbs targeting diverse SARS-CoV-2 epitopes, we sorted immunoglobulin (Ig)G^+ memory B cells from peripheral blood mononuclear cells (PBMCs) of three COVID-19 convalescent individuals (L, M, X) using biotinylated prefusion SARS-CoV-2 S as a bait. The percentage of SARS-CoV-2 S-reactive IgG^+ memory B cells ranged between 1.1%–1.3%. A total of 278 mAbs were isolated and recombinantly produced as human IgG1 (**Fig. 6.1A**). Characterization by ELISA showed that most mAbs isolated from the three donors recognize the RBD (65%–77%), with a smaller fraction targeting the NTD (6%–20%). The remaining mAbs (4%–20%) are expected to bind to either the S_2 subunit or the C-D domains within the S_1 subunit (**Fig. 6.1A**). The low proportion of NTD-specific mAbs isolated from these donors is in line with the previously observed limited NTD immunogenicity in SARS-CoV-2-exposed individuals (Piccoli et al., 2020; Rogers et al., 2020). Overall, we identified 41 mAbs recognizing the SARS-CoV2 NTD with EC_{50}s ranging between 7.6–698 ng/mL and nanomolar binding affinities, as evaluated using ELISA and biolayer interferometry, respectively (**Fig.s 6.1B and S6.1A; Table S6.1**). These NTD-specific mAbs use a large repertoire of V genes, with an over-representation of IGHV3-21 and IGK3-15 genes (**Fig.s S6.1B and S6.1C; Table S6.1**), and harbor few somatic hypermutations (VH and VL are 97.57% and 97.54% identical to V germline genes, respectively; **Fig. S6.1D; Table S6.1**), as previously described for many SARS-CoV-2 neutralizing mAbs binding to the RBD (Piccoli et al., 2020; Seydoux et al., 2020).

CDRH3 lengths range between 10 and 24 amino acid residues (**Fig. S6.1E**). Collectively, these data indicate that the Ab response to the SARS-CoV-2 NTD is polyclonal.

The *in vitro* neutralization activity of the NTD-specific mAbs was subsequently evaluated using a SARS-CoV-2 S pseudotyped murine leukemia virus (MLV) system (Millet and Whittaker, 2016; Walls et al., 2020b). Out of 41 mAbs, 9 are potent neutralizers (IC_{50} < 50 ng/mL) and 6 are moderate neutralizers (IC_{50} of 50–150 ng/mL) (**Fig. 6.1C**). The remaining 25 mAbs were non-neutralizing. Most of the mAbs plateaued around 80%–90% maximum neutralization in this assay (**Fig. 6.1D**). Evaluation of the neutralization potency of a subset of NTD-specific mAbs measured 6 h post-infection of Vero E6 cells infected with authentic SARS-CoV-2 virus confirmed that these mAbs did not completely block viral entry and instead plateaued at 80%–90% neutralization, as opposed to the RBD-specific mAbs S309, S2E12, and S2M11 that achieved 100% neutralization (**Fig. 6.1E**) (Pinto et al., 2020; Tortorici et al., 2020). When the activity was measured at 24 h post-infection, however, all mAbs tested achieved 95%–100% neutralization with a marked enhancement of neutralization potency (**Fig. 6.1F**). For instance, S2X333 neutralized SARS-CoV-2 with an IC_{50} of 2 ng/mL and an IC_{90} of 12 ng/mL, on par with the best-in-class ultrapotent RBD-targeting mAbs S2E12 and S2M11 (**Fig. 6.1F**). The time-dependent difference of the results may reflect inhibition of extensive viral spread during the 24 h assay as opposed to the sole inhibition of viral entry measured after 6 h.

Previous studies established that SARS-CoV-2 infection of Vero E6 cells proceeds through cathepsin-activated endosomal fusion, as opposed to TMPRSS2-dependent

entry, which is supposed to occur at the level of the plasma membrane and to be the most relevant route of lung cells infection (Hoffmann et al., 2020a, 2020b, 2020c). Although S2L28, S2M28, S2X28, and S2X333 efficiently blocked cell-cell membrane fusion of Vero E6 cells transiently transfected with full-length wild-type SARS-CoV-2 S (**Fig. S6.1F**), binding of S2L28, S2M28, and S2X333 to SARS-CoV-2 S was dampened by 2 orders of magnitude at endosomal pH (pH5) compared to neutral pH (pH7) (**Fig. S6.1G**). As cell-cell membrane fusion bypasses the endosomal compartment, S2L28, S2M28, S2X28, and S2X333 efficiently blocked fusion of Vero E6 cells transiently transfected with full-length wild-type SARS-CoV-2 S (**Fig. S6.1F**). Therefore, partial neutralization may have been a result of reduced blocking of the endosomal entry route at 6 h post-infection, whereas cell-cell spread of the virus after 24 h was efficiently blocked.

6.3 NTD-specific neutralizing mAbs delineate an antigenic supersite

To elucidate the mechanism of potent SARS-CoV-2 neutralization by NTD mAbs, we carried out single-particle cryo-EM analysis of the SARS-CoV-2 S ectodomain trimer bound to one NTD-specific mAb from each donor—S2L28, S2M28 or S2X333—in combination with the RBD-specific mAb S2M11. S2M11 was used as it locks the RBDs in the closed state by recognizing a quaternary epitope spanning two adjacent RBDs, thus enabling the use of 3-fold symmetry during reconstruction (Tortorici et al., 2020). 3D classification of the particle images belonging to each dataset revealed the presence of homogeneous ternary complexes with three S2M11 Fabs bound to the RBDs and three bound NTD Fabs radiating from the trimer periphery. We determined reconstructions at 2.6 Å, 2.5 Å, and 2.2 Å for the S2L28/S2M11/S, S2M28/S2M11/S, and S2X333/S2M11/S complexes (**Fig.s 6.2A–6.2I and S6.2A–S2C; Table S6.2**). We

subsequently used local refinement to account for the pronounced conformational dynamics of S2L28, S2M28, and S2X333 and obtained reconstructions at 2.6–3.0 Å resolution for the region comprising the Fab variable domains and their bound epitope in the NTD (**Fig.s 6.2A–2I and S6.2A–S2C; Table S6.2**). In parallel, we determined a crystal structure of the SARS-CoV-2 NTD in complex with the S2M28 Fab at 3.0 Å resolution revealing several additional ordered loops and N-linked glycans (**Fig.s 6.2E–2H; Table S6.3**).

S2L28, S2M28, and S2X333 contact the N-terminal region (residues 14–20, NTD N terminus), a β-hairpin formed by residues 140–158 (supersite β-hairpin), and a loop spanning residues 245–264 (supersite loop). These three regions collectively form an antigenic supersite on the pinnacle of the NTD on the side distal to the viral membrane (**Figs.s 6.2A–6.2I**). The epitopes targeted by S2L28, S2M28, and S2X333 are overlapping and flanked by the oligosaccharides at position N17 and N149, which are located at opposite sides of the antigenic supersite (**Figs.s 6.2D–6.2I**). As a stunning example of convergent binding, each mAb uses hydrophobic residues at the tip of the HCDR3 loop to contact the NTD supersite near residue R246. For S2L28, S2M28, and S2X333, this hydrophobic residue is W105, I93, and W106, respectively (**Figs.s 6.2G–6.2I**).

Although these Fabs recognize overlapping epitopes within this antigenic supersite, the detailed nature of the interactions formed with the NTD varies. In particular, S2L28 primarily binds to residues in the NTD supersite loop including residue D253 (**Fig. 6.2G**), while S2X333 interacts primarily with the N-terminal region including residue N17 and the supersite β-hairpin including residues E156 and R158 (**Fig. 6.2I**).

S2M28 interacts extensively with all three supersite regions, including residues 146–148 in the supersite β-hairpin, D253 in the supersite loop, and N17 in the N-terminal region (**Fig. 6.2H**). As a result of these distinct interactions, whereas the S2L28-bound NTD adopts an identical conformation to that observed in the apo S structures (Walls et al., 2020b; Wrapp et al., 2020; Wrobel et al., 2020b), S2M28 and S2X333 induce remodeling of the NTD supersite loop and β-hairpin, respectively (**Figs.s 6.2G–6.2I**).

Cryo-EM maps of S2L28-, S2M28-, S2X28-, and S2X333-bound SARS-CoV-2 S delineate an antigenic supersite recognized by potent NTD-specific neutralizing mAbs isolated from COVID-19 individuals (**Figs.s 6.2A–2I and S6.3A**). Likewise, the previously described mAb 4A8 (Chi et al., 2020) as well as the recently described FC05 (Zhang et al., 2020) and CM25 (Voss et al., 2020), bind to the same NTD antigenic supersite to neutralize SARS-CoV-2. Collectively, these results show that NTD-specific neutralizing mAbs have adopted a convergent mechanism of viral inhibition through the use of multiple V genes (**Table S6.1**), which points to a key role of the NTD supersite in viral neutralization and putatively also infection.

6.4 Definition of a SARS-CoV-2 NTD antigenic map

To understand the markedly different neutralization potencies across the 41 NTD-specific mAbs identified, we carried out competition biolayer interferometry binding assays using recombinant SARS-CoV-2 S ectodomain trimer or isolated NTD. The data indicated that the mAbs recognize six distinct antigenic sites, which we designated i, ii, iii, iv, v, and vi. Most mAbs clustered within antigenic sites i and iii whereas sites ii, iv, v, and vi each accounted for only one or a small number of mAbs from our panel (**Figs. 6.3A**). All potently neutralizing mAbs competed for binding to the NTD site i,

indicating they recognize overlapping epitopes within the structurally identified antigenic supersite (**Figs.s 6.1B–6.1E, 6.2, and 6.3A**).

To evaluate the presence of site i-specific mAbs in previously infected or vaccinated individuals, we carried out competition ELISA using S2X28. Plasma from infected individuals and convalescent subjects that received a first dose of mRNA vaccine (convalescent-Vx1), but not naive-vaccinated individuals (Vx1), efficiently competed with S2X28 for binding to the antigenic supersite (site i) (**Table S6.5**). These results confirm that in some individuals, antigenic site i-specific Abs can be elicited as a result of SARS-CoV-2 infection and are possibly boosted upon vaccination despite the dense glycosylation of the NTD.

To delineate the different NTD antigenic sites, we mapped the epitopes targeted by representative mAbs through determination of cryo-EM structures of S2M24, S2X316, or S2L20 in complex with SARS-CoV-2 S, together with S2M11 to expedite structural determination at 6–8 Å resolution (**Figs. S6.3; Table S6.4**). S2L20 (site iv) binds to an epitope flanked by glycans at positions N17, N61, and N234 located within the NTD side, which faces the RBD belonging to the same protomer (**Figs. 6.3B–6.3D**). S2M24 (site vi) recognizes the viral membrane proximal side of the NTD, which is devoid of glycans except for the presence of the N61 oligosaccharide on the edge of the epitope (**Figs. 3B–3D**). S2M24 recognizes a similar antigenic site as the one described for the polyclonal Fabs COV57 isolated directly from COVID-19 patient sera (Barnes et al., 2020b). S2X316 (site v) interacts with an epitope residing nearby glycans N74 and N149 at the peripheral-most part of the NTD, located in between antigenic sites i and vi (**Figs. 3B–3D**). As all NTD-specific mAbs with neutralizing activity target

the structurally identified antigenic supersite site (site i), our data points to the existence of a site of vulnerability that is harnessed by the immune system of multiple individuals.

6.5 NTD mAbs are affected by the genetic diversity of circulating SARS-CoV-2 isolates and sarbecoviruses

Analysis of the 508,771 SARS-CoV-2 genome sequences deposited in GISAID as of February 12, 2021 reveals a larger number of prevalent mutations and deletions in the NTD relative to other regions of the S glycoprotein (**Fig. 6.4A**). We therefore set out to evaluate the impact on NTD-specific mAb binding of several mutations identified within and outside the NTD antigenic supersite along with a few mutants designed based on the structural data. The recombinant mutant NTDs tested include S12F, L18F, T19A (removes the N15 glycan), a 69/70 deletion, D80A, a Y144 deletion, H146Y, N149Q (removes the N149 glycan), A222V, R246A, D253G/Y, S254F, S255F, and the B1.351 NTD (comprising the L18F, D80A, D215G, 242–244 deletion, and R246I mutations).

All neutralizing NTD mAbs bound efficiently to NTD variants carrying either a 69/70 deletion or the A222V substitutions, indicating the site i-targeting mAbs would not be affected by these common SARS-CoV-2 mutations (Hodcroft et al., 2020) (**Figs. 6.4A and 6.4B**). The R246A substitution decreased binding of S2L28, S2M28, and S2X333 (to various extents), in agreement with the extensive interactions formed by the R246 side chain with aromatic and hydrophobic residues found in the HCDR3 of each mAb. This substitution also markedly affected the binding of S2X28 and of the previously described mAb 4A8 (Chi et al., 2020) (**Fig. 6.4B**). The L18F, D80A, D253G/Y, or S255F variants only abrogated binding of S2L28 to the NTD, which is relevant considering the L18F substitution is one of the most prevalent mutants

sequenced to date and is present in both the B.1.351 and P.1 lineages (Tegally et al., 2020) (**Fig. 6.4B**). Conversely, the Y144 deletion abrogated binding to S2M28, S2X28, S2X333, and 4A8, but not S2L28, explaining the lost ability of these mAbs to neutralize the B.1.1.7 S pseudovirus, which contains this deletion (Collier et al., 2021; Davies et al., 2020) (**Fig. 6.4B**). The H146Y mutant reduced binding of S2M28, S2X28, and in particular of 4A8 (**Fig. 6.4A**). Binding of all site i-specific NTD mAbs to the B.1.351 NTD was abrogated, in agreement with a recent study showing that 4A8 did not recognize this NTD variant (Wibmer et al., 2021). However, binding of the non-neutralizing S2L20 NTD-specific mAb was not affected by any of these variants, confirming retention of proper folding of the purified NTD mutants (**Fig. 6.4A**). Collectively these data show that S2M28, S2X28, S2X333, and 4A8 have a similar reactivity profile to these NTD mutants, whereas S2L28 has a distinct and complementary pattern. These data demonstrate that several currently circulating variants will partially or completely escape neutralization mediated by mAbs targeting the antigenic supersite (site i), including the B.1.1.7, B.1.351, and P.1 lineages.

To investigate the cross-reactivity of NTD-specific mAbs with sarbecovirus S glycoproteins, we probed individual mAb binding to ExpiCHO cells expressing full-length S from representative members of the three clades of the sarbecovirus subgenus. All mAbs tested bound to RaTG13 S, which is the most closely related S glycoprotein and NTD to SARS-CoV-2 S (**Fig. 6.4C**). S2X333 and S2M28 weakly bound to S from the Pangolin Guanxi 2017 (P-GX) isolate, whereas S2L28 and S2M28 weakly bound to S from the Pangolin Guangdong 2019 (P-GD) isolate (**Fig. 6.4C**). Consistently, weak mAb binding to P-GD S was confirmed by ELISA for S2L28, S2M28, S2X333, and S2M24

(**Fig. S6.4A).** We next evaluated the neutralization breadth of S2L28, S2M28, S2X28, and S2X333 against SARS-CoV-2- and SARS-CoV-related sarbecoviruses using a vesicular stomatitis virus (VSV) pseudotyped neutralization assay. All mAbs tested neutralized RaTG13 S pseudotypes with comparable potency to SARS-CoV-2 S pseudotyped virus but none of them neutralized the more distantly related P-GX S, P-GD S, SARS-CoV S, WIV1 S, or WIV16 S pseudotyped viruses (**Figs. S6.4B–S6.4H**). These findings are in line with the low conservation of the NTD supersite (site i) in phylogenetically distant viruses and indicate that NTD mAbs could be useful countermeasures against a subset of viruses closely related to SARS-CoV-2 but not against divergent viruses. Both species-specific functions and host immunity may play a role in the antigenic diversification of this region compared to other S domains that are more conserved.

6.6 SARS-CoV-2 S neutralization escape mutants reveal unconventional mechanisms of NTD mAb resistance

To investigate potential viral escape from neutralization mediated by the potent neutralizing NTD mAbs, we passaged a replication competent VSV-SARS-CoV-2 chimera in which the native glycoprotein was replaced by the SARS-CoV-2 S Wuhan-1 isolate gene in the presence of S2L28, S2M28, S2X28, and S2X333, as previously described (Case et al., 2020; Liu et al., 2020b). Selective pressure provided by each mAb led to emergence of several escape mutants that clustered within the antigenic supersite i.e., the NTD N terminus, the supersite β-hairpin, and the supersite loop as well as surrounding residues (**Figs. 6.5A and 6.5B**). Similar to what was previously described for RBD mAbs (Starr et al., 2020a, 2020b), all viral escape mutants resulted from a single non-synonymous nucleotide mutation.

The four mAbs selected for many shared escape mutations but with a distinct pattern for S2L28, in agreement with the structural observations indicating distinct residues are important for recognition of this antibody. For instance, S2L28 selected for and was uniquely sensitive to the D80N and D253G mutations, whereas it still bound to the K147T mutation, which abrogated binding by S2M28, S2X28, S2X333, and 4A8. S2X28 selected for Y144C/N substitutions, both S2M28 and S2X28 selected for R246G substitutions, and all four mAbs selected for L18P substitutions (**Fig. 6.5A**). Mutations including L18F, D80A, deletion of residue Y144, R246I/T/K, and D253G/Y have been detected in SARS-CoV-2 clinical isolates, most of them in the B.1.1.7, B.1.351, or P.1 lineages, demonstrating the ability of this system to recapitulate variants at naturally occurring S residues (Davies et al., 2020; Tegally et al., 2020).

Besides residue substitutions that are expected to impact the epitope/paratope interface directly, we observed escape mutants in the signal peptide region and at cysteine residues 15 and 136. The latter two residues are jointly engaged in a disulfide bond that staples the NTD N terminus against the galectin-like β sandwich (**Figs. 6.2G–6.2I**). Alteration of this bond (via C15S or C136Y substitutions) abrogates recognition by S2L28, S2M28, S2X28, S2X333, and 4A8 (**Fig. 6.5B**), likely via dislodging this key supersite component, explaining the observed phenotype. These data further suggest that this mechanism of escape is likely to be effective with all mAbs targeting the NTD antigenic supersite.

The predicted SARS-CoV-2 S signal peptide cleavage site occurs between residues T13 and Q14. Signal peptide cleavage site prediction using SignalP5.0 (Almagro Armenteros et al., 2019) suggests that the selected S12P escape mutation

alters cleavage to occur between residues C15 and V16, thereby eliminating the C15-C136 disulfide bond—similarly to escape mutants at positions C15 and C136. To test the effect of the S12P substitution with purified NTD, we introduced the native signal peptide in the presence or absence of S12P. Similar to direct disruption of the C15-C136 disulfide bond, the S12P signal peptide mutation was found to markedly reduce binding to S2L28, S2M28, S2X28, S2X333, and 4A8 (**Fig. 6.5B**). Mass spectrometry analysis of the S12P mutant was consistent with signal peptide cleavage occurring immediately after residue C15 (**Figs. 6.5C–6.5F**). The S12F mutation observed in clinical isolates, however, did not affect binding of the NTD-specific mAbs tested (**Fig. 6.4B**). Collectively these findings indicate the involvement of the S signal peptide in immune evasion and warrant closer monitoring of NTD variants.

6.7 Mechanism of action of NTD-specific neutralizing mAbs

To understand NTD-specific mAb-mediated inhibition of viral entry, we evaluated the ability of these mAbs to block ACE2 binding as this step correlates with neutralization titers in SARS-CoV-2-exposed individuals (Piccoli et al., 2020). None of the site i-targeting mAbs (S2L28, S2M28, S2X28, and S2X333) blocked binding of SARS-CoV-2 S to immobilized human recombinant ACE2 as measured by biolayer interferometry (**Fig. 6.6A**), excluding interference with engagement of the main entry receptor as mechanism of action. Moreover, these mAbs did not promote shedding of the S_1 subunit from cell-surface-expressed full-length SARS-CoV-2 S (**Fig. 6.6B**), ruling out the possibility of premature S triggering, as previously shown for a SARS-CoV and several SARS-CoV-2 RBD-specific mAbs (Huo et al., 2020; Piccoli et al., 2020; Walls et al., 2019; Wec et al., 2020; Wrobel et al., 2020a). However, S2L28, S2M28, S2X28, and S2X333 blocked SARS-CoV-2 S-mediated cell-cell fusion, indicating they could

prevent interaction with an auxiliary receptor, proteolytic activation, or membrane fusion (**Fig. S6.1G**).

To assess whether steric hindrance plays a role in the neutralization activity of S2L28, S2M28, S2X28, and S2X333, we evaluated the neutralization potency of each mAb, in both Fab and IgG formats, against authentic SARS-CoV-2-Nluc (**Fig. 6.6C**). NTD-specific Fabs displayed a marked potency reduction, both in terms of IC_{50} values and maximal neutralization plateau reached (**Table S6.1**), as compared to IgGs, possibly due to reduced avidity as observed by surface plasmon resonance (**Fig. 6D**). Since the Fabs could still partially neutralize SARS-CoV-2, at least part of the observed neutralization activity results from direct interaction with their respective epitopes. It is possible that NTD-specific mAb-mediated neutralization further relies on the steric hindrance provided by Fc positioning, similar to what was observed for anti-hemagglutinin influenza A virus neutralizing mAbs (Xiong et al., 2015).

We next examined potential additive, antagonistic, or synergistic effects of NTD- and RBD-targeting mAbs, as mAb synergy was previously described for SARS-CoV and SARS-CoV-2 neutralization (Pinto et al., 2020; ter Meulen et al., 2006). Cocktails of S2X333 with S309, S2E12, or S2M11 additively prevented entry of SARS-CoV-2 S-MLV pseudotyped virus in Vero E6 cells **(Figs. S6.5A–S6.5C**). This additive effect was also observed between S2X333 and S309 using authentic SARS-CoV-2 at 24 h post-infection in Vero E6 cells (**Fig. S6.5D**). These results are consistent with RBD- and NTD-targeting mAbs mediating inhibition by distinct mechanisms and demonstrate that they could be used as cocktails for prophylaxis or therapy.

Since Fc-mediated effector functions contribute to protection by promoting viral clearance and anti-viral immune responses *in vivo* (Bournazos et al., 2020, 2016; Schäfer et al., 2021; Winkler et al., 2020), we evaluated the ability of site i-targeting mAbs to trigger activation of FcγRIIa and FcγRIIIa as a proxy for Ab-dependent cellular phagocytosis (ADCP) and Ab-dependent cellular cytotoxicity (ADCC), respectively. S2L28, S2M28, S2X28, and S2X333 promoted dose-dependent FcγRIIa- and FcγRIIIa-mediated signaling to levels comparable to those of the highly effective mAb S309 (Pinto et al., 2020) **(Figs. 6.6E and 6.6F**). In contrast, the non-neutralizing site vi-targeting S2M24 mAb did not promote FcγR-mediated signaling, possibly due to the different orientation relative to the membrane of the effector cells in comparison to site i-specific mAbs (**Figs. 6.6E and 6.6F**). These findings suggest that besides their neutralizing activity, mAbs recognizing site i can exert a protective activity via promoting Fc-mediated effector functions.

6.8 NTD neutralizing mAbs protect against SARS-CoV-2 challenge in hamsters

To evaluate the protective efficacy of NTD-directed mAbs against SARS-CoV-2 challenge *in vivo*, we selected the S2X333 mAb for a prophylactic study in a Syrian hamster model (Boudewijns et al., 2020). The mAb was administered at 4 and 1 mg/kg via intraperitoneal injection 48 h before intranasal SARS-CoV-2 challenge. Four days later, lungs were collected for the quantification of viral RNA and infectious virus titers. Prophylactic administration of S2X333 at a dose of 4 mg/kg decreased the amount of viral RNA detected in the lungs by ~3 orders of magnitude, compared to hamsters receiving a control mAb (**Fig. 6.7A**). Viral replication in the lungs was completely abrogated for most of the animals at both doses tested (**Fig. 6.7B**). However, although all animals had similar serum mAb concentrations within each group, infectious viral

titers were not decreased relative to the control group for 1 and 4 hamsters administered with S2X333 at 1 and 4 mg/kg, respectively (**Figs. 6.7C and 6.7D).** Based on the aforementioned variability and mutation tolerance of the SARS-CoV-2 NTD, we hypothesized that S2X333 escape mutants might have been selected in these animals. Sequencing of the viral RNA extracted from the lungs of one outlier from each group administered with S2X333 revealed an S deletion encompassing position Y144 (within site i) in comparison to the input virus or the animal from the control group. The deletion corresponds to 4 and 2 amino acids for the animals treated with 1 ($LGVY_{144}$) and 4 mg/kg (VY_{144}), respectively (**Fig. S6.6**). These results show that potent NTD mAbs targeting the antigenic supersite (site i) impose a selection pressure driving viral evolution to escape neutralization. Overall, these data show that low doses of anti-NTD mAbs provide remarkable prophylactic activity *in vivo*, comparable to best-in-class RBD-specific mAbs S2E12 and S2M11 (Tortorici et al., 2020), consistent with their potent *in vitro*neutralizing activity. We anticipate that the protection efficacy of S2X333 (and related NTD mAbs) would be further enhanced in humans due to an optimal matching with human Fcγ receptors. However, our results also indicate that potent NTD mAbs targeting the antigenic supersite (site i) impose a selection pressure driving viral evolution to escape neutralization.

6.9 Discussion

The data herein demonstrate that neutralizing NTD-targeting mAbs represent a key aspect of immunity to SARS-CoV-2 and account for 5%–20% of SARS-CoV-2 S-specific mAbs cloned from memory B cells isolated from the PBMCs of three COVID-19 individuals. Analysis of a large panel of neutralizing and non-neutralizing mAbs led us to define an antigenic map of the heavily glycosylated SARS-CoV-2 NTD, in which six

antigenic sites (i–vi) were identified. We found that all the neutralizing mAbs from the three donors investigated, along with the mAbs 4A8 (Chi et al., 2020), FC05 (Zhang et al., 2020), and CM25 (Voss et al., 2020), target the same antigenic supersite (site i) using various germline V genes, thereby providing examples of convergent solutions to NTD-targeted mAb neutralization. We show that a highly potent NTD mAb provides prophylactic protection against SARS-CoV-2 challenge of Syrian hamsters showcasing that this class of mAbs can be a critical barrier to infection and are promising clinical candidates against SARS-CoV-2. Likewise, previous work on Middle East respiratory syndrome coronavirus (MERS-CoV) established that the NTD is the target of several potent and protective neutralizing mAbs and an attractive target for vaccine design (Chen et al., 2017; Jiaming et al., 2017; Wang et al., 2018, 2019; Zhou et al., 2019).

Our cryo-EM structures of SARS-CoV-2 S bound to the ultrapotent RBD-specific mAb S2M11 and to S2X333, S2M28, or S2L28, together with evidence of additive neutralizing effect of RBD- and NTD-targeted mAbs, provide proof of concept for implementing countermeasures using mAb cocktails targeting both the NTD and the RBD and potentially also for vaccine design. As several examples of single amino acid mutations reducing or completely abrogating neutralization by immune sera have been reported (Li et al., 2020; Liu et al., 2020b; Weisblum et al., 2020), we propose that combinations of mAbs targeting distinct domains might reduce the likelihood of emergence of escape mutants, as previously demonstrated for MERS-CoV (Wang et al., 2018).

Here, we show that site i-targeting NTD neutralizing mAbs efficiently activate FcγRIIa and FcγRIIIa *in vitro*. Fc-mediated effector functions can be affected by the

epitope specificity of the mAbs (Piccoli et al., 2020), highlighting the importance of the orientation of the S-bound Fc fragments for efficient FcγR cross-linking and engagement. The site vi-targeting NTD mAb S2M24 did not activate either FcγRIIa or FcγRIIIa, presumably due to the different angle of approach. The contribution of Fc-mediated effector functions could further enhance the prophylactic activity of potent NTD-specific mAbs against SARS-CoV-2 in humans as demonstrated for two RBD-specific SARS-CoV-2 neutralizing mAbs (Schäfer et al., 2021; Winkler et al., 2020). The emergence and fixation of SARS-CoV-2 D614G yielded a virus mutant with enhanced ACE2 binding, transmissibility, and replication but without marked changes of antigenicity (Hou et al., 2020; Korber et al., 2020; Plante et al., 2020; Yurkovetskiy et al., 2020). In contrast, we observed that several variants detected in circulating SARS-CoV-2 clinical isolates reduced or abrogated recognition by some NTD-specific neutralizing mAbs. For instance, the L18F substitution, which has a prevalence of 10% among sequenced SARS-CoV-2 genomes and is present in the B.1.351 and P.1 lineages initially identified in South Africa and Brazil (Tegally et al., 2020), would escape S2L28-mediated neutralization. The deletion of residue Y144 found in 13% of circulating SARS-CoV-2 clinical isolates, including the B.1.1.7 lineage originally detected in the UK before being identified in numerous other countries (Davies et al., 2020), would circumvent neutralization by S2M28, S2X28, S2X333, and 4A8 (Collier et al., 2021).

Using selection of viral escape mutants and structural analysis, we identified additional key residues for binding of neutralizing mAbs, such as S12 (indirectly) and R246, for which natural variants are reported and are thus expected to participate in neutralization escape. The finding that multiple circulating SARS-CoV-2 variants map to

the NTD, including several of them in the antigenic supersite (site i), suggests that the NTD is under selective pressure from the host humoral immune response. This is further supported by the identification of NTD deletions within the antigenic supersite (site i) in immunocompromised hosts with prolonged infections (Avanzato et al., 2020; Choi et al., 2020; McCarthy et al., 2021), the *in vitro* selection of SARS-CoV-2 S escape variants with NTD mutations that decrease the neutralization potency of COVID-19 convalescent patient sera (Andreano et al., 2020; Weisblum et al., 2020), and the escape mutants reported here. Collectively, these data indicate that the mutations and deletions accumulated in the NTDs of the B.1.1.7, B.1.351, and P.1 lineages are a direct result of mAb-mediated selective pressure in infected hosts. We propose that mAb cocktails comprising S2L28 plus S2M28, S2X28, or S2X333 could provide broad protection against some SARS-CoV-2 variants as these mAbs are susceptible to distinct sets of SARS-CoV-2 S mutations, as previously proposed for RBD-specific mAbs (Greaney et al., 2020).

6.10 Methods

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
E. coli DH10B Competent Cells	ThermoFisher Scientific	Cat# EC0113
Deposited Data		
S + S2M11 + S2X333 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23579
S + S2M11 + S2X333 Cryo-EM Model	https://www.rcsb.org	PDB 7LXY
S + S2M11 + S2X333 Focused Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23577
S + S2M11 + S2X333 Focused Cryo-EM Model	https://www.rcsb.org	PDB 7LXW
S + S2M11 + S2M28 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23582
S + S2M11 + S2M28 Cryo-EM Model	https://www.rcsb.org	PDB 7LY2
S + S2M11 + S2M28 Focused Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23581
S + S2M11 + S2M28 Focused Cryo-EM Model	https://www.rcsb.org	PDB 7LY0
S + S2M11 + S2L28 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23580
S + S2M11 + S2L28 Cry-oEM Model	https://www.rcsb.org	PDB 7LXZ
S + S2M11 + S2L28 Focused Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23578
S + S2M11 + S2L28 Focused Cryo-EM Model	https://www.rcsb.org	PDB 7LXX
2P-DS-S + S2X28 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23584
S + S2M11 + S2L20 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23586
S + S2M11 + S2X316 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23585
S + S2M11 + S2M24 Cryo-EM Map	https://www.ebi.ac.uk/pdbe/emdb/	EMD-23583

REAGENT or RESOURCE	SOURCE	IDENTIFIER
S NTD + S2M28 Crystal Structure	https://www.rcsb.org	PDB 7LY3
Experimental Models: Cell Lines		
FreeStyle™ 293-F Cells	ThermoFisher Scientific	Cat# R79007
Expi293F™ Cells	ThermoFisher Scientific	Cat# A14527
HEK293T	ATCC	Cat# CRL-11268
Vero-E6	ATCC	CRL-1586
Expi-CHO cells	ThermoFisher Scientific	Cat# A29127
Expi-CHO cells stably expressing SARS-CoV-2	This study	N/A
Jurkat cells stably expressing FcγRIIIa receptor (V158 variant) and NFAT-driven luciferase gene (effector cells)	Promega	Cat# G7018
Jurkat cells stably expressing FcγRIIa receptor (H131 variant) and NFAT-driven luciferase gene	Promega	Cat# G9995
Oligonucleotides		
NTD_fwd: TATAATGGTACCGCCACCATGGGCAT	Integrated DNA Technologies, Inc.	N/A
NTD_rev: TATAATAAGCTTTCAGTGGTGGTGGTGGTGGTGAT GATGCGCGGTAAAGCTCTTCGCTGTACACTT	Integrated DNA Technologies, Inc.	N/A
N149Q_fwd: AAAAGTCCTGGATGGAGTCTGAGT	Integrated DNA Technologies, Inc.	N/A
N149Q_rev: GGTTCTTGTGATAGTACACGCCC	Integrated DNA Technologies, Inc.	N/A
D253G_fwd: GCTCCTCTAGCGGATGGAC	Integrated DNA Technologies, Inc.	N/A
D253G_rev: CGCCTGGTGTCAGGTAGG	Integrated DNA Technologies, Inc.	N/A
D253Y_fwd: ACTCCTCTAGCGGATGGAC	Integrated DNA Technologies, Inc.	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
D253Y_rev: AGCCTGGTGTCAGGTAGG	Integrated DNA Technologies, Inc.	N/A
T19A_fwd: ACAAGGACCCAGCTGCC	Integrated DNA Technologies, Inc.	N/A
T19A_rev: GGCCAGGTTCACGCACTG	Integrated DNA Technologies, Inc.	N/A
R246A_fwd: CATCCTACCTGACACCAGGC	Integrated DNA Technologies, Inc.	N/A
R246A_rev: CGTGCAGGGCCAGCAG	Integrated DNA Technologies, Inc.	N/A
L18F_fwd: CTTTACCACAAGGACCCAGC	Integrated DNA Technologies, Inc.	N/A
L18F_rev: TTCACGCACTGTGTGCC	Integrated DNA Technologies, Inc.	N/A
H146Y_fwd: ACAAGAACAATAAGTCCTGGATGGAGTC	Integrated DNA Technologies, Inc.	N/A
H146Y_rev: AATAGTACACGCCCAGGAATGGAT	Integrated DNA Technologies, Inc.	N/A
A222V_fwd: TCCTGGAGCCACTGGTG	Integrated DNA Technologies, Inc.	N/A
A222V_rev: CGGAGAATCCCTGTGGCAG	Integrated DNA Technologies, Inc.	N/A
SP_rev: AGGACGAGGAAGACGAACATGGTGGCGGTACCTCCG	Integrated DNA Technologies, Inc.	N/A
SP_fwd: GCTCCCTCTGGTTTCTTCCCAGTGCGTGAACCTGACCA	Integrated DNA Technologies, Inc.	N/A
SP_S12P_fwd: GCTCCCTCTGGTTCCATCCCAGTGCGTGAACCTGACCA	Integrated DNA Technologies, Inc.	N/A
Y144del_fwd: CACAAGAACAATAAGTCCTGGATGGAGT	Integrated DNA Technologies, Inc.	N/A
Y144del_rev: GTACACGCCCAGGAATGGATC	Integrated DNA Technologies, Inc.	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
S254F_fwd: CTCTAGCGGATGGACCGCA	Integrated DNA Technologies, Inc.	N/A
S254F_rev: AAGTCGCCTGGTGTCAGGTAG	Integrated DNA Technologies, Inc.	N/A
K147T_fwd: CGAACAATAAGTCCTGGATGGAGTCT	Integrated DNA Technologies, Inc.	N/A
K147T_rev: TGTGATAGTACACGCCCAGGAAT	Integrated DNA Technologies, Inc.	N/A
C136Y_fwd: ATAATGATCCATTCCTGGGCGTGTA	Integrated DNA Technologies, Inc.	N/A
C136Y_rev: AAAACTGGAACTCGCACACCTTG	Integrated DNA Technologies, Inc.	N/A
L19P_fwd: GGGTTCACGCACTGTG	Integrated DNA Technologies, Inc.	N/A
L19P_rev: GACCACAAGGACCCAGCT	Integrated DNA Technologies, Inc.	N/A
C15S_fwd: CICTGTGTGCCGGTCTCTG	Integrated DNA Technologies, Inc.	N/A
C15S_rev: CGTGAACCTGACCACAAGGAC	Integrated DNA Technologies, Inc.	N/A
G142D_fwd: CCAGGAATGGATCATTACAAAACTGGAAC	Integrated DNA Technologies, Inc.	N/A
G142D_rev: ACGTGTACTATCACAAGAACAATAAGTCCTGG	Integrated DNA Technologies, Inc.	N/A
S255F_fwd: AAGGAGTCGCCTGGTGTCA	Integrated DNA Technologies, Inc.	N/A
S255F_rev: TAGCGGATGGACCGCAG	Integrated DNA Technologies, Inc.	N/A
S12F_fwd: TTCCCAGTGCGTGAACCTGA	Integrated DNA Technologies, Inc.	N/A
S12F_rev: AAAACCAGAGGGAGCAGGACG	Integrated DNA Technologies, Inc.	N/A
D80N_fwd: CAATCCCGTGCTGCCTTTTAAC	Integrated DNA Technologies, Inc.	N/A
D80N_rev: TTGAACCGCTTTGTGCCATTG	Integrated DNA Technologies, Inc.	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
D80A_fwd: CCAATCCCGTGCTGCCTTTTAAC	Integrated DNA Technologies, Inc.	N/A
D80A_rev: CGAACCGCTTTGTGCCATTG	Integrated DNA Technologies, Inc.	N/A
69/70del_2fwd: TCCGGCACCAATGGCACA	Integrated DNA Technologies, Inc.	N/A
69/70del_2rev: GATGGCGTGGAACCAGGTCAC	Integrated DNA Technologies, Inc.	N/A
Spike-cDNA-R: CAAGAATCTCAAGTGTCTGTGG	Microsynth AG	N/A
Spike-SEQ1-F: CACTAGTCTCTAGTCAGTGTG	Microsynth AG	N/A
Spike-SEQ1-R: GACTCAGTAAGAACACCTGTGC	Microsynth AG	N/A
Spike-SEQ2-F: AGTGCGTGATCTCCCTCAGG	Microsynth AG	N/A
Spike-SEQ3-F: AGTCAGACAAATCGCTCCAGG	Microsynth AG	N/A
Spike-SEQ-NTD-R: AGCACCAGCTGTCCAACCTG	Microsynth AG	N/A
Spike-SEQ2-NTD-R: CATAAGAAAAGGCTGAGAGA	Microsynth AG	N/A
Recombinant DNA		
pCMV::SARS-CoV-2_S_ecto_hexapro	Ref: PMID: https://.ncbi.nlm.nih.gov/32577660/	N/A
pCMV::SARS-CoV-2_S_ecto_2P_DS	Ref PMID: https://.ncbi.nlm.nih.gov/32753755/	N/A
pCMV::SARS-CoV-2_S_NTD	This study	N/A
pCMV::P-GD_S_ecto	This study	N/A
pCMV::SARS-CoV-2_S_ecto_avi	PMID: https://.ncbi.nlm.nih.gov/32972994/	N/A
pcDNA3.1::SARS-CoV-2-S-D19	(Ou et al., 2020)	N/A
phCMV1::SARS-CoV-2	PMID: https://.ncbi.nlm.nih.gov/32972994/	N/A
phCMV1::RatG13_S	This study	N/A
phCMV1::P-GX_S	This study	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
phCMV1::P-GD_S	This study	N/A
phCMV1::ZC45_S	This study	N/A
phCMV1::ZXC21_S	This study	N/A
phCMV1::YN2013_S	This study	N/A
phCMV1::SARS-CoV_S	This study?	N/A
phCMV1::RmYN02_S	This study	N/A
phCMV1::Bt-kY72_S	This study	N/A
phCMV1::BM48-31_S	This study	N/A
pSARS-CoV-2-Nluc	(Xie et al., 2020)	N/A
Biological Samples		
PE-Cy7 anti-CD19 antibody	BD Bioscience	Cat# 341113
Anti-PE MicroBeads	Miltenyi Biotec	Cat# 130-048-801
LS separation cloumns	Miltenyi Biotec	Cat# 130-042-401
PE anti-human IgM	Lucerna-Chem AG	Cat# 314508
PE anti-human IgD	BD Biosciences	Cat# 555779
PE anti-human IgA	Southern Biotech	Cat# 2050-09
PE anti-human CD14	BD Biosciences	Cat# 562691
PEI MAX	Polysciences	Cat# POL24765-1
4-Nitrophenyl phosphate disodium salt hexahydrate (pNPP)	Sigma-Aldrich	Cat# N2765-100TAB
Tween 20	Sigma-Aldrich	Cat# 93773
Bovine Serum Albumine (BSA)	Sigma-Aldrich	Cat# 3059
Blocker Casein	ThermoFisher Scientific	Cat# 37528
EZ-Link™ NHS-PEG solid phase biotinylation kit	ThermoFisher Scientific	Cat# 21450
Zeba ™ Spin Desalting Columns	ThermoFisher Scientific	Cat# 89892
TPCK treated Trypsin	Worthington Biochem	Cat# LS003750
Alexa Fluor 647-AffiniPure F(ab')2 Fragment Goat Anti-Human IgG, Fcg Fragment Specific	Jackson ImmunoResearch	Cat# 109-606-098
Goat Anti-Human IgG-AP	Southern Biotech	Cat# 2040-04
Streptavidin, Alexa Fluor 647 conjugate	Life technologies	Cat# S21374
Streptavidin, alkaline-phospatase conjugate	Jackson ImmunoResearch	Cat# 016-050-084

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bio-Glo™ Luciferase Assay System	Promega	Cat# G7940
ExpiFectamine™ CHO Transfection Kit	Life technologies	Cat# A29130
ADCC Reporter Bioassays, F Variant	Promega	Cat# G9798
ADCC Reporter Bioassays, V Variant	Promega	Cat# G7018
FcgammaRIIa-H ADCP Reporter Bioassay	Promega	Cat# G9995
10% Igepal CA-630	Sigma-Aldrich	Cat# I8896
25 mM dNTPs	Cytiva	Cat# 28406560
RNase OUT	Life Technologies	Cat# 10777-019
SupeScript III RT	Life Technologies	Cat# 18080-044
DNase/RNase free water	Life Technologies	Cat# 10977-035
10 mM dNTPs mix	Cytiva	Cat# 28406564
Q5 high fidelity DNA polymerase	Bioconcept	Cat# M0493S
AgeI-HF	New England Biolabs	Cat# R3552
SalI-HF	New England Biolabs	Cat# R3138
XhoI	New England Biolabs	Cat# R0146
BsiWI	New England Biolabs	Cat# R3553
NotI	New England Biolabs	Cat# R0189
KpnI	New England Biolabs	Cat# R0142
Cutsmart	New England Biolabs	Cat# B7204S
T4 DNA Ligase	New England Biolabs	Cat# M0202
T4 Ligase buffer	New England Biolabs	Cat# B0202S
RNeasy Mini Kit	QIAGEN	Cat# 74104
OneStep RT-PCR Kit	QIAGEN	Cat# 210210
iTaq Universal One-Step RT-qPCR Kits	Biorad	Cat# 1725150

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NucleoSpin Plasmid, Mini kit for plasmid DNA	Macherey-Nagel	Cat# 740588.50
GFX PCR DNA and Gel Band Purification Kit	Cytiva	Cat# 28903470
Software and Algorithms		
cryoSPARC v3.0.1	(Punjani et al., 2017)	https://cryosparc.com
Relion v3.0	(Zivanov et al., 2018)	https://www3.mrc-lmb.cam.ac.uk/relion
Coot	(Casañal et al., 2019)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Phenix-Refine	(Adams et al., 2010)	https://www.phenix-online.org/download/
Phenix-Phaser	(McCoy et al., 2007)	https://www.phenix-online.org/download/
XDS	(Kabsch, 2010)	http://xds.mpimf-heidelberg.mpg.de
Prism 8	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Biacore Evaluation software	Cytiva	http://www.cytivalifesciences.com/en/us/shop/protein-analysis/spr-label-free-analysis/software/biacore-insight-evaluation-software-p-23528

Cell lines

Cell lines used in this study were obtained from ATCC (HEK293T and Vero-E6) or ThermoFisher Scientific (Expi CHO cells, FreeStyle 293-F cells and Expi293F cells).

Sample donors

Samples from three SARS-CoV-2 recovered individuals, designates as donors L (male, 34 years-old), M (female, 50 years-old), and X (male, 52 years-old), were obtained under study protocols approved by the local Institutional Review Boards

(Canton Ticino Ethics Committee, Switzerland, the Ethical committee of Luigi Sacco Hospital, Milan, Italy). All donors provided written informed consent for the use of blood and blood components (such as PBMCs, sera or plasma). Samples were collected 14 and 52 days after symptoms onset for donor L and M, respectively. Blood drawn from donor X was obtained at day 36, 48, 75 and 125 after symptoms onset.

Cloning and mutant generation

SARS-CoV-2 NTD was sub-cloned with *E. coli* DH10B Competent Cells into pCMV using primers NTD_fwd and NTD_rev. The resulting construct was mutated by PCR mutagenesis to generate C15S, 69/70del, D80A/N, G142D, N149Q, D253G/Y, T19A, R246A, L18F/P, H146Y, A222V, Y144del, S254F, S255F, K147T, C136Y, and the NTD construct with native signal peptide with and without S12P/F, using the eponymously named primers (Key Resources Table). The genes encoding for the Sarbecovirus S proteins tested were cloned in the phCMV1 or pcDNA.3 vectors, and the gene for the C-terminally his-tagged ectodomain of P-GD S was cloned into pCMV (Key Resources Table). Plasmid sequences were verified by Genewiz sequencing facilities (Brooks Life Sciences).

Recombinant ectodomains production

All SARS-CoV-2 S spike ectodomains were produced in 500 mL cultures of FreeStyle 293-F cells (ThermoFisher Scientific) grown in suspension using FreeStyle 293 expression medium (ThermoFisher Scientific) at 37°C in a humidified 8% CO2 incubator rotating at 130 rpm Cells grown to a density of 2.5 million cells per mL were transfected using PEI (9 µg/mL) and pCMV::SARS-CoV-2_S_ecto_hexapro, pCMV::SARS-CoV-2_S_ecto_2P_DS, pCMV::P-GD_S_ecto, pCMV::SARS-CoV-

2_S_ecto_avi, pCMV::SARS-CoV-2_S_D614G_ecto_avi and cultivated for 4 days. The supernatant was harvested and cells were resuspended for another three days, yielding two harvests. S ectodomains were purified from clarified supernatants using a Cobalt affinity column (Cytiva, HiTrap TALON crude), washing with 20 column volumes of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl and eluted with a gradient of 600 mM imidazole. The same protocol was followed for P-GD spike ectodomain purification, except that 25 mM sodium phosphate pH 7 and 300 mM sodium chloride were used instead of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl. At this stage, SARS-CoV-2 S with the avi tag (from pCMV::SARS-CoV-2_S_ecto_avi) was biotinylated (BirA biotin-protein ligase standard reaction kit, Avidity) and further purified by size exclusion chromatography (Superose6, GE Healthcare). All purified proteins were then concentrated using a 100 kDa centrifugal filter (Amicon Ultra 0.5 mL centrifugal filters, MilliporeSigma), residual imidazole was washed away by consecutive dilutions in the centrifugal filter unit with 20 mM Tris-HCl pH 8.0 and 150 mM NaCl, and finally concentrated to 5 mg/mL and flash frozen.

All SARS-CoV-2 S NTD domain constructs (residues 14-307) with a C-terminal 8XHis-tag were produced in 100 mL culture of Expi293F Cells (ThermoFisher Scientific) grown in suspension using Expi293 Expression Medium (ThermoFisher Scientific) at 37°C in a humidified 8% CO2 incubator rotating at 130 rpm) (Walls et al., 2020b). Cells grown to a density of 3 million cells per mL were transfected using pCMV::SARS-CoV-2_S_NTD derivative mutants with the ExpiFectamine 293 Transfection Kit (ThermoFisher Scientific) with and cultivated for five days at which point the supernatant was harvested. His-tagged NTD domain constructs were purified from clarified

supernatants using 2 mL of cobalt resin (Takara Bio TALON), washing with 50 column volumes of 20 mM HEPES-HCl pH 8.0 and 150 mM NaCl and eluted with 600 mM imidazole. Purified protein was concentrated using a 30 kDa centrifugal filter (Amicon Ultra 0.5 mL centrifugal filters, MilliporeSigma), the imidazole was washed away by consecutive dilutions in the centrifugal filter unit with 20 mM HEPES-HCl pH 8.0 and 150 mM NaCl, and finally concentrated to 20 mg/mL and flash frozen. For crystallization, the purified NTD was not frozen but was further purified by size exclusion chromatography (Superdex Increase 75 10/300 G, GE Healthcare), concentrated using a new 30 kDa centrifugal filter, and used immediately.

Intact mass spectrometry analysis of purified NTD constructs

The purpose of intact MS was to verify the n-terminal sequence on four constructs. N-linked glycans were removed by PNGase F after overnight non-denaturing reaction at room temperature. 4 µg of deglycosylated protein was used for each injection on the LC-MS system to acquire intact MS signal after separation of protease and protein by LC (Agilent PLRP-S reversed phase column). Thermo MS (Q Exactive Plus Orbitrap) was used to acquire intact protein mass under denaturing condition. BioPharma Finder 3.2 software was used to deconvolute the raw m/z data to protein average mass.

Non-reducing Peptide Mapping mass spectrometry analysis of purified NTD constructs

The purpose of peptide mapping was to identify the post modification on unpaired cysteines. Tryptic digestion was used without adding reducing reagent. 50 µg of deglycosyated protein was denatured (6 M guanidine hydrochloride), alkylated

(Iodoacetamide), and buffer exchanged (Zeba spin desalting column) before trypsin digestion. 10 µg of digested peptide was analyzed on the LC-MS system (Agilent AdvanceBio peptide mapping column and Thermo Q Exactive Plus Orbitrap MS) to acquire both MS1 and MS2 data under HCD fragmentation. Peptide mapping data was analyzed on Biopharma Finder 3.2 by searching the possible modifications such as cysteinylation and carbamidomethylation on the cysteines.

Isolation of peripheral blood mononuclear cells (PBMCs), plasma and sera

PBMCs were isolated from blood draw performed using tubes pre-filled with heparin, followed by Ficoll density gradient centrifugation. PBMCs were either used freshly along SARS-CoV2 Spike protein specific memory B cells sorting or stored in liquid nitrogen for later use. Sera were obtained from blood collected using tubes containing clot activator, followed by centrifugation and stored at −80°C.

B cell isolation and recombinant mAb production

Starting from freshly isolated PBMCs or upon cells thawing, B cells were enriched by staining with CD19 PE-Cy7 and incubation with anti-PE beads, followed by positive selection using LS columns. Enriched B cells were stained with anti-IgM, anti-IgD, anti-CD14 and anti-IgA, all PE labeled, and prefusion SARS-CoV-2 S with a biotinylated avi tag conjugated to Streptavidin Alexa Fluor 647 (Life Technologies). SARSCoV-2 S-specific IgG+ memory B cells were sorted by flow cytometry via gating for PE negative and Alexa Fluor 647 positive cells. Cells were cultured for the screening of positive supernatants. Antibody VH and VL sequences were obtained by RT-PCR and mAbs were expressed as recombinant human Fab fragment or as IgG1 (G1m3 allotype) carrying the half-life extending M428L/N434S (LS) mutation in the Fc region.

ExpiCHO cells were transiently transfected with heavy and light chain expression vectors as previously described (Pinto et al., 2020).

Affinity purification was performed on ÄKTA Xpress FPLC (Cytiva) operated by UNICORN software version 5.11 (Build 407) using HiTrap Protein A columns (Cytiva) for full length human and hamster mAbs and CaptureSelect CH1-XL MiniChrom columns (ThermoFisher Scientific) for Fab fragments, using PBS as mobile phase. Buffer exchange to the appropriate formulation buffer was performed with a HiTrap Fast desalting column (Cytiva). The final products were sterilized by filtration through 0.22 µm filters and stored at 4°C.

Enzyme-linked immunosorbent assay (ELISA)

To determine specificity of recombinantly produced mAbs, 96 half area well-plates (Corning) were coated over-night at 4°C with of SARS-CoV-2 S, NTD or RBD proteins prepared at 1 µg/mL, 2 µg/mL and 5 µg/mL in PBS pH 7.2, respectively. Plates were then blocked with PBS 1% BSA (Sigma) and subsequently incubated with mAbs serial dilutions for 1 h at room temperature. After 2 washing steps with PBS 0.05% Tween 20 (PBS-T) (Sigma-Aldrich) goat anti-human IgG secondary antibody (Southern Biotech) was added and incubated for 1 h at room temperature. Plates were then washed again with PBS-T and 4-NitroPhenyl phosphate (pNPP, Sigma-Aldrich) substrate added. After 30 min incubation, absorbance at 405 nm was measured by a plate reader (Biotek) and data plotted using Prism GraphPad.

For all other applications reported, the following ELISA procedure was followed: 30 µl of ectodomains (stabilized prefusion trimer) of S or NTD from SARS-CoV-2 were coated on 384 well ELISA plates at 1 ng/µl for 16 h at 4°C. Plates were washed with a

405 TS Microplate Washer (BioTek Instruments) then blocked with 80 µl SuperBlock (PBS) Blocking Buffer (Thermo Scientific) for 1 h at 37°C. Plates were then washed and 30 µl antibodies were added to the plates at concentrations between 0.001 and 100,000 ng/mL and incubated for 1 h at 37°C. Plates were washed and then incubated with 30 µl of 1/5000 diluted goat anti-human Fc IgG-HRP (Invitrogen). Plates were washed and then 30 µl Substrate TMB microwell peroxidase (Seracare) was added for 4 min at room temperature. The colorimetric reaction was stopped by addition of 30 µl of 1 N HCl. A_{450} was read on a Varioskan Lux plate reader (Thermo Scientific).

MLV-based pseudotyped virus production and neutralization

To generate SARS-CoV-2 S murine leukemia virus pseudotyped virus, HEK293T cells were seeded in 10-cm dishes in DMEM supplemented with 10% FBS. The next day cells were transfected with a SARS-CoV-2 S glycoprotein-encoding plasmid harboring the D19 C-terminal truncation (Ou et al., 2020), an MLV Gag-Pol packaging construct and the reporter vector pTG-Luc, using the X-tremeGENE HP DNA transfection reagent (Roche) according to the manufacturer's instructions. Cells were then incubated at 37°C with 5% CO_2 for 72 h. Supernatant was harvested and cleared from cellular debris by centrifugation at 400 x g, and stored at −80°C.

For neutralization assays, Vero E6 cells were seeded into white 96-well plates (PerkinElmer) at 20,000 cells/well and cultured overnight at 37°C with 5% CO_2 in 100 µl DMEM supplemented with 10% FBS and 1% penicillin/streptomycin. The next day, MLV-SARS-CoV-2 pseudovirus was activated with 10 µg/mL TPCK treated-Trypsin (Worthington Biochem) for 1 h at 37°C. Next, recombinant antibodies at various concentrations were incubated with activated pseudovirus for 1 h at 37°C. The Vero E6

cells were then washed with DMEM, and 50 µL of pseudovirus/mAbs mixes were added and incubated for 2 h at 37°C with 5% CO_2. After incubation, 50 µl of DMEM containing 20% FBS and 2% penicillin/streptomycin was added and the cells were incubated 48 h at 37°C with 5% CO_2. Following these 48 h of infection, culture medium was removed from the cells and 50 µl/well of Bio-Glo (Promega) diluted 1:2 with PBS with $Ca^{2+}Mg^{2+}$ (Thermo Fisher) was added to the cells and incubated in the dark for 15 min before reading on a Synergy H1 Hybrid Multi-Mode plate reader (Biotek). Measurements were done in duplicate, RLU values were converted to percentage of neutralization and plotted with a nonlinear regression curve fit in Graph Prism.

VSV-based pseudotype virus production and neutralization assay

SARS-CoV-2 S (YP 009724390.1), RaTG13 S (QHR63300.2), Pangolin-Guangdong S(QLR06867.1), Pangolin-Guanxi S (EPI ISL 410539), SARS-CoV S (YP 009825051.1), WIV1 S (AGZ48831.1) and WIV16 S (ALK02457.1) pseudotyped VSV viruses were prepared using 293T cells seeded in 10-cm dishes. Briefly, cells in DMEM supplemented with 10% FBS, 1% PenStrep were transfected with the plasmid encoding for the corresponding S glycoprotein using lipofectamine 2000 (Life Technologies) following manufacturer's indications. One day post-transfection, cells were infected with VSV(G*ΔG-luciferase) and after 2 h, infected cells were washed four times with DMEM before adding medium supplemented with anti-VSV-G antibody (I1- mouse hybridoma supernatant diluted 1 to 50, from CRL- 2700, ATCC). Particles were harvested 18 h post-inoculation, clarified from cellular debris by centrifugation at 2,000 x g for 5 min and concentrated 10 times using a 30 kDa cut off membrane and used for neutralization experiments, aliquoted and frozen at −80°C until use in neutralization experiments.

For neutralization, stable 293T cells expressing ACE2 (Crawford et al., 2020) in DMEM supplemented with 10% FBS, 1% PenStrep were seeded at 40,000 cells/well into clear bottom white walled 96-well plates and cultured overnight at 37°C. Twelve-point 3-fold serial dilutions of the corresponding mAb were prepared in DMEM and pseudotyped VSV viruses were added 1:1 to each mAb dilution in the presence of anti-VSV-G antibody from I1- mouse hybridoma supernatant diluted 50 times. After 45 min incubation at 37℃, 40 µl of the mixture was added to the cells and 2 h post-infection, 40 µL DMEM was added to the cells. After 17-20 h 50 µL/well of One-Glo-EX substrate (Promega) was added to the cells and incubated in the dark for 5-10 min prior reading on a Varioskan LUX plate reader (ThermoFisher). Measurements were done in duplicate with two independent productions of pseudotyped viruses and RLU values were converted to percentage of neutralization and plotted with a nonlinear regression curve fit in Graph Prism.

Neutralization of authentic SARS-CoV-2-Nluc virus

Neutralization of authentic SARS-CoV-2 by entry-inhibition assay Neutralization was determined using SARS-CoV-2-Nluc, an infectious clone of SARSCoV-2 (based on strain 2019-nCoV/USA_WA1/2020) which encodes nanoluciferase in place of the viral ORF7 and demonstrated comparable growth kinetics to wildtype virus (Xie et al., 2020). Vero E6 cells were seeded into black-walled, clear-bottom 96-well plates at 2 × 104 cells/well and cultured overnight at 37°C. The next day, 9-point 4-fold serial dilutions of mAbs were prepared in infection media (DMEM + 10% FBS). SARS-CoV-2-Nluc was diluted in infection media at a final MOI of 0.1 or 0.01 PFU/cell, added to the mAb dilutions and incubated for 30 min at 37°C. Media was removed from the Vero E6 cells,

mAb-virus complexes were added and incubated at 37°C for 6 or 24 h. Media was removed from the cells, Nano-Glo luciferase substrate (Promega) was added according to the manufacturer's recommendations, incubated for 10 min at room temperature and the luciferase signal was quantified on a VICTOR Nivo plate reader (Perkin Elmer).

Binding and affinity determination by Biolayer Interferometry (BLI)

BLI measurements were performed using an Octet Red96 (ForteBio). All reagents were prepared in kinetics buffer (PBS plus 0.01% BSA) at the indicated concentrations.

BLI was used to assess antibody binding affinity to SARS-CoV-2 NTD. IgG antibodies were prepared at 2.7 µg/mL and captured on pre-hydrated Protein A biosensors (Sartorius) for 1 min. The biosensors with immobilized antibodies were moved into kinetics buffer with SARS-CoV-2 NTD (concentrations tested: 333.3, 166.6, 83.3, 41.7, 20.8, 10.4, 5.2 nM) for 5 min (i.e., association). The dissociation of the SARS-CoV-2 NTD was then recorded for 9 min in wells containing kinetics buffer. Affinity constants were calculated using a global fit model and results were plotted using GraphPad Prism.

BLI was also used to assess antibody competition studies to define the NTD antigenic map. Biotinylated SARS-CoV-2 S protein was prepared at 10 µg/mL in kinetics buffer and loaded on pre-hydrated High Precision Streptavidin SAX Biosensors (Sartorius) for 3 min. NTD mAbs at 20 µg/mL in kinetics buffer were then sequentially added to observe binding competition and signal recorded for 5 min (or 7 min). Given the low responses detected for binding of S2L11 and S2X176 to the S ectodomain trimer, mapping of these mAbs was performed by capturing the purified NTD to BLI

biosensors. NTD was prepared at 8 µg/mL and assay performed using the same kinetics conditions described above.

BLI was also used to assess mAb-mediated inhibition of SARS-CoV-2 S binding to human recombinant ACE2. Before the assay SARS-CoV2 S ectodomain trimer (5 µg/mL) was incubated with tested mAbs (30 µg/mL) or no mAb for 30 min at 37°C. Biotinylated recombinant human ACE2 protein (2 µg/mL) was immobilized on High Precision Streptavidin SAX Biosensors (Sartorius). Next, an association step with S/mAb complexes was performed for 10 min. Results were plotted using GraphPad Prism.

S2X28 blockade-of-binding to NTD

S2X28 mAb was biotinylated using EZ-Link NHS-PEG solid phase biotinylation kit (ThermoFisher Scientific) and binding to NTD tested to set optimal concentration to be used in the assay after sample desalting using Zeba Spin Desalting Columns (ThermoFisher Scietific). Half area 96 well-plates were coated over-night at 4°C with SARS-CoV-2 NTD diluted at 2 µg/mL in PBS. After a blocking step with Blocker Casein (ThermoFisher Scientific), serial plasma dilutions in Blocker Casein were incubated 1 h at room temperature. Biotinylated S2X28 was added at a concentration achieving 70% of maximal binding and the mixture was incubated for 45 min at room temperature. Alkaline-phosphatase conjugated streptavidin (Jackson ImmunoResearch) was diluted at 0.5 µg/mL in Blocker Casein and added on plates previously washed 4 times with PBS 0.05%Tween 20. After 30 min incubation, plates were washed and 4-NitroPhenyl phosphate substrate incubated for 45 min at room temperature. Absorbance at 405 nm

was measured and percentage of inhibition was calculated as follows: (1-(OD sample-OD neg ctr)/ (OD pos ctr-OD neg ctr)) x100.

Affinity determination by Surface Plasmon Resonance (SPR)

SPR binding measurements were performed using a Biacore T200 instrument where purified avi-tagged SARS-CoV-2 S D614G ectodomain trimer was captured using anti-AviTag pAb covalently immobilized on a CM5 sensor chip. The running buffer was Cytiva HBS-EP+ pH 7.4; measurements were performed at 25þC. Affinity/avidity determinations were run as single-cycle kinetics, with a 3-fold dilution series of mAb starting from 300 nM, and each concentration injected for 180 s. Double reference-subtracted data were fit to a 1:1 binding model using Biacore Evaluation software. Fit results for IgG yielded apparent equilibrium dissociation constants due to avidity. For dissociation rates that were too slow to fit, equilibrium dissociation constants are reported as an upper limit.

Transient Expression of Sarbecovirus S protein in ExpiCHO-S Cells

Immediately before transfection, ExpiCHO-S cells were seeded at 6×10^6 cells cells/mL in a volume of 5 mL in a 50 mL bioreactor. Spike coding plasmids were diluted in cold OptiPRO SFM, mixed with ExpiFectamine CHO Reagent (Life Technologies) and added to the cells. Transfected cells were then incubated at 37°C with 8% CO_2 with an orbital shaking speed of 120 RPM (orbital diameter of 25 mm) for 42 h

Binding to cell surface expressed Sarbecovirus S proteins by Flow Cytometry

Transiently transfected ExpiCHO cells were harvested and washed two times in wash buffer (PBS 1% BSA, 2 mM EDTA). Cells were counted, distributed into round bottom 96-well plates (Corning) and incubated with the NTD antibodies at the final

concentration of 5 µg/mL. Alexa Fluor647-labeled Goat Anti-Human IgG secondary Ab (Jackson Immunoresearch) was prepared at 1.5 µg/mL added onto cells after two washing steps. Cells were then washed twice and resuspended in wash buffer for data acquisition at ZE5 cytometer (Biorad).

Fusion inhibition assay

Vero E6 cells were seeded in 96 well plates at 15,000 cells per well in 70 µL DMEM with high glucose and 2.4% FBS (Hyclone). After 16 h at 37°C with 8% CO_2, the cells were transfected with SARS-CoV-2-S-D19_pcDNA3.1 as follows: for 10 wells, 0.57 µg plasmid SARS-CoV-2- S-D19_pcDNA3.1 were mixed with 1.68 µl X-tremeGENE HP in 30 µl OPTIMEM. After 15 min incubation, the mixture was diluted 1:10 in DMEM medium and 30 µL was added per well. A 4-fold serial dilution mAbs was prepared and added to the cells, with a starting concentration of 20 µg/mL. The following day, 30 µL 5X concentrated DRAQ5 in DMEM was added per well and incubated for 2 h at 37°C. Nine images of each well were acquired with a Cytation 5 equipment for analysis.

Measurement of Fc-effector functions

mAb-dependent activation of human FcγRs was performed with a bioluminescent reporter assay. ExpiCHO cells stably expressing full-length wild-type SARS-CoV-2 S (target cells) were incubated with different amounts of mAbs. After a 15-min incubation, Jurkat cells stably expressing FcγRIIIa receptor (V158 variant) or FcγRIIa receptor (H131 variant) and NFAT-driven luciferase gene (effector cells) were added at an effector to target ratio of 6:1 for FcγRIIIa and 5:1 for FcγRIIa. Signaling was quantified by the luciferase signal produced as a result of NFAT pathway activation.

Luminescence was measured after 20 h of incubation at 37þC with 5% CO2 with a luminometer using the Bio-Glo-TM Luciferase Assay Reagent according to the manufacturer's instructions (Promega).

Cell-surface mAb-mediated S_1 shedding

CHO cells stably expressing wild-type SARS-CoV-2 S were resuspended in wash buffer (PBS 1% BSA, 2 mM EDTA) and treated with 10 µg/mL TPCK-trypsin (Worthington Biochem) for 30 min at 37°C. Cells were then washed and distributed into round bottom 96-well plates (90,000 cells/well). MAbs were added to cells at 15 µg/mL final concentration for 180 min at 37°C. Cells were collected at different time points (5, 30, 60, 120 and 180), washed with wash buffer at 4°C, and incubated with 1.5 mg/mL secondary goat anti-human IgG, Fc fragment specific (Jackson ImmunoResearch) on ice for 20 min. Cells were washed and resuspended in wash buffer and analyzed with ZE5 FACS (Bio-rad).

Selection of SARS-CoV-2 monoclonal antibody escape mutants (MARMS)

VSV-SARS-CoV-2 chimera was used to select for SARS-CoV-2 S monoclonal antibody resistant mutants (MARMS) as previously described (Case et al., 2020; Liu et al., 2020b). Briefly, MARMS were recovered by plaque isolation on Vero cells with the indicated mAb in the overlay. The concentration of mAb in the overlay was determined by neutralization assays at a multiplicity of infection (MOI) of 100. Escape clones were plaque-purified on Vero cells in the presence of mAb, and plaques in agarose plugs were amplified on MA104 cells with the mAb present in the medium. Viral stocks were amplified on MA104 cells at an MOI of 0.01 in Medium 199 containing 2% FBS and 20 mM HEPES pH 7.7 (Millipore Sigma) at 34°C. Viral supernatants were harvested

upon extensive cytopathic effect and clarified of cell debris by centrifugation at 1,000 x g for 5 min. Aliquots were maintained at −80°C. Viral RNA was extracted from VSV-SARS-CoV-2 mutant viruses using RNeasy Mini kit (QIAGEN), and S was amplified using OneStep RT-PCR Kit (QIAGEN). The mutations were identified by Sanger sequencing (GENEWIZ). Their resistance was verified by subsequent virus infection in the presence or absence of antibody. Briefly, Vero cells were seeded into 12 well plates overnight. The virus was serially diluted using DMEM and cells were infected at 37°C for 1 h. Cells were cultured with an agarose overlay in the presence or absence of mAb at 34°C for 2 days. Plates were scanned on a biomolecular imager and expression of eGFP is show at 48 h post-infection.

Cryo-EM sample preparation and data collection

2.5 µL of 12 mg/mL S2M11 Fab, 2.5 µL of 12 mg/mL NTD Fab, and 6 µL of 5 mg/mL SARS-CoV-2 S (specifically the hexapro construct (Hsieh et al., 2020)) were incubated together for 30 min at 37°C. Alternatively, 3 µL of 3 mg/mL S2X28 Fab, and 3 µL of 1 mg/mL SARS-CoV-2 2P DS S (McCallum et al., 2020) were incubated together for 30 min at 37°C. Unbound Fab was then washed away with three consecutive dilutions in 400 µL of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl over a 100 kDa centrifugal filter (Amicon Ultra 0.5 mL centrifugal filters, MilliporeSigma). The complex was concentrated to 1.2 mg/mL and 3 µL was applied onto a freshly glow discharged 2.0/2.0 UltraFoil grid (200 mesh), plunge frozen using a vitrobot MarkIV (ThermoFisher Scientific) using a blot force of −1 and 6.5 s blot time at 100% humidity and 23°C.

Data were acquired using the Leginon software (Suloway et al., 2005) to control a FEI Titan Krios or Glacios transmission electron microscope equipped with a Gatan K2 Summit direct detectors and operated at 300 kV with a Gatan Quantum GIF energy filter or at 200 kV, respectively. For both microscopes, the dose rate was adjusted to 8 counts/pixel/s, and each movie was acquired in 50 frames of 200 ms. For the Krios, ~3000 micrographs were acquired for each session with a super-resolution pixel size of 0.525 Å with a defocus range between −0.8 and −2.0 µm. For the Glacios, ~100 micrographs with a pixel size of 1.16 Å were collected in a single session with a defocus range between −0.8 and −2.0 µm.

Cryo-EM data processing

Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking and extraction (with a box size of 400 pixels^2) were carried out using Warp (Tegunov and Cramer, 2019). At this stage, the pixel size was binned to 1.05 Å for data collected from the Krios. Reference-free 2D classification was performed using cryoSPARC (Punjani et al., 2017) to select well-defined particle images. 3D classification with 50 iterations each (angular sampling 7.5þ for 25 iterations and 1.8þ with local search for 25 iterations) were carried out using Relion (Zivanov et al., 2018) without imposing symmetry to separate distinct SARS-CoV-2 S conformations. 3D refinements were carried out using non-uniform refinement along with per-particle defocus refinement in cryoSPARC (Punjani et al., 2020) before particle images were subjected to Bayesian polishing using Relion (Zivanov et al., 2019). To accommodate the resolution of the S/S2M11/S2X333 dataset, at the polishing stage the box size was adjusted to 600 Å while the pixel size was binned to 0.896 Å. Another round of non-

uniform refinement in cryoSPARC was performed, followed by global and per-particle defocus refinement and again non-uniform refinement. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (Scheres and Chen, 2012).

Cryo-EM model building and analysis

UCSF Chimera (Goddard et al., 2007) and Coot (Casañal et al., 2019) were used to fit atomic models (PDB 7K43 and 6ZGE) into the cryo-EM maps. The model was then refined into the map using Rosetta (DiMaio et al., 2015; Frenz et al., 2019; Wang et al., 2016), Phenix (Liebschner et al., 2019) and ISOLDE (Croll, 2018). Analysis used MolProbity(Chen et al., 2010), EMringer(Barad et al., 2015), and Phenix (Liebschner et al., 2019). Figs. were generated using UCSF ChimeraX(Goddard et al., 2018) and UCSF Chimera (Goddard et al., 2007).

S2M28 Fab and NTD co-crystallization and structure determination

Crystals of S2M28 Fab with SARS-CoV-2 NTD were identified by the sitting-drop vapor diffusion method, set-up by hand with MCSG-1 crystallization screen (Anatrace). Optimized crystals were generated with 1 µl of 5.7 mg/mL Fab and 4 mg/mL NTD in 20 mM HEPES-HCl pH 8.0 and 150 mM NaCl plus 1 µl mother liquor solution containing 0.2 M Ammonium Sulfate, 0.1 M Sodium Citrate pH 4.75, 25% (w/v) PEG4000. Crystals were flash cooled in liquid nitrogen using the mother liquor solution supplemented with 17.5% (w/v) xylitol as a cryoprotectant. Diffraction data were collected on synchrotron beamline 5.0.2 at the Advanced Light Source, and processed with the XDS software package (Kabsch, 2010). Initial phases were obtained by molecular replacement using

Phenix-Phaser (McCoy et al., 2007), using the S/S2M11/S2M28 cryoEM structure. Several subsequent rounds of model building and refinement were performed using Coot (Casañal et al., 2019) and Phenix-Refine (Adams et al., 2010).

In vivo mAb testing using a Syrian hamster model

KU LEUVEN R&D has developed and validated a SARS-CoV-2 Syrian Golden hamster infection model (Boudewijns et al., 2020). The SARS-CoV-2 isolate used in this study (BetaCoV/Belgium/GHB-03021/2020-EPI ISL 109 407976|2020-02-03), was recovered from a nasopharyngeal swab taken from a RT-qPCR confirmed asymptomatic patient who returned from Wuhan, China at the beginning of February 2020. A close relatedness with the prototypic Wuhan-Hu-1 2019 SARS-CoV-2 isolate was confirmed by sequencing and phylogenetic analysis. Infectious virus was isolated by serial passaging on Huh7 and Vero E6 cells and passage 6 virus was used for the study described here. The titer of the virus stock was determined by end-point dilution on Vero E6 cells by the Reed and Muench method (Reed and Muench, 1938). This work was conducted in the high-containment A3 and BSL3+ facilities of the KU Leuven Rega Institute (3CAPS) under licenses AMV 30112018 SBB 219 2018 0892 and AMV 23102017 SBB 219 20170589 according to institutional guidelines.

Syrian hamsters (Mesocricetus auratus) were purchased from Janvier Laboratories and were housed per two in ventilated isolator cages (IsoCage N Biocontainment System, Tecniplast) with *ad libitum* access to food and water and cage enrichment (wood block). Housing conditions and experimental procedures were approved by the ethical committee of animal experimentation of KU Leuven (license P065-2020). 6-10 week-old female hamsters were administered by intraperitoneal

injection with S2X333 mAb at 1 mg/kg and 4 mg/kg 48 h before intranasal infection with 1.89×10^6 50% tissue culture infectious dose ($TCID_{50}$) in 50 µl inoculum. Hamsters were monitored for appearance, behavior and weight. At day 4 post infection hamsters were euthanized by intraperitoneal injection of 500 µL Dolethal (200 mg/mL sodium pentobarbital, Vétoquinol SA). Lungs were collected, homogenized using bead disruption (Precellys) in 350 µL RLT buffer (RNeasy Mini kit, QIAGEN) and centrifuged (10,000 rpm, 5 min, 4°C) to pellet the cell debris. RNA was extracted using a NucleoSpin kit (Macherey-Nagel) according to the manufacturer's instructions. RT-qPCR was performed on a LightCycler96 platform (Roche) using the iTaq Universal Probes One-Step RTqPCR kit (BioRad) with N2 primers and probes targeting the nucleocapsid (Boudewijns et al., 2020). Standards of SARS-CoV-2 cDNA (IDT) were used to express viral genome copies per mg tissue or per mL serum. To quantify infectious SARS-CoV-2 particles, endpoint titrations were performed on confluent Vero E6 cells in 96-well plates. Viral titers were calculated by the Reed and Muench method (Reed and Muench, 1938) and were expressed as $TCID_{50}$per mg tissue.

To detect escape mutants in the *in vivo* tests, cDNAs were synthesized from 100 ng vRNAs extracted from lung homogenates using Superscript III Reverse Transcriptase (Invitrogen) and Spike-cDNA-R oligonucleotide. A gene-specific fragment of 1664 bp encompassing the NTD region of SARS-CoV-2-S was amplified by PCR using Spike-SEQ1-F and Spike-SEQ1-R primers with Q5® Hot Start High-Fidelity DNA Polymerase (New England Biolabs). Reaction conditions were as follows: 50 µl reaction volume containing 2 µl of cDNA, 1X Q5 Reaction buffer, 200 µM each of four dNTPs (Cytiva Europe GmbH), 0.5 µM of each primer (Microsynth) and 0.02 U/µl of polymerase. PCR

conditions: 98°C for 30 s (initial denaturation); 45 cycles of 98°C for 10 s (denaturation), 62°C for 30 s (annealing), 72°C for 60 s (elongation), followed by 1 cycle of 72°C for 2 min (final elongation).PCR products were purified with GFX PCR DNA and Gel Band Purification Kit (Cytiva Europe GmbH), and sequencing was outsourced at Microsynth (Sanger sequencing) with Spike-SEQ2-F, Spike-SEQ3-F, Spike-SEQ-NTD-R, and Spike-SEQ2-NTD-R primers. Sequence analysis was performed with CLC Main Workbench 21 (QIAGEN) using NCBI NC_045512.2 and virus stock SARS-CoV-2 S - p6-20200525 as reference sequences.

Quantification and statistical analysis

Statistical analyses were performed using GraphPad Prism (v9) and Microsoft Excel for Windows 10 (v16.0.13001.20254). Statistical differences were analyzed with Mann-Whitney U-test. Statistical significance was defined as $^{*}p < 0.05$, $^{**}p < 0.01$. EC_{50} and IC_{50}, values were determined by non-linear regression analysis (log(agonist) versus response - Variable slope (four parameters)).

6.12 Figures and Tables

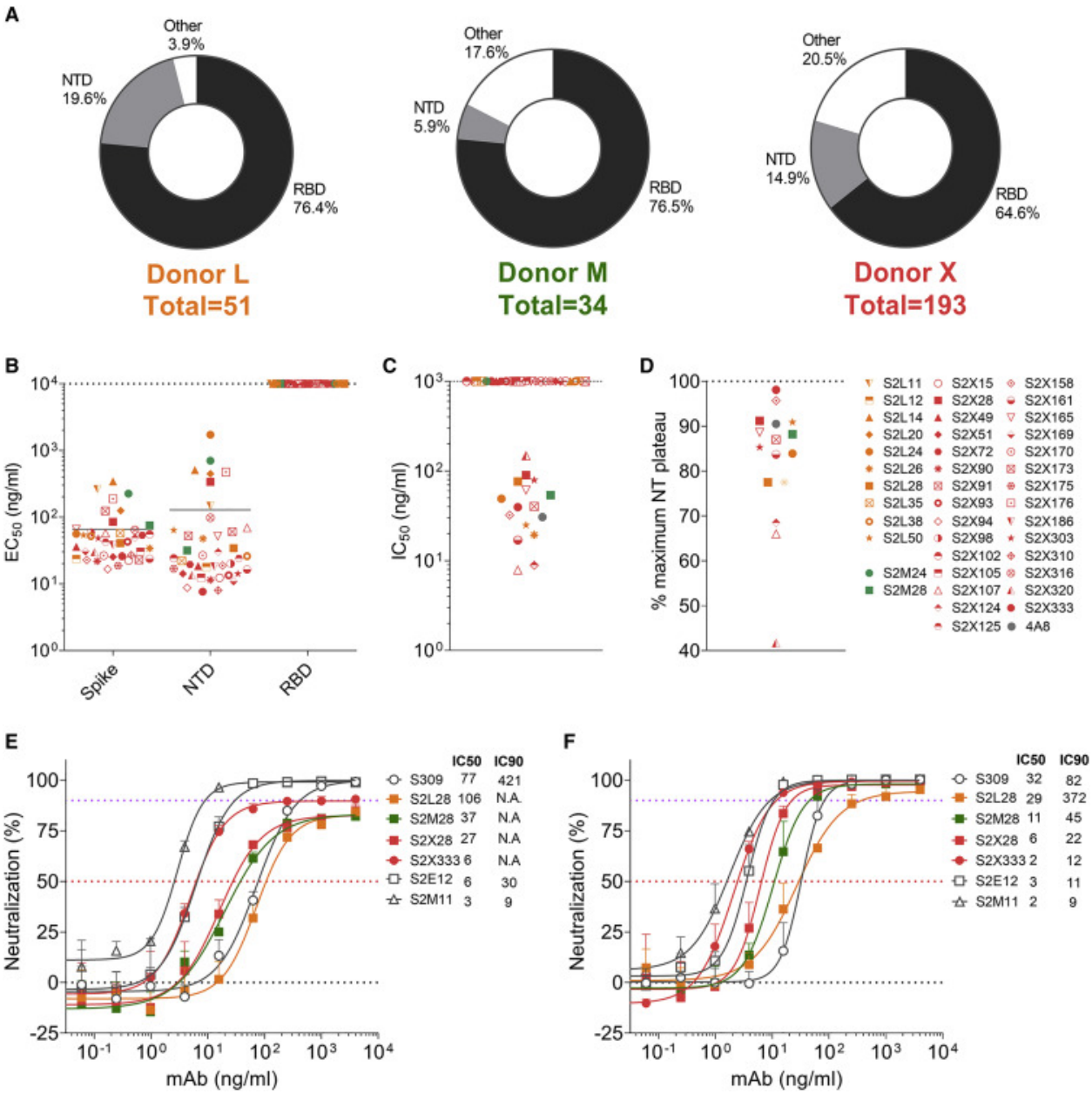


Figure 6.1 Discovery of potent NTD-specific SARS-CoV-2 neutralizing mAbs from three convalescent individuals.

(**A**) Pie charts showing the frequency of mAbs (cloned from IgG$^+$memory B cells) recognizing the SARS-CoV-2 NTD, RBD, or other S regions for patients L, M, and X. (**B**) Binding of the 41 isolated NTD mAbs to immobilized SARS-CoV-2 S, NTD, or RBD analyzed by ELISA. (**C and D**) Neutralization potencies (IC_{50}, C) and maximal neutralization plateau (NT, D) of 15 NTD-specific neutralizing mAbs against SARS-CoV-2 S MLV pseudotyped virus. Data are from one out of two independent experiments performed. (**E and F**) Dose-dependent neutralization of selected NTD- and RBD-specific mAbs against authentic SARS-CoV-2-Nluc assessed 6 h (at MOI 0.1) (**E**) or 24 h (at MOI of 0.01) (**F**) after infection. One independent experiment out of at least two

is shown. Error bars indicate standard deviation of triplicates. See also Fig. S1, Table S1, and Data S1.

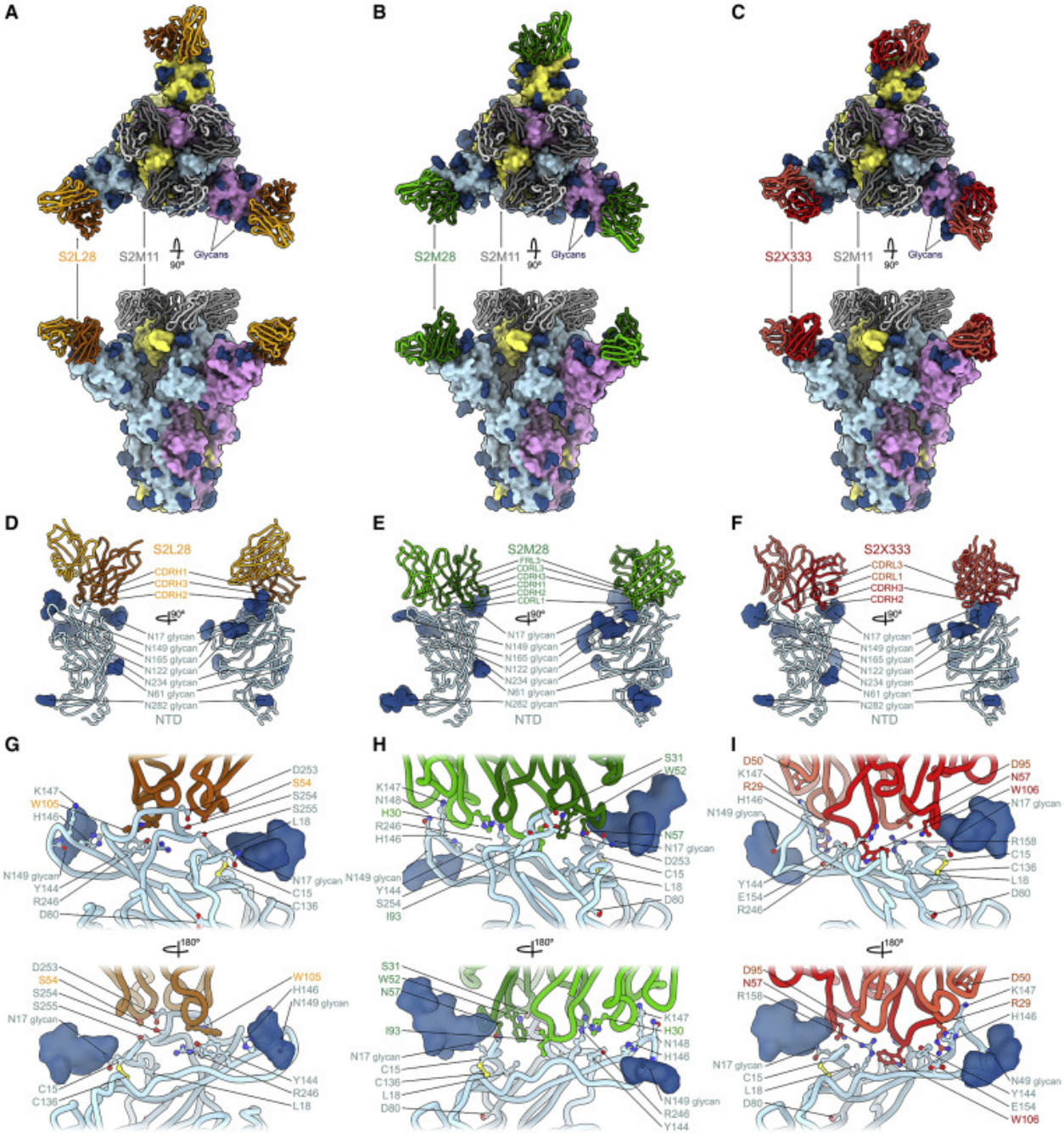


Figure 6.2: SARS-CoV-2 NTD neutralizing mAbs target the same antigenic supersite.

(A–C) Ribbon diagrams in two orthogonal orientations of the SARS-CoV-2 S ectodomain trimer (surface) bound to the RBD-specific S2M11 Fab (gray) and to the NTD-targeted Fab S2L28 (A), S2M28 (B), or S2X333 (C).
(D–F) S2L28 (D), S2M28 (E), and SX333 (F) binding pose relative to the NTD.

(G–I) Zoomed-in views showing selected interactions of S2L28 (G), S2M28 (H), or S2X333 (I) with the NTD. SARS-CoV-2 S protomers are colored pink, cyan, and gold whereas N-linked glycans are rendered as dark blue surfaces.
See also Fig. S2 and Tables S2–S4.

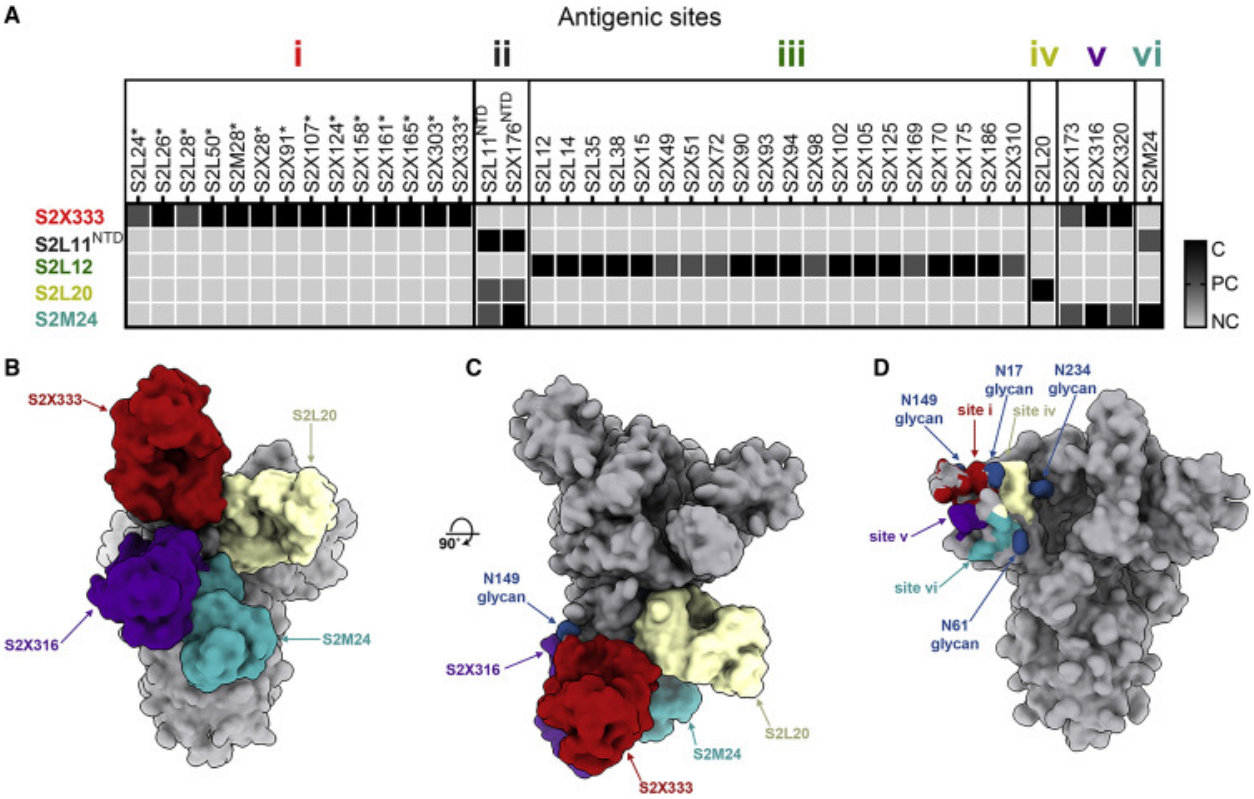


Figure 6.3: 3 The SARS-CoV-2 NTD comprises multiple antigenic sites.

(**A**) Epitope binning of the 41 NTD-specific mAbs isolated led to the identification of six antigenic sites based on competition binding assays using biolayer interferometry. The competition assays were performed either using the S ectodomain trimer or the NTD (^{NTD}). NC, no competition; PC, partial competition; C, competition; *, potently neutralizing mAbs. One independent experiment out of two is shown.
(**B and C**) Composite model of the SARS-CoV-2 S trimer viewed along two orthogonal orientations with four bound Fab fragments representative of antigenic site i (S2X333), site iv (S2L20), site v (S2X316), and site vi (S2M24).
(**D**) Footprints of the antigenic sites identified structurally are shown along with neighboring N-linked glycans (blue spheres).
See also **Fig. S3**, **Table S5**, and **Data S1**.

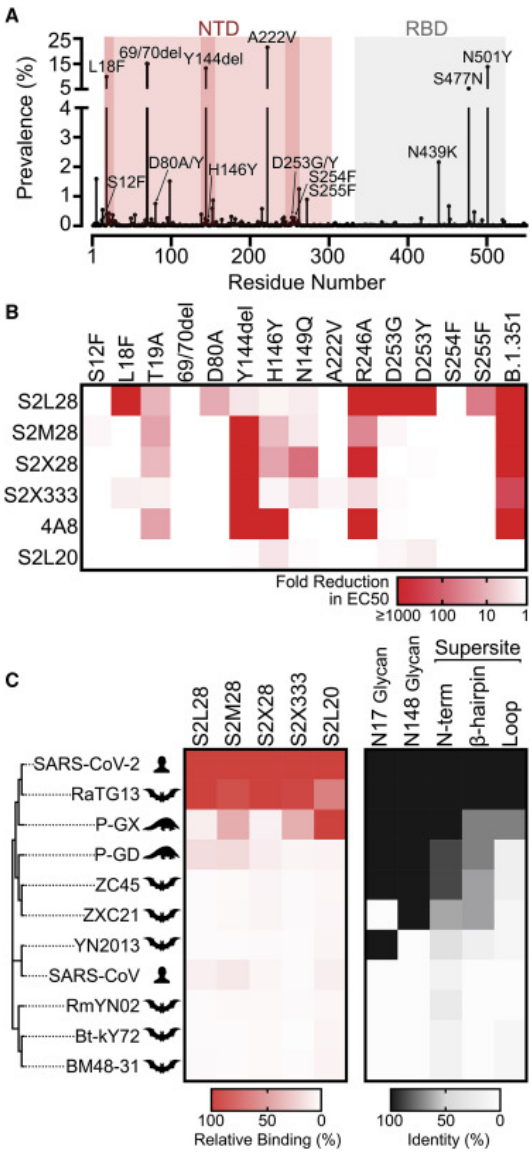


Figure 6.4: NTD mAbs are affected by the genetic diversity of circulating SARS-CoV-2 variants and sarbecoviruses.

(A) Prevalence of SARS-CoV-2 S variants among circulating isolates as of February 12, 2021 (508,771 sequences). The NTD and RBD are highlighted with red and gray backgrounds, respectively. The NTD supersite residues are highlighted by darker red background shading.
(B) S2L28, S2M28, S2X28, S2X333, 4A8, and S2L20 binding to recombinant SARS-CoV-2 NTD variants analyzed by ELISA and displayed as a heatmap. 69/70del, deletion of residues 69/70; Y144del, deletion of residue Y144. EC_{50}s were calculated as mean of duplicates from two technical replicates.
(C) Heatmap of NTD mAb binding to a panel of sarbecovirus S glycoproteins expressed at the surface of ExpiCHO cells analyzed by flow cytometry. Binding is expressed as mean fluorescence intensities relative to SARS-CoV-2 S binding for each mAb. One independent experiment out of two is shown. A coronavirus NTD cladogram is shown

on the left. The conservation of relevant glycans and NTD supersite regions is also shown (right).
See also Fig. S4 and Data S1.

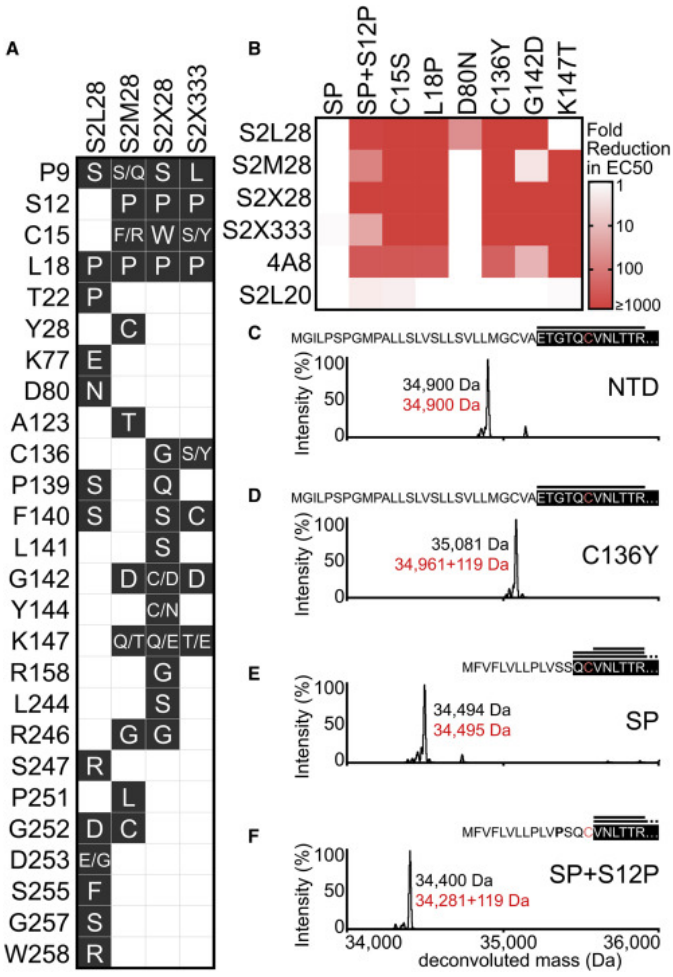


Figure 6.5: Analysis of SARS-CoV-2 S neutralization escape mutants.

(A) Escape mutations selected with each mAb (black background white font) and residues that remained unchanged (white background) upon passaging a VSV-SARS-CoV-2 S chimeric virus.
(B) S2L28, S2M28, S2X28, S2X333, 4A8, and S2L20 binding to selected SARS-CoV-2 NTD escape mutants analyzed by ELISA and displayed as a heatmap. The native signal peptide was introduced using the wild-type sequence (SP) or with the S12P mutation (SP+S12P). EC_{50}s were calculated as mean of duplicates from two technical replicates.
(C–F) Deconvoluted mass spectra of purified NTD constructs, including the NTD construct with an optimized signal peptide (C, NTD), the NTD construct with an optimized signal peptide and the C136Y mutation (D, NTD), the NTD construct with the native signal peptide (E, NTD), and the NTD construct with the native signal peptide carrying the S12P mutation (F, NTD). The empirical mass (black) and theoretical mass

(red) are shown beside the corresponding peak. Additional 119 Da were observed for the C136Y and SP+S12P NTDs corresponding to cysteinylation of the free cysteine residue in these constructs (as L-cysteine was present in the expression media). The cleaved signal peptide (black text white background) and subsequent N-terminal sequence (white text black background) are also shown based on the mass spectrometry (MS) results; C15 is highlighted in light red. Peptides identified by MS/MS analysis consistent with the mass of N-terminal peptides are shown above the N-terminal sequence (black horizontal lines). See also **Data S1**.

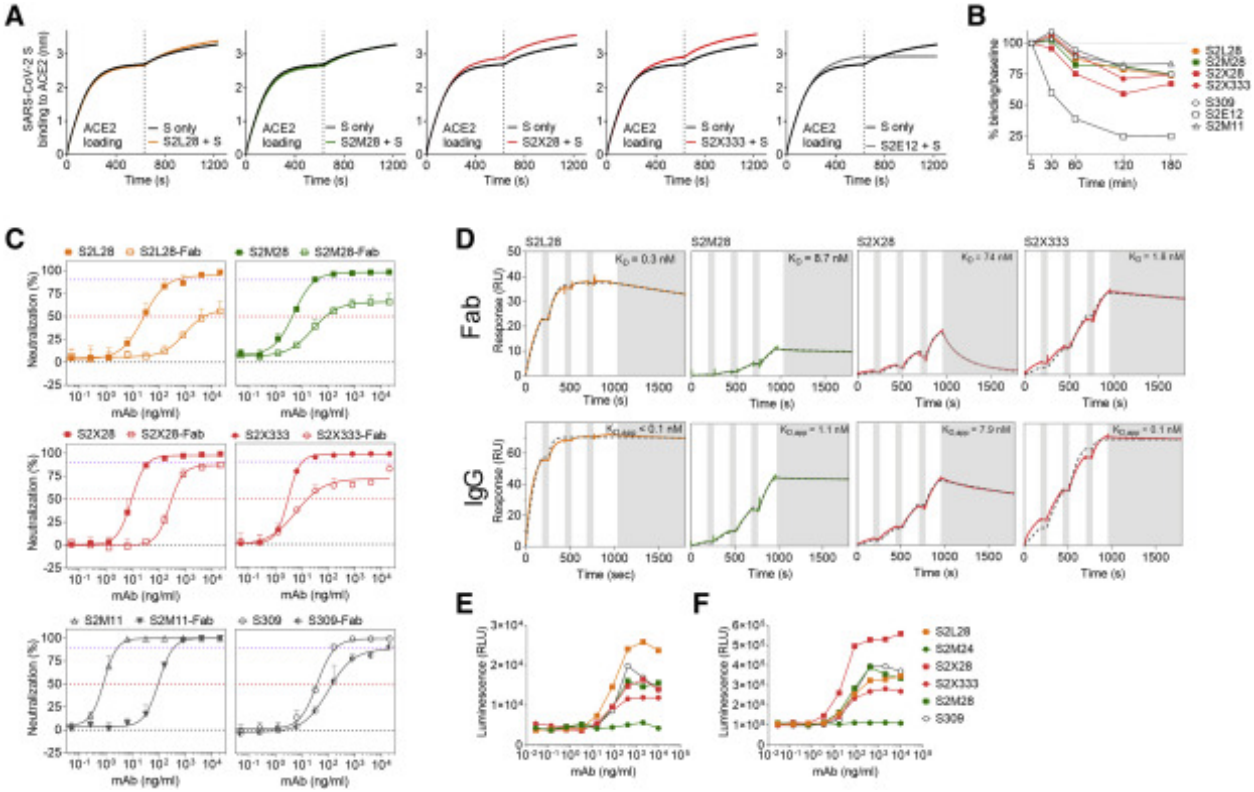


Figure 6.6: Mechanism of action of NTD-specific neutralizing mAbs.

(A) Competition of S2L28, S2M28, S2X28, and S2X333 with ACE2 to bind to SARS-CoV-2 S as measured by biolayer interferometry. ACE2 was immobilized at the surface of the biosensors before incubation with the S ectodomain trimer alone or pre-complexed with mAbs. The vertical dashed line indicates the start of the association of free S or S/mAb complexes with solid-phase ACE2. The anti-RBD S2E12 mAb was included as positive control.
(B) mAb-mediated S_1 subunit shedding from cell surface expressed SARS-CoV-2 S as determined by flow cytometry. The anti-RBD S2M11 and S2E12 mAbs were included as negative and positive controls, respectively.
(C) Neutralization of authentic SARS-CoV-2 (SARS-CoV-2-Nluc) by S2L28, S2M28, S2X28, and S2X333 IgG or Fab. S309 and S2M11 Fab and IgG were also included as controls. Symbols are means ± SD of triplicates. Dotted lines indicate IC_{50} and IC_{90}values.

(D) SPR analysis of mAbs binding to the SARS-CoV-2 S ectodomain trimer. Gray dashed line indicates a fit to a 1:1 binding model. The equilibrium dissociation constants (K_D) or apparent equilibrium dissociation constants (K_D, app) are indicated. White and gray stripes indicate association and dissociation phases, respectively.
(E and F) Activation of FcγRIIa H131 (E) and FcγRIIIa V158 (F) induced by NTD-specific mAbs. The anti-RBD mAb S309 was included as positive control. One independent experiment out of at least two is shown. See also **Figs. S1 and S5 and Table S1**.

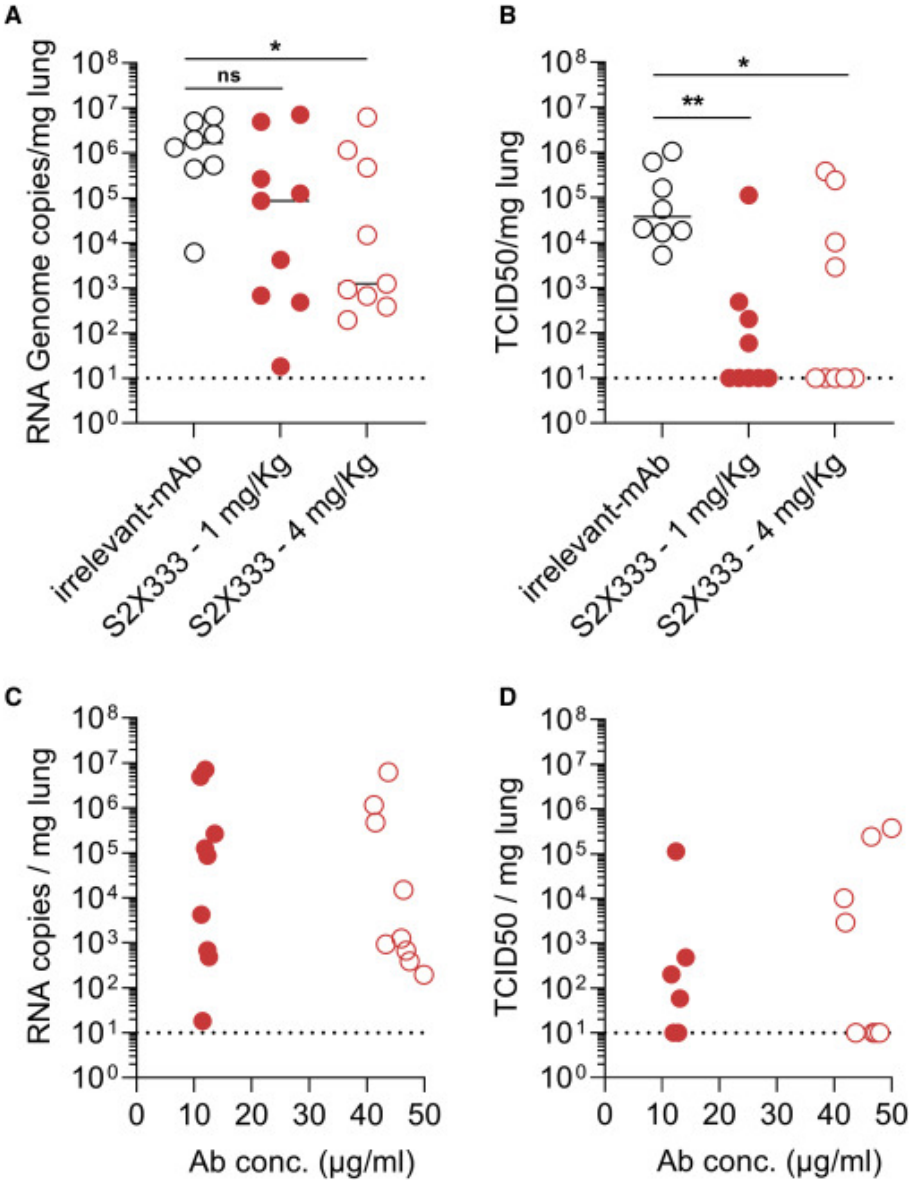
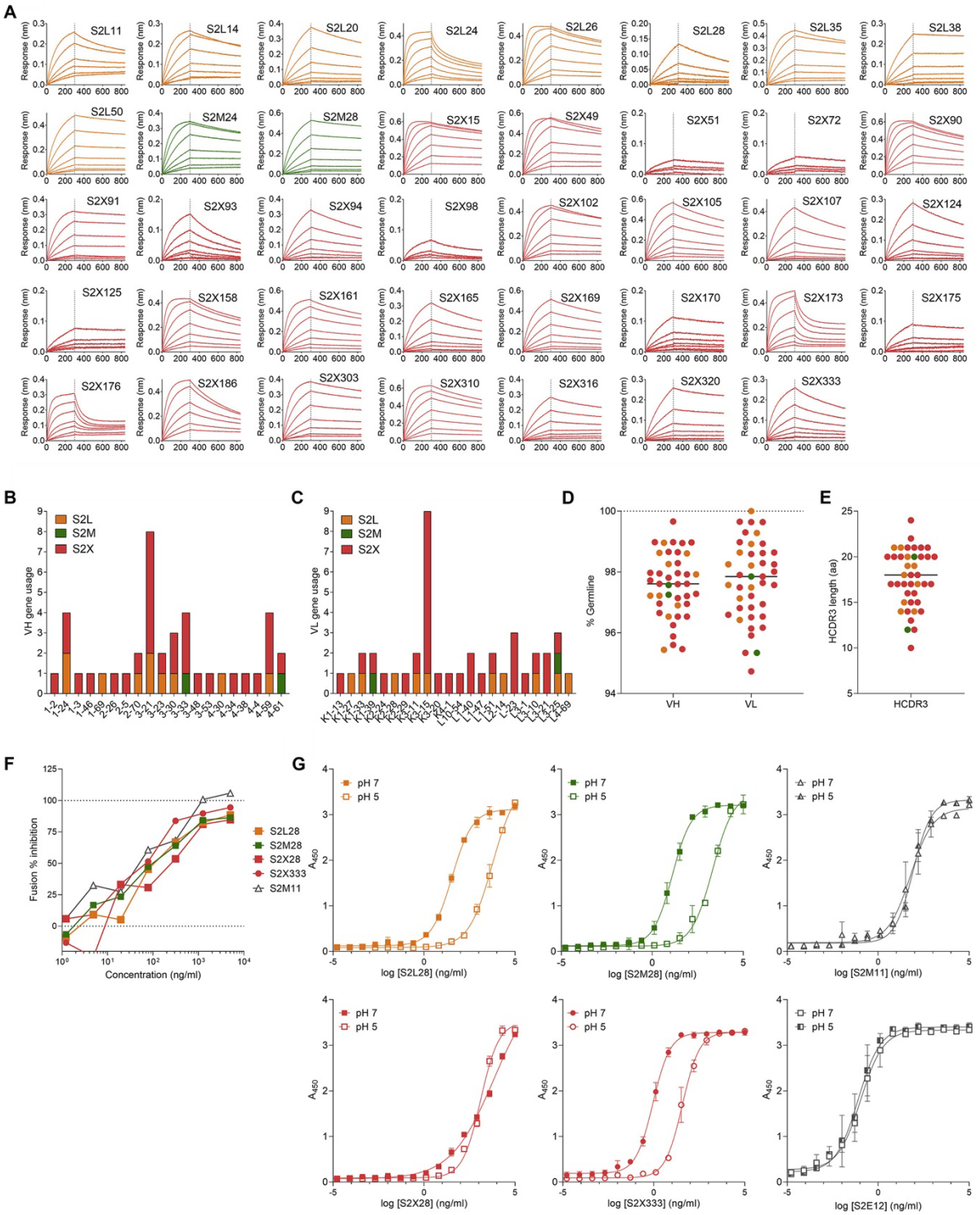


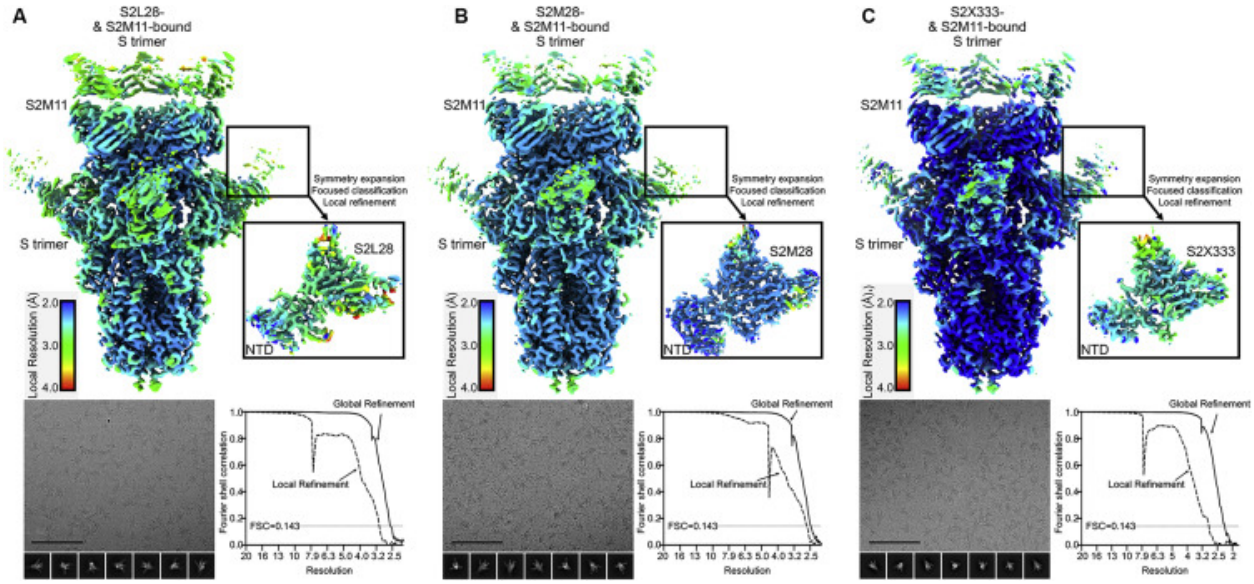
Figure 6.7: S2X333 provides robust in vivo protection against SARS-CoV-2 challenge but selects for escape mutants.

Syrian hamsters were injected with the indicated amount of mAb 48 h before intra-nasal challenge with SARS-CoV-2.
(A) Quantification of viral RNA in the lungs 4 days post-infection.
(B) Quantification of replicating virus in lung homogenates harvested 4 days post-infection using a $TCID_{50}$ assay.
(C and D) Viral RNA loads and replicating virus titers in the lung 4 days post-infection plotted as a function of serum mAb concentrations before infection (day 0). Mann-Whitney test was used. $^{*}p < 0.05$, $^{**}p < 0.01$. Data from two independent experiment are presented. (Irrelevant mAb n = 8; S2X333 1 mg/kg n = 9; S2X333 4 mg/kg n = 9). See also **Fig. S6**.



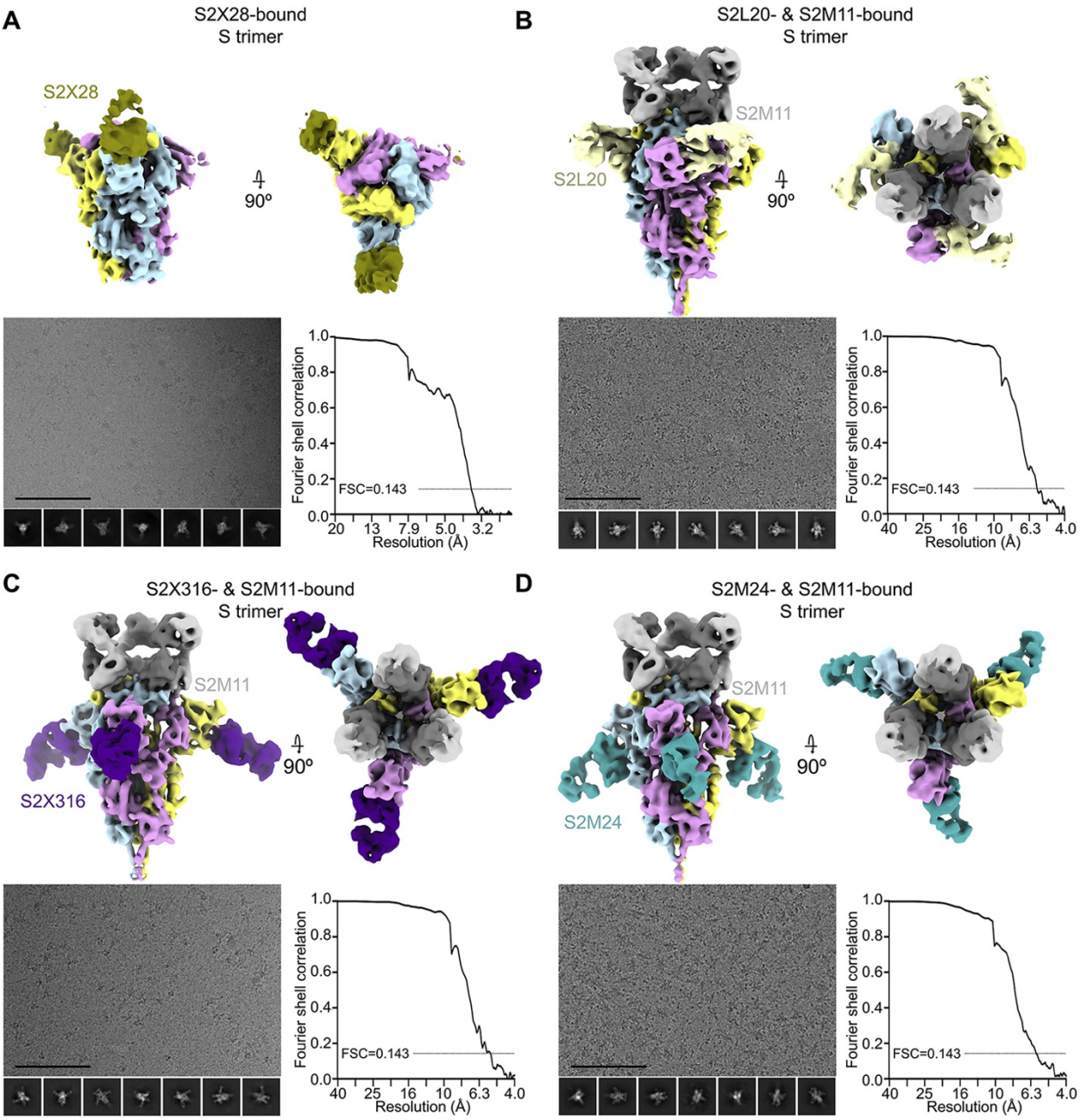
Supplemental Figure 6.1: Characteristics of SARS-CoV-2 NTD mAbs related to Fig. 1.

(A) Biolayer interferometry binding kinetic analysis of the SARS-CoV-2 NTD to immobilized mAbs. (B and C) V gene usage for the heavy (B) and light (C) chains of the NTD mAbs.
(D) Nucleotide sequence identity of the mAbs isolated relative to the respective V germline genes. (E) HCDR3 amino acid length for individual mAbs. (F) Cell-to-cell fusion inhibition assay with Vero E6 cells transfected with SARS-CoV-2 S and incubated with varying concentrations of S2L28, S2M28, S2X28, S2X333 or the RBD-specific mAb S2M11. (G) Binding of NTD- and RBD-specific mAbs to immobilized SARS-CoV-2 S at pH7 and pH5 as analyzed by ELISA. One independent experiment out of two is shown.



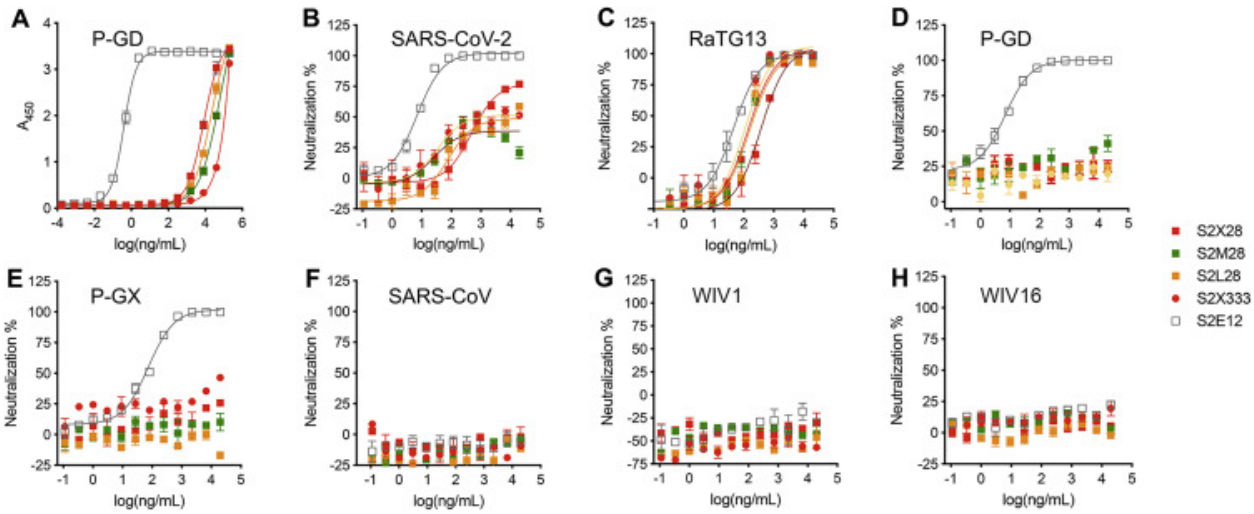
Supplemental Figure 6.2: Cryo-EM data processing of SARS-CoV-2 S bound to S2L28, S2M28, or S2X333, related to Fig. 2.

(Top) Unsharpened maps colored by local resolution calculated using cryoSPARC for the S trimer bound to S2M11 Fab and either S2L28 (A), S2M28 (B), or S2X333 Fab (C), as well as the locally refined reconstruction of NTD-bound Fab (inset). (Bottom) Representative electron micrograph and class averages (bottom left of each panel) are shown for SARS-CoV-2 S in complex with the indicated Fabs embedded in vitreous ice (Scale bar: 100 nm).



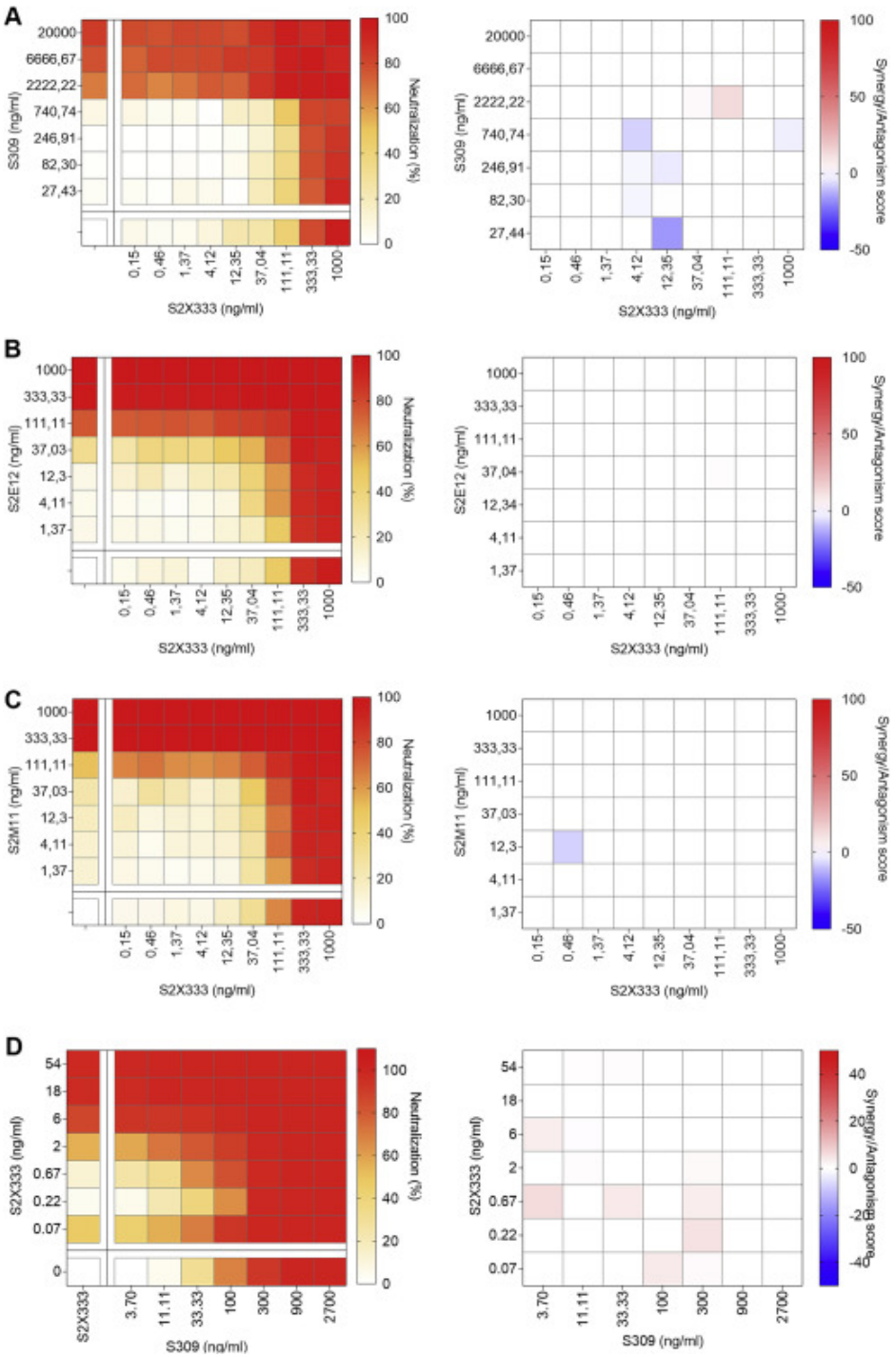
Supplemental Figure 6.3: Cryo-EM data processing of SARS-CoV-2 S bound to S2X28, S2L20, S2X316, or S2M24, related to Figs. 2 and 3.

(A) 6 Å low-pass filtered map of the SARS-CoV-2 2P DS S trimer (McCallum et al., 2020) bound to S2X28 (gold).
(B–D) Sharpened maps of the S trimer bound to S2M11 Fabs and S2L20 (B, gray), S2X316 (C, purple), or S2M24 Fabs (D, sky blue). Representative electron micrographs and class averages are shown at the bottom left of each panel (Scale bar: 100 nm). The corresponding Fourier shell correlation curves (bottom right of each panel) are shown with the 0.143 cutoff indicated by horizontal dashed lines.



Supplemental Figure 6.4: NTD neutralizing mAbs inhibit S-mediated entry of the closely related RaTG13 but not of more distant sarbecoviruses, ***related to Fig. 4***

SARS-CoV-2 S neutralization escape mutants reveal unconventional mechanisms of NTD mAb resistance



Supplemental Figure 6.5: Neutralizing activity of SARS-CoV-2 NTD- and RBD-targeting mAb cocktails, related to Fig.s 1 and 6.

(A–C) SARS-CoV-2-MLV pseudotypes neutralization (left) and synergy score (right) measured combining S2X333 with the RBD-targeting mAb S309 (A), S2E12 (B), or S2M11 (C).

(D) Neutralization matrix to assess the synergistic activity of S2X333 and S309 mAb cocktails *in vitro* with authentic SARS-CoV-2-Nluc. Data for authentic SARS-CoV-2-Nluc are from one representative experiment performed in triplicate each.

Supplemental Figure 6.6: Sequence alignment of SARS-CoV-2 S vRNA from input virus, control group, and selected outlier animals administered with S2X333, related to Fig. 7.

Nucleotide alignment of a region of interest reveals a deletion encompassing amino acid at position 144 only in S2X333-treated animals. SARS-CoV-2 S Wuhan = reference sequence (NCBI Reference Sequence NC_045512.2); SARS-CoV-2 S -p6-20200525 = virus stock; hamster 172 = animal from the control group; hamster 394 = outlier (TCID50) from the group administered with S2X333 at 1 mg/kg; hamster 397 = outlier (TCID50) from the group administered with S2X333 at 4 mg/kg. Alignment was performed using CLC Main Workbench 21 (QIAGEN)

CHAPTER 7. BROAD SARBECOVIRUS NEUTRALIZATION BY A HUMAN MONOCLONAL ANTIBODY.

In this next study, I will describe our efforts to identify and characterize broadly neutralizing antibodies that are not affected by the ongoing antigenic drift and that can prevent or treat future zoonotic infections. Here we describe a human monoclonal antibody designated S2X259, which recognizes a highly conserved cryptic epitope of the receptor-binding domain and cross-reacts with spikes from all clades of sarbecovirus. S2X259 broadly neutralizes spike-mediated cell entry of SARS-CoV-2, including variants of concern, as well as a wide spectrum of human and potentially zoonotic sarbecoviruses through inhibition of angiotensin-converting enzyme 2 (ACE2) binding to the receptor-binding domain. Our data reveal a key antigenic site that is targeted by broadly neutralizing antibodies and will guide the design of vaccines that are effective against all sarbecoviruses.

Adapted from: Tortorici, M.A., Czudnochowski, N., Starr, T.N. *et al.* Broad sarbecovirus neutralization by a human monoclonal antibody. *Nature* **597**, 103–108 (2021). https://doi.org/10.1038/s41586-021-03817-4

7.1 Chapter Introduction

SARS-CoV-2 has caused a global pandemic with over 182 million infections and more than 3 million fatalities as of July 2021. SARS-CoV-2 genetic drift has resulted in

emerging variants of concern that are characterized by higher transmissibility, immune evasion and disease severity relative to the ancestral isolate (Deng, 2021; Challen, 2021; Davies, 2021; Tegally, 2021; McCallum, 2021; Wang, 2021; Collier, 2021; Wibmer, 2021; Chen, 2021; Madhi, 2021). Countermeasures, such as vaccines and therapeutic agents, are needed to cope with SARS-CoV-2 evolution and to protect against future sarbecovirus spillovers.

The sarbecovirus spike (S) glycoprotein mediates viral entry into host cells, represents the main target of neutralizing antibodies and is the focus of vaccine design(Tortorici, 2019). S comprises an S1 subunit, which recognizes host cell receptors, and an S2 subunit that promotes viral–cell mem- brane fusion (Walls, 2020. Wrapp, 2020). The S1 subunit includes the receptor-binding domain (RBD), which in the case of SARS-CoV and SARS-CoV-2 interacts with ACE2 to allow virus entry (Walls, 2020; Kuhn, 2004; Hoffmann, 2020; Letko, 2020).

A highly conserved region on sarbecovirus RBDs, designated antigenic site II (Piccoli, 2020), has previously been shown to elicit SARS-CoV and SARS-CoV-2 cross-neutralizing antibodies (Cohen, 2021; Yuan, 2020). However, site II becomes accessible only when at least two RBDs in the S trimer adopt an open conformation and thus has limited immunogenicity(Piccoli, 2020; Yuan, 2020). Here we describe a site-II-targeting monoclonal antibody (mAb) designated S2X259 that possesses exceptional neutralization breadth within the *Sarbecovirus* subgenus (including all SARS-CoV-2 variants of concern), and an escape profile that is limited to a sole and rare substitution. In addition, we show that prophylactic and therapeutic administration of S2X259 protects Syrian hamsters against SARS-CoV-2 challenge with the prototypic virus and

the B.1.351 variant of concern. Our findings identify S2X259 as a promising countermeasure to protect against SARS-CoV-2 antigenic drift as well as potentially new zoonotic infections, and highlight the importance of RBD site II for the design of vaccines eliciting antibodies against all sarbecoviruses.

7.2 A broadly neutralizing sarbecovirus mAb

We identified the mAb designated S2X259 from the memory B cells of an individual who had recovered from COVID-19 (**Supplementary Table** 7.**1**); this mAb cross-reacted with 25 out of 26 S glycoproteins representative of all clades of sarbecovirus (Cohen, 2021; Rappazzo, 2021) (**Fig. 1a, Extended Data Fig. 7.1a–c**). However, no binding was observed to embecovirus S (HCoV-HKU1 and HCoV-OC43) or merbecovirus S (MERS-CoV) glycoproteins (**Extended Data Fig. 7.1a, b**).

The cross-reactivity of S2X259 within the *Sarbecovirus* subgenus was further confirmed using enzyme-linked immunosorbent assay (ELISA) with a panel of 12 RBDs (**Fig. 7.1b**). S2X259 bound tightly to all of the RBDs tested (half maximal effective concentration (EC50) of less than 32 ng ml−1) with the exception of two bat strains, ZC45 (EC50 of 539 ng ml−1) and Anlong-112 (EC50 of 1,281 ng ml−1) (**Fig. 7.1b**). The broad recognition of S from the divergent bat Southeast Asian and non-Southeast Asian clades (clades 2 and 3, respectively) highlights the exceptional breadth of S2X259 in comparison with previously reported mAbs (Piccoli, 2020; Tortorici, 2020; Pinto, 2020) (**Extended Data Fig. 7.1d**). Surface plasmon resonance indicated that the S2X259 Fab fragment bound to clade 1a, 1b and 3 RBDs with nano- to picomolar affinities, and to clade 2 RBDs with micro- to nanomolar affinities. Of note, the affinity of S2X259 for the

RBD of the Guanxi pangolin coronavirus was 10–100-fold lower compared to other clade 1b RBDs (**Extended Data Fig. 7.2a**). The S2X259 Fab recognized both the SARS-CoV-2 prefusion-stabilized S ectodomain trimer and the recombinant RBD with picomolar affinities, albeit with a slower on-rate for binding to S (presumably owing to limited epitope exposure) (Walls, 2020; Wrapp, 2020) (**Extended Data Fig. 7.2a**). Finally, S2X259 binding was unaffected by several point mutants or constellation of mutations identified in circulating clinical isolates (K417V, N439K, Y453F and E484K) and in the RBDs of the B.1.1.7 (N501Y), B.1.351 (K417N/E484K/N501Y), B.1.427/B.1.429 (L452R) and P.1 (K417T/E484K/N501Y) lineages (**Fig. 7.1b, Extended Data Fig. 7.2b**).

We next evaluated the neutralization potency of S2X259 using authentic SARS-CoV-2-Nluc as well as murine leukaemia virus (MLV) and vesicular stomatitis virus (VSV) pseudotyping systems. S2X259 potently neutralized authentic SARS-CoV-2, SARS-CoV-2 S MLV and SARS-CoV-2 S VSV with half-maximal inhibitory concentrations (IC50) of 144 ng ml−1, 87 ng ml−1 and 213 ng ml−1, respectively (**Fig. 7.1c, d**). We subsequently evaluated S2X259-mediated neutralization against VSV pseudotypes containing SARS-CoV-2 S from the B.1.1.7, B1.351, P.1, B.1.427/B.1.429 and B.1.617.1 and other lineages, as well as the N439K and Y453F single mutations (McCallum, 2021; Davies, 2021; Thomson, 2021; Faria, 2021). S2X259 neutralized all of these mutants with a similar potency, consistent with the identical binding affinities determined using RBD mutants (**Fig. 7.1d, e, Extended Data Figs. 7.2b, 7.3a, 7.b, Supplementary Table 7.2**). Furthermore, S2X259 also neutralized a broad panel of ACE2-using sarbecoviruses with comparable potencies, including Guangdong pangolin

coronavirus, RaTG13, SARS-CoV, WIV1, WIV16 and RsSHC014 (**Fig. 7.1f, g, Supplementary Table 7.2**). Consistent with the surface plasmon resonance data, S2X259 neutralized VSV pseudo- types containing S of Guanxi pangolin coronavirus but with reduced potency as compared to SARS-CoV-2 (**Fig. 7.1g, Extended Data Fig. 7.2a, Supplementary Table 7.2**).

7.3 Structural basis of sarbecovirus neutralization

Three-dimensional classification of the cryo-electron microscopy data showed the presence of SARS-CoV-2 S ectodomain trimers saturated with three S2X259 Fabs bound to open RBDs swung out to various extents, for which we determined a consensus structure at 3.2 Å resolution (**Fig. 7.2a, b, Extended Data Fig. 7.4, Supplementary Table 7.3**). We used local refinement to determine a 3.7 Å resolution map of the region corresponding to the S2X259 variable domains and RBD. In parallel, we crystallized the S2X259 Fab-bound SARS-CoV-2 RBD in the presence of the S2H97 Fab26 and determined a 2.65 Å resolution X-ray structure (**Fig. 7.2c–e, Supplementary Table 7.4**).

S2X259 recognizes a glycan-free, cryptic epitope within antigenic site IIa, which was previously defined on the basis of the nonclonally related S2X35 mAb that was isolated from the same donor (Piccoli, 2020). S2X259 binding requires opening of two RBDs to grant access to the Fab in the context of the S trimer (**Fig. 7.2a, b**). S2X259 contacts the RBD using both heavy and light chains, which respectively contribute two thirds and one third of the approximately 950 Å2 paratope surface that is buried upon binding. Specifically, S2X259 uses complementary determining regions (CDRs) H1, H2, H3, L1 and L3 to contact RBD residues 369–386, which form two α-helices and an

intervening β-strand belonging to the structurally conserved RBD β-sheet, as well as the residues 404–411 and 499–508, which form a continuous surface made up of an α-helix and a loop followed by an α-helix, respectively.

The S2X259 epitope is conserved in circulating SARS-CoV-2 isolates and does not comprise prevalent RBD mutations, such as S477N, N439K or L452R. S2X259 also circumvents residues 417 and 484, and contacts the backbone of residue N501 but not its side chain, explaining the high-affinity binding to the B.1.1.7, B.1.351 and P.1 RBDs and potent neutralization of pseudotyped variant of concern viruses (**Fig. 7.1b, d Extended Data Fig. 7.2b**). S2X259 broad neutralization results from the conservation of its epitope among sarbecoviruses and from the angle of approach of the Fab that circumvents the SARS-CoV N357 glycan present in all sarbecoviruses, except for SARS-CoV-2 (SARS-CoV-2 N370) (**Extended Data Fig. 7.5**). The SARS-CoV-2 RBD and that of Guanxi pangolin coronavirus differ at only two positions (G504N and Y508F), and the structural data indicate that the former substitution would probably dampen mAb binding through steric hindrance as indicated by in vitro neutralization data (**Fig. 7.1g**).

S2X259 binding to the RBD does not affect or prevent recognition by site-I or site-IV-targeting mAbs (**Extended Data Fig. 7.6a, b**), which represent the majority of mAbs currently approved or in the clinic (Piccoli, 2020; Pinto, 2020; Barnes, 2020; Catchcart, 2021). Competition binding assays suggest that serum antibodies directed to the RBD antigenic site II are subdominant in individuals who have recovered from SARS-CoV-2 infection and vaccinated individuals, as compared to antibodies directed towards site I (**Extended Data Fig. 7.6c**). Therefore, S2X259 can be used in combination with other

site I or site IV mAbs to maximize breadth against currently circulating and emerging SARS-CoV-2 variants, as well as against future sarbecovirus infections from animal reservoirs.

7.4 S2X259 has an unusually narrow escape profile

Using deep-mutational scanning with a yeast-display RBD mutant library (Starr, 2020), we exhaustively mapped mutations that escape S2X259 binding. Notably, binding was strongly reduced by only a restricted number of amino acid substitutions, as compared to previously described neutralizing mAbs26, with mutations at RBD site 504 yielding the most marked reduction in binding (**Fig. 7.3a, Extended Data Fig. 7.7a–c**). Indeed, substitution of G504 with most other amino acids reduced S2X259 binding compared to the wild-type (Wuhan-Hu-1) RBD, which highlights the importance of this residue for mAb recognition.

To further explore the functional importance of the deep-mutational scanning data, we passaged a replication-competent VSV-SARS-CoV-2 S chimera (Case, 2020; Lui, 2021) in Vero cells in the presence of S2X259 mAb. The selective pressure imposed by S2X259 led to the emergence of viral mutants containing the G504D mutation, which was the sole mutation present in each of the 18 neutralization-resistant plaques we sampled. This substitution did not impair viral fitness of the VSV-SARS-CoV-2 S chimera **(Extended Data Fig. 7.7d, e, Supplementary Table 7.5**). Under the same conditions, S2X3517 (which also targets antigenic site IIa) selected for seven escape mutants at five different positions (Y369S, K378E, D405N, V503D or V503G and G504D or G504S)26 (**Supplementary Table 7.5**).

The G504D substitution is very rare: it has been detected in only 27 out of 1,345,708 (0.002%) SARS-CoV-2 genome sequences available on GISAID, as of May 2021 (**Fig. 7.3b**). Furthermore, in agreement with the deep-mutational scanning results, S2X259 efficiently neutralized VSV pseudoviruses containing S mutations located in S2X259-contact residues and found with high frequency in human isolates (**Fig. 7.3b, Extended Data Fig. 7.7f**). In the same analysis, we confirmed the loss of S2X259 neutralizing activity against the G504D mutant. S2X259 CDRL1 and CDRL3 contact G504 or surrounding residues, which explains why substituting a glycine with any other residue prevents mAb binding (via steric hindrance along with potential disruption of the interactions of the Fab with RBD residue D405, the latter being in close proximity to G504) (**Extended Data Fig. 7.7g**).

These data also support the hypothesis proposed above for the markedly reduced affinity and neutralization potency of S2X259 against Guanxi pangolin pseudovirus which contains the G504N mutation but maintains an aspartate at position 405 (**Fig. 7.1g, Extended Data Figs. 7.2a, 7.7f, 7.8**). However, S2X259 binding to clade 2 S glycoproteins that bear G504D or G504E mutations (for example, those of Shanxi2011 or YN2013) was not abrogated (**Fig. 7.1b**), possibly owing to substitution at the SARS-CoV-2 equivalent position 405 with a serine and to substantial deletions in the receptor-binding motif (**Extended Data Fig. 7.8**).

The finding that the only escape mutation for S2X259 is the rare G504D substitution suggests this mAb might have a high barrier for the emergence of SARS-CoV-2 escape mutants, which might prove essential during the next stages of the pandemic.

7.5 S2X259 protects hamsters from SARS-CoV-2 challenge

Consistent with structural data that indicates a clash between S2X259 and ACE2 for binding to the RBD (**Extended Data Fig. 7.9a**), we found that S2X259 efficiently blocked binding of SARS-CoV-2 and SARS-CoV RBDs to human ACE2 (Extended Data Fig. 9b).

In addition, S2X259 binding efficiently promoted shedding of the S1 subunit from cell-surface-expressed full-length SARS-CoV-2 S, as has previously been shown with other RBD-specific mAbs17 (Huo, 2020; Walls, 2019; Wec, 2020; Wrobel, 2020), but not with S2M11 (which locks S in the prefusion closed state) (Tortorici, 2020) (**Extended Data Fig. 7.9c).** The efficient S1 shedding induced by S2X259 might be responsible for the lack of FcγRIIa and FcγRIIIa activation, which we used as a proxy for antibody-dependent cellular phagocytosis and antibody-dependent cellular cytotoxicity, respectively (**Extended Data Fig. 7.9d, e).** Indeed, performing the same assays using target cells that express an uncleavable prefusion stabilized SARS-CoV-2 S protein (unable to release the S1 subunit), we confirmed the ability of S2X259 mAb to induce activation of FcγRIIa and FcγRIIIa in these conditions (**Extended Data Fig. 7.9f, g**).

These data show that the primary mechanism of S2X259-mediated neutralization of SARS-CoV-2 and other sarbecoviruses relies on blocking viral attachment to host-cell receptors through competitive inhibition of ACE2 binding. Premature triggering of S1 subunit shedding could inactivate viruses before encountering target cells, albeit reducing activation of S2X259-mediated effector functions.

Next, we evaluated the prophylactic activity of S2X259 against challenge with the prototypic (Wuhan-Hu-1-related) SARS-CoV-2 in a Syrian hamster model (Boudewijns, 2020). S2X259 was administered at 1, 5 and 25 mg per kg body weight via intraperitoneal injection before intranasal challenge, and the lungs of the hamsters were collected 4 days later for the quantification of viral RNA and replicating virus. Despite the lack of Fc-dependent effector functions in vitro and their proposed contribution to in vivo protection against SARS-CoV-2 (Schäfer, 2021; Winkler, 2021), administration of S2X259 at 1 mg kg−1 reduced the amount of viral RNA and infectious virus detected in the lungs by approximately 2 and 2.5 orders of magnitude, respectively, compared to hamsters that received a control mAb (**Fig. 7.4a, b**). Furthermore, complete abrogation of virus replication was observed in the lungs of the hamsters that were treated with S2X259 at 5 and 25 mg kg−1 relative to the control group (**Fig. 7.4b**). Therapeutic administration (after exposure) of S2X259 at 5 mg kg−1 reduced the amounts of viral RNA detected by 3 log and abrogated viral replication in the lungs of all of the hamsters but one (**Fig. 7.4c, d**). Hamsters treated with S2X259 at 10 or 20 mg kg−1 displayed a further approximately 1 log reduction of viral RNA and complete abrogation of viral replication (**Fig. 7.4c, d**). Viral replication in the lungs detected at day 4 inversely correlated with serum mAb concentration measured at the time of infection or at day 4 after infection in prophylactic and therapeutic settings, respectively (**Extended Data Fig. 7.10a, b**).

We next assessed the prophylactic activity of S2X259 against challenge with B.1.351 SARS-CoV-239. S2X259 at 1 and 4 mg kg−1, or in combination with S309 (each mAb at 1 mg kg−1), significantly decreased the amount of viral RNA detected in

the lungs relative to hamsters that received a control mAb (**Fig. 7.4e**). Moreover, S2X259 at 4 mg kg−1 or in combination with S309 (each mAb at 1 mg kg−1) completely abrogated SARS-CoV-2 B.1.351 virus replication in most of the hamsters (**Fig. 7.4f**). As was the case for the prototypic SARS-CoV-2, the amount of infectious virus at day 4 detected in the group administered only with S2X259 also inversely correlated with serum mAb concentration measured at the time of infection (**Extended Data Fig. 7.10c**). Moreover, hamsters that received the mAb cocktail appeared to benefit from the additive contribution of the S309 mAb (**Fig. 7.4f, Extended Data Fig. 7.10d, e**).

7.6 Discussion

The large number of mutations in the S glycoprotein of SARS-CoV-2 variants of concern (Tegally, 2021; Davies, 2021; Faria, 2021) has been shown to abrogate or reduce the neutralization potency of several mAbs and vaccine-elicited antibodies(McCallum, 2021; Wang, 2021; Collier, 2021; Wibmer, 2021; Chen, 2021; Madhi, 2021; Walls, 2021). Moreover, the detection of a large pool of sarbecoviruses in bats and other mammals across several continents suggests that future cross-species transmission events are likely to occur (Menachery, 2015; Menachery, 2016; Ge, 2013; Menachery, 2015). However, approved SARS-CoV-2 therapeutic agents and vaccines predominantly target or elicit immunity against immunodominant, but highly variable, epitopes and therefore have a limited efficacy against antigenic drift and genetically distinct strains with zoonotic potential (Walls, 2021; Martinez, 2021).

Strategies for eliciting broadly neutralizing and protective sarbecovirus antibodies targeting the RBD have recently been proposed (Cohen, 2021; Walls, 2021), and previous work has described mAbs that are now in the clinic with extended coverage

within clade 1 sarbecoviruses(Rappazzo, 2021; Pinto, 2020). The recent results from a phase III study demonstrating efficacy (that is, 85% reduction of hospitalizations or deaths) of the cross-reactive derivative of the S309 mAb (sotrovimab) paves the way for the development of additional neutralizing mAbs targeting conserved RBD epitopes that may broaden the coverage across sarbecoviruses and increase the barrier to resistance.

Here we identify the S2X259 mAb that targets the RBD antigenic site II and uniquely crossreacts with all of the SARS-CoV-2 variants of concern that we evaluated, as well as with 30 S trimers or RBDs from all four sarbecovirus clades. S2X259 broadly neutralizes pseudoviruses containing S proteins from the B.1.1.7, B.1.351, P.1, B.1.427/B.1.429, and B.1.617.1 variants, and from other representative members of SARS-CoV-2 and SARS-CoV clades (clades 1b and 1a, respectively). Of note, S2X259 neutralizes pseudotyped WIV1 S and RsSHC014 S (clade 1a) viruses, which have previously been suggested to be potential zoonotic threats to human health (Menachery, 2015; Menachery, 2016). Moreover, S2X259 also recognizes S and RBDs from bat Southeast Asian and non-Southeast Asian sarbecovirus clades (clades 2 and 3, respectively). The coverage of clade 2 might be particularly important, as a previous study (Letko, 2020) has shown that exogenous protease addition to pseudoviruses from this clade resulted in detectable S-mediated entry in human cells. Furthermore, although clade 2 and clade 3 sarbecoviruses have not been reported to bind human ACE2, mutation or recombination events could lead to human ACE2 binding and entry (Letko, 2020; Makarenkov, 2021). Overall, these findings indicate that S2X259 is a promising candidate for clinical development and could be stockpiled as part of a

pandemic preparedness toolbox. On the basis of the growing body of data demonstrating that antibodies targeting the SARS-CoV-2 RBD account for most neutralizing activity in sera from patients who have recovered from COVID-19 (Piccoli, 2020; Greaney, 2021), we propose that RBD-based vaccines could better promote elicitation of high titres of S2X259-like neutralizing antibodies, owing to the enhanced accessibility of this antigenic site compared to S-based (Walls, 2021; Walls 2020) vaccines . Our serological competition data confirm that the antibody responses against the conserved site targeted by S2X259 mAb is subdominant in donors who have recovered from COVID-19 as well as in individuals who have received S-based mRNA vaccines. We anticipate these data will guide future efforts to develop vaccines that overcome the emergence of SARS-CoV-2 variants and that are effective against all sarbecoviruses.

7.7 Methods

Cell Lines

Cell lines used in this study were obtained from ATCC (HEK293T, Vero and Vero E6), ThermoFisher Scientific (ExpiCHO-S cells, FreeStyle 293-F cells and Expi293F cells) or were generated via lentiviral transduction (ExpiCHO-S and HEK293T expressing ACE2 (HEK293T-ACE2)). None of the cell lines used was authenticated. Cell lines were routinely tested for mycoplasma contamination.

Sample Donors

Samples from an individual who had recovered from COVID-19 (designated as donor X (a 52-year-old man)) was obtained 75 days after symptom onset under study protocols approved by the local Institutional Review Boards (Canton Ticino Ethics Committee, Switzerland). Samples from additional nine individuals who had recovered from COVID-19 and nine vaccinated individuals were obtained under study protocols approved by the Canton Ticino Ethics Committee, Switzerland and the Ethical committee of Luigi Sacco Hospital, Milan. Samples from individuals who had recovered from COVID-19 were collected between 14 and 75 days after symptom onset; samples from vaccinated individuals were collected between 13 and 20 days after the second dose of an mRNA-based SARS-CoV-2 vaccine. All donors provided written informed consent for the use of blood and blood components (such as peripheral blood mononuclear cells (PBMCs), sera or plasma).

Isolation of PBMCs, plasma and sera

PBMCs were isolated from blood draw performed using tubes prefilled with heparin, followed by Ficoll density gradient centrifugation. PBMCs were either used freshly along SARS-CoV-2 S-protein-specific memory B cell sorting or stored in liquid nitrogen for later use. Sera were obtained from blood collected using tubes containing clot activator, followed by centrifugation and stored at −80 °C.

B cell isolation and recombinant S2X259 mAb production

Starting from freshly isolated PBMCs or upon cells thawing, B cells were enriched by staining with CD19 PE–Cy7 and incubation with anti-PE beads, followed by

positive selection using LS columns. Enriched B cells were stained with anti-IgM, anti-IgD, anti-CD14 and anti-IgA, all PE labelled, and prefusion SARS-CoV-2 S with a biotinylated avi tag conjugated to streptavidin Alexa-Fluor 647 (Life Technologies). SARS-CoV-2 S-specific IgG^+ memory B cells were sorted by flow cytometry via gating for PE-negative and Alexa-Fluor 647-positive cells. Sorted IgG^+ memory B cells were seeded in 384-well microtitre plates on a monolayer of mesenchymal stromal cells at 0.5 cells per well in the presence CpG2006, IL-2, IL-6 and IL-21, as previously described (Pinto, 2019). After 7 days, the supernatants were screened for binding to SARS-CoV-2 RBD and prefusion S as well as for neutralization activity using a high-throughput VSV SARS-CoV-2 S-based microneutralization assay. Selection of the antibody of interest in secondary screening was based on reactivity to a panel of sarbecovirus RBDs as well as on the neutralizing activity. Antibody VH and VL sequences were obtained by RT–PCR and mAbs were expressed as recombinant human Fab fragment or as IgG1 (G1m3 allotype). ExpiCHO cells were transiently transfected with heavy and light chain expression vectors as previously described(Pinto, 2020).

Affinity purification was performed on ÄKTA Xpress FPLC (Cytiva) operated by UNICORN software version 5.11 (build 407) using HiTrap Protein A columns (Cytiva) for full length human and hamster mAbs and CaptureSelect CH1-XL MiniChrom columns (ThermoFisher Scientific) for Fab fragments, using PBS as mobile phase. Buffer exchange to the appropriate formulation buffer was performed with a HiTrap Fast desalting column (Cytiva). The final products were sterilized by filtration through 0.22-µm filters and stored at 4 °C.

Sarbecovirus sequence analysis and SARS-CoV-2 mutant counts

Alignment and phylogenetic tree of the strains within the *Sarbecovirus* subgenus was generated using MEGA X and CLC Main work bench 21.0.3 (Qiagen). The following sequence was retrieved from GISAID and NCBI: A021 (AAV97986.1). The viral sequences were obtained from GISAID EpiCoV project (https://www.gisaid.org/). Analysis was performed on sequences submitted to GISAID up to 2 April 2021. The S protein sequences were either obtained directly from the protein dump provided by GISAID or, for the latest submitted sequences that were not yet incorporated in the protein dump on the day of data retrieval, from the genomic sequences with exonerate(Slater, 2005) 2 2.4.0–haf93ef1_3 (https://quay.io/repository/biocontainers/exonerate?tab=tags) using protein to DNA alignment with parameters -m protein2dna --refine full --minintron 999999 --percent 20 and using accession YP_009724390.1 as a reference. Multiple sequence alignment of all human S proteins was performed with mafft50 7.475–h516909a_0 (https://quay.io/repository/biocontainers/mafft?tab=tags) with parameters --auto --reorder --keeplength --addfragments using the same reference as above. S sequences that contained more than 10% ambiguous amino acids or that were less than 80% of the canonical protein length were discarded. A total of 923,686 sequences were used for analysis. Variants were then extracted as compared to the reference with R 4.0.2 (https://www.r-project.org/) using Biostrings 2.56.0.

Binding to cell-surface-expressed sarbecovirus S proteins by flow cytometry

ExpiCHO-S cells were seeded at 6 × 10^6 cells per ml in a volume of 5 ml in a 50-ml bioreactor. S-encoding plasmids HKU3 (QND76020.1), Rs3367 (AGZ48818.1),

YN2013 (AIA62330.1), Rs4874 (ATO98205.1), Rs4255 (ATO98193.1), Rs4231 (ATO98157.1), Rs4084 (ATO98132.1), ZXC21 (AVP78042.1), ZC45 (AVP78031.1), Rp/Shaanxi2011 (AGC74165.1), Rm1/2004 (ABD75332.1), Rf1-2004 (ABD75323.1), BM48-31/BGR/2008 (YP_003858584.1), RaTG13 (QHR63300.2), PC4-127 (AAU93318.1), SARS-CoV-2 (YP_009724390.1), LYRa3 (AHX37569.1), Cp/Yunnan2011 (AGC74176.1), SARS-CoV Urbani (AAP13441.1), As6526 (ATO98108.1), BtKY72 (APO40579.1), RmYN02 (EPI_ISL_412977), Pangolin_Guangdong-2019 (EPI_ISL_410721), Pangolin-Guanxi-2017 (EPI_ISL_410539) and HKU3-12 (ADE34812.1) were diluted in cold OptiPRO SFM (Life Technologies, 12309-050), mixed with ExpiFectamine CHO Reagent (Life Technologies, A29130) and added to the cells. Transfected cells were then incubated at 37 °C with 8% CO_2 with an orbital shaking speed of 250 r.p.m. (orbital diameter of 25 mm) for 42 h. Transiently transfected ExpiCHO cells were collected and washed two times in wash buffer (PBS 2% FBS, 2 mM EDTA). Cells were counted and distributed into round-bottom 96-well plates (Corning, 3799) and incubated with S2X259 mAb serial dilutions from 50 μg ml^{-1}. Alexa Fluor647-labelled goat anti-human IgG secondary antibody (Jackson Immunoresearch, 109-606-098) was prepared at 2 μg ml^{-1} and added onto cells after two washing steps. Cells were then washed twice and resuspended in wash buffer for data acquisition at ZE5 cytometer (Biorad).

Recombinant protein production

Wild-type SARS-CoV-2 RBD (with N-terminal signal peptide and 'ETGT', and C-terminal 8×His tag) was expressed in Expi293F cells at 37 °C and 8% CO_2 in the presence of 10 μM kifunensine. Transfection was performed using the ExpiFectamine

293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatant was collected four days after transfection and supplemented with 10× PBS to a final concentration of 2.5× PBS (342.5 mM NaCl, 6.75 mM KCl and 29.75 mM phosphates). For crystallization, wild-type SARS-CoV-2 RBD was purified using a 5-ml HisTALON superflow cartridge (Takara Bio) followed by size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column pre-equilibrated in 20 mM Tris-HCl pH 7.5, 150 mM NaCl. RBD was deglycosylated with EndoH and mixed with a 1.3-fold molar excess of S2X259 Fab and S2H97 Fab. The complex was purified on a Superdex 200 Increase 10/300 GL column pre-equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl. SARS-CoV-2 S hexapro protein, used for cryo-electron microscopy (cryo-EM) single particle studies, was expressed and purified as previously described (Tortorici, 2020).

ELISA

Ninety-six half area well-plates (Corning, 3690) were coated overnight at 4 °C with 25 µl of sarbecovirus RBD proteins WIV1 (AGZ48831.1), Anlong-112 (ARI44804.1), YN2013 (AIA62330.1), SC2018 (QDF43815.1), ZC45 (AVP78031.1), Rp/Shaanxi2011 (AGC74165.1), BM48-31/BGR/2008 (YP_003858584.1), RaTG13 (QHR63300.2), SARS-CoV-2 (YP_009724390.1), SARS-CoV Urbani (AAP13441.1), BtKY72 (APO40579.1), Pangolin_Guangdong-2019 (EPI_ISL_410721) and SARS-CoV-2 RBD mutants, prepared at 5 µg ml^{-1} in PBS pH 7.2. Plates were then blocked with PBS 1% BSA (Sigma-Aldrich, A3059) and subsequently incubated with mAb serial dilutions for 1 h at room temperature. After 4 washing steps with PBS 0.05% Tween 20 (PBS-T) (Sigma-Aldrich, 93773), goat anti-human IgG secondary antibody (Southern

Biotech, 2040-04) was added and incubated for 1 h at room temperature. Plates were then washed four times with PBS-T and 4-nitrophenyl phosphate (pNPP, Sigma-Aldrich, 71768) substrate was added. After 30 min incubation, absorbance at 405 nm was measured by a plate reader (Biotek) and data were plotted using Prism GraphPad 9.1.0.

S2X259 blockade-of-binding to RBD

S2X259 and S2H14 mAbs were biotinylated using EZ-Link NHS-PEG solid phase biotinylation kit (ThermoFisher Scientific) and binding to RBD was tested to set optimal concentration to be used in the assay after sample desalting using Zeba Spin Desalting Columns (ThermoFisher Scientific). Half-area 96-well-plates were coated overnight at 4 °C with SARS-CoV-2 RBD–mouse Fc Tag diluted at 1 μg ml^{-1} in PBS. After a blocking step with Blocker Casein (ThermoFisher Scientific), serial plasma dilutions in Blocker Casein were incubated 30 min at room temperature. Biotinylated S2X259 or S2H14 were added at a concentration achieving 80% of maximal binding and the mixture was incubated for 30 min at room temperature. Alkaline-phosphatase conjugated streptavidin (Jackson ImmunoResearch) was diluted at 0.5 μg ml^{-1} in Blocker Casein and added on plates previously washed 4 times with PBS 0.05% Tween 20. After 45 min incubation, plates were washed and 4-nitrophenyl phosphate substrate incubated for 1 h at room temperature. Absorbance at 405 nm was measured and percentage of inhibition was calculated as follows: (1 − (optical density (OD) sample − OD negative control)/(OD positive control − OD negative control)) × 100.

MLV-based pseudotyped virus production and neutralization

To generate SARS-CoV-2 S MLV pseudotyped virus, HEK293T cells were seeded in 10-cm dishes in DMEM supplemented with 10% FBS. The next day cells were transfected with a SARS-CoV-2 S glycoprotein-encoding plasmid containing the D19 C-terminal truncation (Katoh, 2013), an MLV Gag-Pol packaging construct and the reporter vector pTG-Luc, using the X-tremeGENE HP DNA transfection reagent (Roche) according to the manufacturer's instructions. Cells were then incubated at 37 °C with 5% CO_2 for 72 h. Supernatant was collected and cleared from cellular debris by centrifugation at 400*g*, and stored at −80 °C.

For neutralization assays, Vero E6 cells were seeded into white 96-well plates (PerkinElmer) at 20,000 cells per well and cultured overnight at 37 °C with 5% CO_2 in 100 µl DMEM supplemented with 10% FBS and 1% penicillin–streptomycin. The next day, MLV-SARS-CoV-2 pseudovirus was activated with 10 µg ml^{-1} TPCK-treated trypsin (Worthington Biochem) for 1 h at 37 °C. Recombinant antibodies at various concentrations were incubated with activated pseudovirus for 1 h at 37 °C. The Vero E6 cells were then washed with DMEM, and 50 µl of pseudovirus–mAbs mixes were added and incubated for 2 h at 37 °C with 5% CO_2. After incubation, 50 µl of DMEM containing 20% FBS and 2% penicillin–streptomycin were added and the cells were incubated 48 h at 37 °C with 5% CO_2. Following these 48 h of infection, culture medium was removed from the cells and 50 µl per well of Bio-Glo (Promega) diluted 1:2 with PBS with $Ca^{2+}Mg^{2+}$ (Thermo Fisher) was added to the cells and incubated in the dark for 15 min before reading on a Synergy H1 Hybrid Multi-Mode plate reader (Biotek). Measurements were done in duplicate, Relative light unit (RLU) values were converted

to percentage of neutralization and plotted with a nonlinear regression curve fit in GraphPad Prism 9.1.0.

VSV-based pseudotype virus production and neutralization assay

SARS-CoV-2 S (CAD0240757.1), RaTG13 S (QHR63300.2), Pangolin-Guangdong S (QLR06867.1), Pangolin-Guanxi S (QIA48623.1), SARS-CoV S (YP 009825051.1), WIV1 S (AGZ48831.1), WIV16 S (ALK02457.1), RsSHCO14 S (AGZ48806.1), the variant of concern B.1.429 S (QTC60823) and the variants of concern, N437K, Y453F, B.1.1.7 S, B.1.351 S and P.1 S with their corresponding mutations inserted in the SARS-CoV-2 S (CAD0240757.1) backbone were used to pseudotype VSV. Pseudotyped viruses were prepared using 293T cells seeded in 10-cm dishes. In brief, cells in DMEM supplemented with 10% FBS, 1% penicillin–streptomycin were transfected with the plasmid encoding for the corresponding S glycoprotein using lipofectamine 2000 (Life Technologies) following manufacturer's indications. One day after transfection, cells were infected with VSV (G*ΔG-luciferase) and after 2 h, infected cells were washed 4 times with DMEM before adding medium supplemented with anti-VSV-G antibody (I1- mouse hybridoma supernatant diluted 1 to 50, from CRL- 2700, ATCC). Particles were collected 18 h after inoculation, clarified from cellular debris by centrifugation at 2,000*g* for 5 min and concentrated 10 times using a 30-kDa cut-off membrane and used for neutralization experiments, aliquoted and frozen at −80 °C until use in neutralization experiments.

For neutralization, stable 293T cells expressing ACE252 in DMEM supplemented with 10% FBS, 1% penicillin–streptomycin were seeded at 40,000 cells per well into clear-bottom white-walled 96-well plates and cultured overnight at 37 °C. Twelve-point

threefold serial dilutions of the corresponding mAb were prepared in DMEM and pseudotyped VSV viruses were added 1:1 to each mAb dilution in the presence of anti-VSV-G antibody from I1- mouse hybridoma supernatant diluted 50 times. After 45 min incubation at 37 °C, 40 µl of the mixture was added to the cells and 2 h after infection, 40 µl DMEM was added to the cells. After 17–20 h 50 µl per well of One-Glo-EX substrate (Promega) was added to the cells and incubated in the dark for 5–10 min before reading on a Varioskan LUX plate reader (ThermoFisher). Measurements were done in duplicate with two independent productions of pseudotyped viruses and RLU values were converted to percentage of neutralization and plotted with a nonlinear regression curve fit in GraphPad Prism 9.1.0.

Generation of VSV SARS-CoV-2 S pseudotypes from a selection of circulating variants

Replication-defective VSV pseudovirus expressing SARS-CoV-2 S protein were generated as previously described (Takada, 1997) with some modifications. Plasmids encoding the SARS-CoV-2 S glycoprotein corresponding to the variants of concern used to produce SARS-CoV-2-VSV were obtained using a multistep overlap extension PCR protocol (McCallum, 2021; Collier 2021). In brief, the mutations of the different variant of concern lineages were encoded on each primer pair used to amplify sequential, overlapping fragments of the SARS-CoV-2 S encoding DNA sequence from a pCDNA_SARS-CoV-2_S D19 plasmid encoding the C-terminally truncated S protein that has been shown to be better expressed at the cell (Case, 2020). The amplified overlapping fragments were separated on agarose gel and purified using Illustra GFX PCR DNA and Gel Band Purification Kit (Merck KGaA). The purified overlapping

fragments were joined by performing 10 cycles of PCR without the addition of primers, followed by 30 PCR cycles using a single pair of external primers carrying 5′ overhangs homologous to the vector backbone. The final PCR product was then cloned into the pCDNA3 vector using the Takara In-fusion HD cloning kit following manufacturer's instructions.

Lenti-X 293T cells (Takara, 632180) were seeded in 10-cm dishes at a density of 5×10^6 cells per cm^2 and the following day transfected with 10 µg of S expression plasmid with TransIT-Lenti (Mirus, 6600) according to the manufacturer's instructions. One day after transfection, cells were infected with VSV-luc (VSV-G) (Kerafast, EH1020-p.m.) for 1 h, rinsed three times with PBS, then incubated for an additional 24 h in complete medium at 37 °C. The cell supernatant was clarified by centrifugation, filtered (0.45 µm), aliquoted and frozen at −80 °C.

For VSV pseudovirus neutralization assays, Vero-E6 cells were grown in DMEM supplemented with 10% FBS and seeded into CulturePlate-96, White Opaque 96-well Microplate (PerkinElmer) at a density of 2×10^4 cells per well. The next day, mAbs were serially diluted in prewarmed complete medium, mixed at 1:1 ratio with pseudovirus and incubated for 1 h at 37 °C in round-bottom polypropylene plates. Medium from cells was aspirated and 50 µl of virus–mAb complexes were added to cells and then incubated for 1 h at 37 °C. An additional 100 µl of prewarmed complete medium was then added on top of complexes and cells incubated for an additional 16–24 h. Conditions were tested in duplicate wells.

Virus–mAb-containing medium was then aspirated from cells and 50 µl of a 1:2 dilution of SteadyLite Plus (Perkin Elmer) in PBS with $Ca^{2+}Mg^{2+}$ was added to cells.

Plates were incubated for 10 min at room temperature and then analysed on the Synergy H1 Hybrid Multi-Mode reader (Biotek). Measurements were done in two independent experiments and RLU values were converted to percentage of neutralization and plotted with a nonlinear regression curve fit with an upper constraint equal to 100 and a lower constraint equal to 0 in Prism (version 9.1.0). Half-maximal inhibitory concentration (IC_{50}) values across biological replicates are presented as arithmetic mean ± s.d.

Neutralization of SARS-CoV-2 B.1.351 pseudotyped VSV by S2X259 + S309 antibody cocktail

Vero E6 cells were plated into white 96-well plates (PerkinElmer) at 20,000 cells per well and cultured overnight at 37 °C 5% CO_2 in 100 µl culture medium (DMEM 10% FBS 1% penicillin–streptomycin). The next day, S2E12 and S309 antibodies were serially diluted in infection medium (DMEM 10% heat-inactivated FBS 1% penicillin–streptomycin) and mixed in a matrix format before incubation with SARS-CoV-2 B.1.351 pseudovirus (multiplicity of infection (MOI) 0.1) at 37 °C 5% CO_2 for 1 h. Cell culture medium was aspirated and 50 µl of mAbs–pseudovirus mix added onto cells. After incubation at 37 °C 5% CO_2 for 1 h, 100 µl of infection medium was added to cell and incubation continued for the following 20 h. For results read out, medium was removed and 50 µl per well of Steadylite PLUS (PerkinElemer) 1:2 diluted in PBS + Ca^{2+} and Mg^{2+} was added and incubated 15 min at room temperature in the dark. Luminescence signals were measured using a Synergy H1 reader (Biotek) with 1-s integration time. Data were analysed using MacSynergy II and synergy plots obtained with 99.9% confidence were used for graphical elaboration with GraphPad Prism 9.1.0.

Neutralization of authentic SARS-CoV-2-Nluc virus

Neutralization of authentic SARS-CoV-2 by entry-inhibition assay was determined using SARS-CoV-2-Nluc, an infectious clone of SARS-CoV-2 (based on strain 2019-nCoV/USA_WA1/2020) that encodes nanoluciferase in place of the viral ORF7 and demonstrated comparable growth kinetics to wild-type virus (Xie, 2020). Vero E6 cells were seeded into black-walled, clear-bottom 96-well plates at 2×10^4 cells per well and cultured overnight at 37 °C. The next day, 9-point fourfold serial dilutions of mAbs were prepared in infection medium (DMEM + 10% FBS). SARS-CoV-2-Nluc was diluted in infection medium at a final MOI of 0.01 plaque-forming units per cell, added to the mAb dilutions and incubated for 30 min at 37 °C. Medium was removed from the Vero E6 cells, mAb–virus complexes were added and incubated at 37 °C for 24 h. Medium was removed from the cells, Nano-Glo luciferase substrate (Promega) was added according to the manufacturer's recommendations, incubated for 10 min at room temperature and the luciferase signal was quantified on a VICTOR Nivo plate reader (Perkin Elmer).

Affinity determination by surface plasmon resonance

Surface plasmon resonance binding measurements were performed using a Biacore T200 instrument, running buffer was Cytiva HBS-EP+ (pH 7.4), and all measurements were performed at 25 °C. For Fab binding, anti-AviTag pAb (for capturing S proteins) or StrepTactin XT (for capturing RBDs) were covalently immobilized on CM5 chips. S2X259 Fab concentrations were 11, 33, 100, and 300 nM run as single-cycle kinetics. Double reference-subtracted data were fit to a 1:1 binding model using Biacore Evaluation software, which yields an 'apparent KD' ($K_{D,app}$) for the

S-binding data because the kinetics also reflect S conformational dynamics. For SARS-CoV-2 S the dissociation rate was too slow to fit, so the $K_{D,app}$ is reported as an upper limit. The K_D above 1 μM is approximate and was determined from a fit in which R_{max} was set to a constant based on theoretical R_{max}.

Affinity determination using biolayer interferometry

Biotinylated RBD (wild type, N501Y, K417N-E484K-N501Y, or K417T-E484K-N501Y) were immobilized at 5 ng μl^{-1} in undiluted 10X Kinetics Buffer (Pall) to SA sensors until a load level of 1.1 nm. A 1:3 dilution series of Fab in undiluted kinetics buffer starting at 10 nM was used for 600 s association before 600 s dissociation to determine protein–protein affinity. The data were baseline-subtracted and the plots fitted using the Pall FortéBio/Sartorius analysis software (version 12.0). Data were plotted in Prism 9.1.0.

Competition assay by biolayer interferometry

Biolayer interferometry was used to assess S2X259 competition with S309 and S2E12 using an Octet Red96 (ForteBio). All reagents were prepared in kinetics buffer (PBS 0.01% BSA) at the indicated concentrations. His-tagged SARS-CoV-2 RBD was prepared at 8 μg ml^{-1} and loaded on prehydrated anti-penta-HIS biosensors (Sartorius) for 2.5 min. Biosensors were then moved into a solution containing S2X259 mAb and association recorded for 7 min. A second association step was subsequently performed into S2X259 (as control), S309 and S2E12 mAbs solutions at 20 μg ml^{-1} and recorded for 7 min. Response values were exported and plotted using GraphPad Prism 9.1.0.

Cell-surface mAb-mediated S1 shedding

CHO cells stably expressing wild-type SARS-CoV-2 S were resuspended in wash buffer (PBS 1% BSA, 2 mM EDTA) and treated with 10 μg ml^{-1} TPCK-trypsin (Worthington Biochem) for 30 min at 37 °C. Cells were then washed and distributed into round bottom 96-well plates (90,000 cells per well). S2X259 was added to cells at 15 μg ml^{-1} final concentration for 180 min at 37 °C. Cells were collected at different time points (5, 30, 60, 120 and 180 min), washed with wash buffer at 4 °C, and incubated with 1.5 μg ml^{-1} secondary goat anti-human IgG, Fc fragment specific (Jackson ImmunoResearch) on ice for 20 min. Cells were washed and resuspended in wash buffer and analysed with ZE5 FACS (Biorad).

Measurement of Fc-effector functions

S2X259-dependent activation of human FcγRIIIa was performed with a bioluminescent reporter assay. ExpiCHO cells transiently expressing full-length wild-type SARS-CoV-2 S (target cells) or full-length prefusion stabilized SARS-CoV-2 S, which contains the 2P mutation and S1/S2 furin cleavage site mutation (RRARS to SGAG) as previously described

(Walls, 2020), were incubated with different amounts of mAbs. After a 15-min incubation, Jurkat cells stably expressing FcγRIIIa receptor (V158 variant) or FcγRIIa receptor (H131 variant) and NFAT-driven luciferase gene (effector cells) were added at an effector to target ratio of 6:1 for FcγRIIIa and 5:1 for FcγRIIa. Signalling was quantified by the luciferase signal produced as a result of NFAT pathway activation. Luminescence was measured after 20 h of incubation at 37 °C with 5% CO_2 with a

luminometer using the Bio-Glo Luciferase Assay Reagent according to the manufacturer's instructions (Promega).

In vivo mAb testing using a Syrian hamster model

KU LEUVEN R&D has developed and validated a SARS-CoV-2 Syrian hamster infection model (Boudewijns, 2020; Abdelnabi, 2021). The SARS-CoV-2 (BetaCov/Belgium/GHB-03021/2020-EPI ISL 109 407976|2020-02-03), which is most closely related to the prototypic Wuhan-Hu-1 2019-nCoV (GenBank accession number MN908947.3) strain, and B.1.351 (hCoV105 19/Belgium/rega-1920/2021; EPI_ISL_896474, 2021-01-11) isolates used in this study were recovered from nasopharyngeal swabs taken from an asymptomatic returning traveller with RT–qPCR-confirmed SARS-CoV-2 infection in February 2020, and from a patient with respiratory symptoms in January 2021, respectively. A close relatedness with the prototypic Wuhan-Hu-1 2019 SARS-CoV-2 and with B.1.351 lineage was confirmed by sequencing and phylogenetic analysis. Infectious viruses were isolated by serial passaging on Vero E6 cells and passage 6 for SARS-CoV-2 Wuhan-Hu-1 and passage 2 for B.1.351 viruses were used for the study. The titre of the virus stock was determined by end-point dilution on Vero E6 cells by a previously published method (Reed, 1938). This work was conducted in the high-containment A3 and BSL3+ facilities of the KU Leuven Rega Institute (3CAPS) under licences AMV 30112018 SBB 219 2018 0892 and AMV 23102017 SBB 219 20170589 according to institutional guidelines.

Syrian hamsters were purchased from Janvier Laboratories and were housed per two in ventilated isolator cages (IsoCage N Biocontainment System, Tecniplast) with ad libitum access to food and water and cage enrichment (wood block). Housing conditions

and experimental procedures were approved by the ethical committee of animal experimentation of KU Leuven (licence P065-2020). Hamsters were randomized and the sample size was chosen on the basis of previous experience with this animal model. To evaluate the prophylactic efficacy 6–10-week-old female hamsters were administered by intraperitoneal injection with S2X259 mAb at 1, 4, 5 and 25 mg kg^{-1} or a combination of S2X259 and S309 mAbs (1 mg kg^{-1} + 1 mg kg^{-1}) 48 h before intranasal infection with 1.89×10^6 $TCID_{50}$ in 50 µl inoculum. Blood samples were collected before infection and serum was obtained for pharmacokinetics analysis. To assess the therapeutic efficacy, 6–8-week-old Syrian hamsters were administered by intraperitoneally with isotype control antibody (20 mg kg^{-1}) or S2X259 mAb (at 20, 10 or 5 mg kg^{-1}) 24 h after infection with SARS-CoV-2. Blood samples were collected when hamsters were killed, and serum was obtained for pharmacokinetics analysis. Hamsters were monitored for appearance, behaviour and weight. At day 4 after infection, hamsters were euthanized by intraperitoneal injection of 500 µl Dolethal (200 mg ml^{-1} sodium pentobarbital, Vétoquinol SA). Lungs were collected, homogenized using bead disruption (Precellys) in 350 µl RLT buffer (RNeasy Mini kit, Qiagen) and centrifuged (10,000 r.p.m., 5 min, 4 °C) to pellet the cell debris. RNA was extracted using a NucleoSpin kit (Macherey-Nagel) according to the manufacturer's instructions. RT–qPCR was performed on a LightCycler96 platform (Roche) using the iTaq Universal Probes One-Step RTqPCR kit (BioRad) with N2 primers and probes targeting the nucleocapsid[36]. Standards of SARS-CoV-2 cDNA (IDT) were used to express viral genome copies per mg tissue or per ml serum. To quantify infectious SARS-CoV-2 particles, end-point titrations were performed on confluent Vero E6 cells in 96-well

plates. The hamster-derived samples for RNA quantification and viral load titration were run by technicians who were blinded to treatment group of the analysed samples. Viral titres were calculated by a previously published method(Reed, 1938) and were expressed as $TCID_{50}$ per mg tissue.

Evaluation of escape mutants using deep-mutational scanning

A previously described deep-mutational scanning approach(Greaney, 2021) was used to identify RBD mutations that escape S2X259 binding. In brief, duplicate libraries containing virtually all possible amino acid changes compatible with ACE2 binding and RBD folding within the Wuhan-Hu-1 SARS-CoV-2 RBD sequence were expressed on the surface of yeast (Starr 2020; Greaney, 2021). Libraries were labelled at 59 ng ml^{-1} S2X259 antibody (the 90% effective concentration for binding to yeast-displayed SARS-CoV-2 RBD determined in isogenic pilot binding experiments), and fluorescence-activated cell sorting (FACS) was used to select RBD^{+} cells that exhibit reduced antibody binding, as previously described(Greaney, 2021; Starr, 2021) . Libraries were sequenced before and after selection to determine per-mutation escape fractions, as previously described (Starr, 2021). Experiments were performed in duplicate with independently generated mutant libraries, and we report the average mutant escape fraction across the duplicates. Raw escape fractions are available on GitHub: https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Vir_mAbs/blob/main/results/supp_data/s2X259_raw_data.csv. Detailed steps of analysis and code for the deep mutational scanning selections are available on GitHub: https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Vir_mAbs.

Selection of SARS-CoV-2 mAb escape mutants

VSV-SARS-CoV-2 chimera was used to select for SARS-CoV-2 S mAb-resistant mutants, as previously described (Case, 2020, Liu, 2021). In brief, mAb-resistant mutants were recovered by plaque isolation on Vero cells with the indicated mAb in the overlay. The concentration of mAb in the overlay was determined by neutralization assays at a MOI of 100. Escape clones were plaque-purified on Vero cells in the presence of mAb, and plaques in agarose plugs were amplified on MA104 cells with the mAb present in the medium. Viral stocks were amplified on MA104 cells at an MOI of 0.01 in Medium 199 containing 2% FBS and 20 mM HEPES pH 7.7 (Millipore Sigma) at 34 °C. Viral supernatants were collected upon extensive cytopathic effect and clarified of cell debris by centrifugation at 1,000*g* for 5 min. Aliquots were maintained at −80 °C. Viral RNA was extracted from VSV-SARS-CoV-2 mutant viruses using RNeasy Mini kit (Qiagen), and S was amplified using OneStep RT–PCR Kit (Qiagen). The mutations were identified by Sanger sequencing (GENEWIZ). Their resistance was verified by subsequent virus infection in the presence or absence of antibody. In brief, Vero cells were seeded into 12-well plates for overnight. The virus was serially diluted using DMEM and cells were infected at 37 °C for 1 h. Cells were cultured with an agarose overlay in the presence or absence of mAb at 34 °C for 2 days. Plates were scanned on a biomolecular imager and expression of eGFP is show at 48 h after infection.

Crystallization, data collection, structure determination and analysis

Crystals of the SARS-CoV-2-RBD–S2X259–S2H97 Fab complex were obtained at 20 °C by sitting drop vapour diffusion. A total of 200 nl complex at 5.7 mg ml^{-1} was mixed with 200 nl mother liquor solution containing 0.12 M monosaccharides mix, 20%

(v/v) ethylene glycol, 10% (w/v) PEG 8000, 0.1 M Tris (base)/bicine pH 8.5, 0.02 M sodium chloride, 0.01 M MES pH 6 and 3% (v/v) Jeffamine ED-2003. Crystals were flash-frozen in liquid nitrogen. Data were collected at Beamline 9-2 of the Stanford Synchrotron Radiation Lightsource facility. Data were processed with the XDS software package (Kabsch, 2010) for a final dataset of 2.65 Å in space group P21. The RBD–S2X259–S2H97 Fab complex structure was solved by molecular replacement using phaser (McCoy, 2007 from a starting model consisting of SARS-CoV-2 RBD (PDB code 7JX3) and homology models of the S2X259 and S2H97 Fabs built using the Molecular Operating Environment (MOE) software package from the Chemical Computing Group (https://www.chemcomp.com). Several subsequent rounds of model building and refinement were performed using Coot(Emsley), ISOLDE(Croll, 2018), Refmac5 (Murshudov) and MOE (https://www.chemcomp.com), to arrive at a final model for the ternary complex.

Cryo-EM sample preparation, data collection and data processing

Recombinantly expressed and purified Fab S2X259 and SARS-CoV-2 S hexapro were incubated at 1 mg ml^{-1} with a 1.2 molar excess of Fab at 4 °C during 1 h. Three microlitres of the complex mixture were loaded onto freshly glow discharged R 2/2 UltrAuFoil grids (200 mesh) or lacey grids covered with a thin layer of manually added carbon, before plunge-freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 7–7.5 s blot time (for the UltrAuFoil grids) or with a blot force of −1 and 2.5 s blot time (for the lacey thin carbon grids) at 100% humidity and 21 °C.

Data were acquired on a FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K2 Summit direct detector and Gatan

Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon (Suloway, 2005) at a nominal magnification of 130,000× with a super-resolution pixel size of 0.525 Å. The dose rate was adjusted to 8 counts per pixel per s, and each movie was fractionated in 50 frames of 200 ms. Two datasets were collected from UltrAuFoil grids with the stage tilted 30° and 55° to circumvent particle preferential orientation. The third dataset was collected on lacey grids covered with a thin layer of carbon. The three datasets, with a total of 6,786 micrographs, were collected with a defocus range of between −0.8 and −2 µm. For each dataset, movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking and extraction were carried out using Warp (Tegunov, 2019). Particle images were extracted with a box size of 800 pixels2 and binned to 400 yielding a pixel size of 1.05 Å. The three datasets were merged and two rounds of reference-free 2D classification were performed using cryoSPARC (Punjani). Subsequently, one round of 3D classification with 50 iterations, using PDB code 6VXX as initial model, was carried out using Relion (Zivanov, 2018; Kimanius ,2016) without imposing symmetry. Three-dimensional refinements were carried out using nonuniform refinement (Punjani, 2020). Particle images from each dataset were subjected to Bayesian polishing (Zivanov, 2019) using Relion before merging them to perform another round of nonuniform refinement in cryoSPARC followed by per-particle defocus refinement and again nonuniform refinement. To improve the density of the S–S2X259 interface, the particles were symmetry-expanded and subjected to a Relion focus 3D classification without refining angles and shifts using a soft mask encompassing the RBD and S2X259 variable domains. Particles belonging

to classes with the best-resolved local density were selected and subjected to local refinement using cryoSPARC. Local resolution estimation, filtering and sharpening were carried out using CryoSPARC. Reported resolutions are based on the gold-standard Fourier shell correlation of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (Chen, 2013).

Cryo-EM model building and analysis

UCSF Chimera (Pettersen, 2004) and Coot (Emsley, 2010) were used to fit atomic models (PDB code 6VXX or 6VYB) into the cryo-EM maps and the Fab variable domains were manually built. S2E12 was built in the locally refined map and subsequently validated using the Fab crystal structure. Models were refined and relaxed using Rosetta using both sharpened and unsharpened maps (Frenz, 2019; Wang, 2016). Validation used Phenix (Liebschner, 2010), Molprobity (Chen, 2010) and Privateer (Agirre,2015). Figures were generated using UCSF ChimeraX (Goddard, 2018).

Material availability

Materials generated in this study can be available on request and may require a material transfer agreement. The SARS-CoV-2 RBD mutant libraries (no. 1000000172) and unmutated parental plasmid (no. 166782) are available on Addgene

Data availability

All datasets generated and information presented in the study are available from the corresponding authors on reasonable request. The cryo-EM structure data and

model are available from the PDB (and Electron Microscopy Data Bank) under accession codes 7RA8 (EMD-24347) for the global refinement of the SARS-CoV-2 S and S2X259 complex and 7RAL (EMD-24365) for the local refinement of the SARS-CoV-2 S RBD and S2X259 Fab variable domains. The X-ray structure data and model has been deposited at PDB code 7M7W for RBD–S2H97–S2X259. Interactive escape maps and structural visualizations can be found at: https://jbloomlab.github.io/SARS-CoV-2-RBD_MAP_Vir_mAbs/. Raw Illumina sequencing data from deep mutational scanning experiments are available on NCBI SRA, BioSample SAMN18315604 (SARS-CoV-2 mutant selection data). A complete table of deep-mutational scanning antibody escape fractions is provided on GitHub: https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Vir_mAbs/blob/main/results/supp_data/all_antibodies_raw_data.csv.

Code availability

A repository containing all code, analysis and summary notebooks for the analysis of the SARS-CoV-2 deep-mutational scanning escape selections is available on GitHub at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_Vir_mAbs.

7.8 Figures and Tables

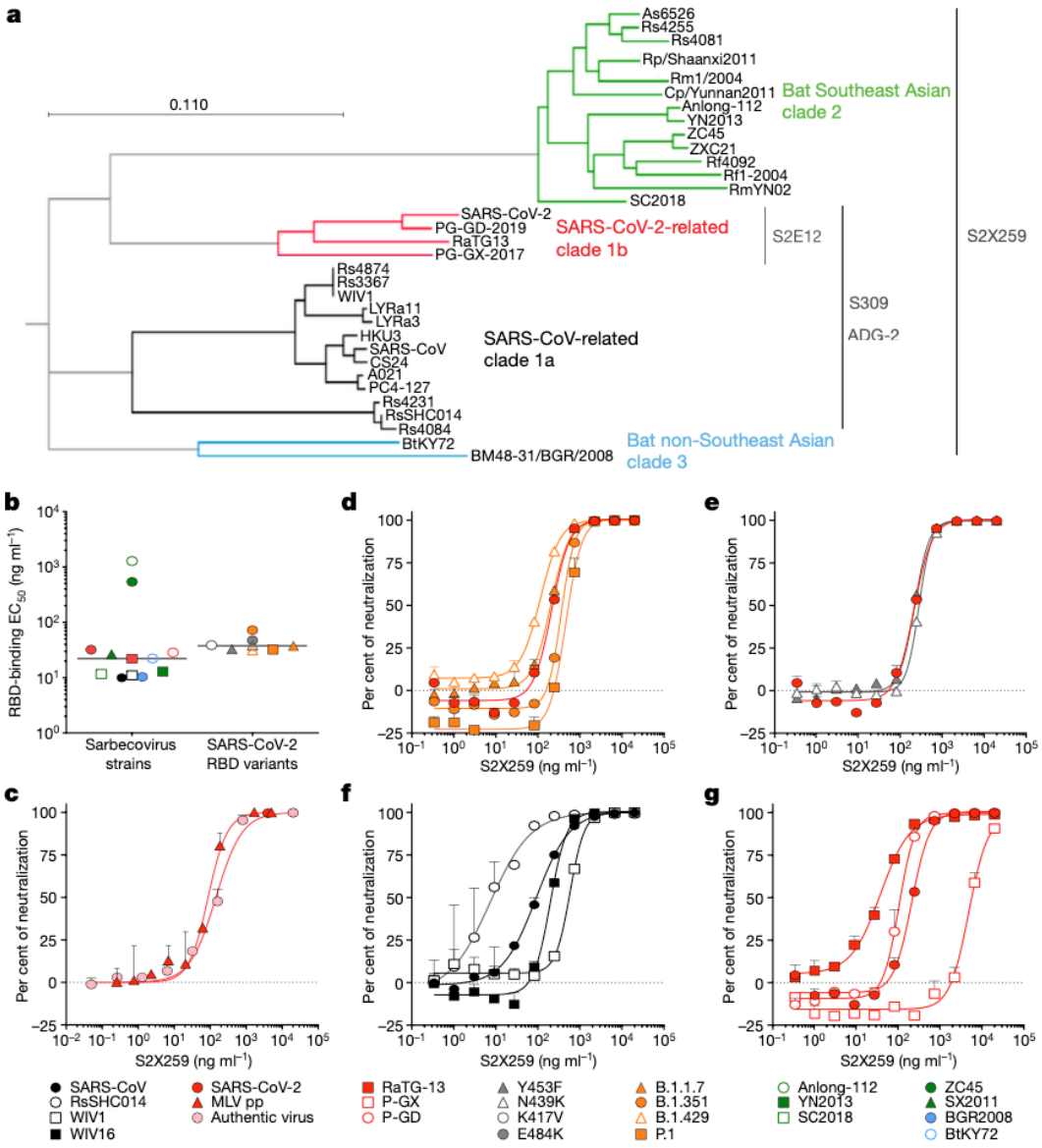


Figure 7.1: Identification of a broadly neutralizing sarbecovirus mAb.

a, Phylogenetic tree of sarbecovirus RBDs constructed using maximum likelihood analysis of amino acid sequences retrieved from GISAID and GenBank. Cross-reactivity within the *Sarbecovirus* subgenus for S2E1221, S30922 and ADG-220 is included for comparison. b, S2X259 binding to RBDs representative of the different sarbecovirus clades and SARS-CoV-2 variants as measured by ELISA. c, S2X259-mediated neutralization of SARS-CoV-2-Nluc authentic virus and SARS-CoV-2 S MLV-pseudotyped virus (MLV pp). d, e, S2X259-mediated neutralization of VSV pseudotypes containing SARS-CoV-2 S from isolates representing the B.1.1.7, B.1.351, P.1 and

B.1.429 variants of concern (d) as well as single RBD mutants (e). f, g, S2X259-mediated neutralization of VSV pseudotypes containing SARS-CoV-related (clade 1a) (f) or SARS-CoV-2-related (clade 1b) (g) S glycoproteins. *n* = 2 independent experiments. Error bars indicate s.d. of duplicates or triplicates.

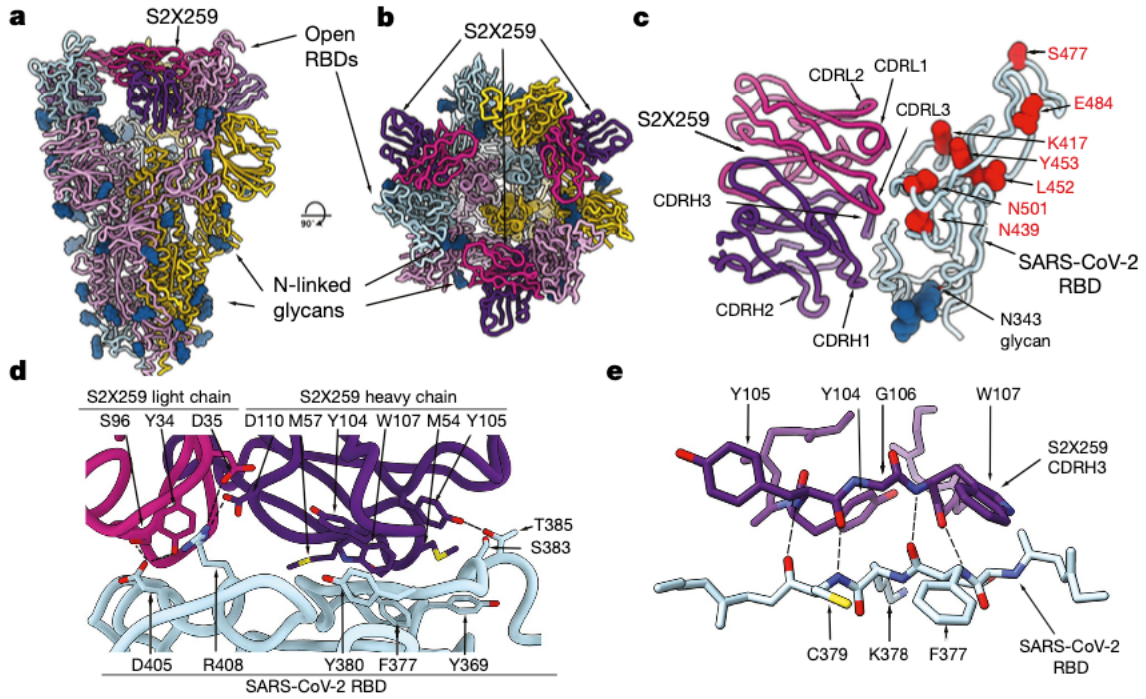


Figure 7.2: The S2X259 broadly neutralizing sarbecovirus mAb recognizes RBD antigenic site II.|

a, b Cryo-electron microscopy structure of the prefusion SARS-CoV-2 S ectodomain trimer with three S2X259 Fab fragments bound to three open RBDs, viewed along two orthogonal orientations (Protein Data Bank (PDB) codes 7RA8 and 7RAL). **c**, The S2X259 binding mode involving contacts with multiple RBD regions. Residues corresponding to prevalent RBD mutations are shown as red spheres (PDB code 7M7W). **d**, **e**, Close-up views showing selected interactions formed between S2X259 and the SARS-CoV-2 RBD. In **a**–**e**, each SARS-CoV-2 S protomer is coloured distinctly (cyan, pink and gold), and the S2X259 light and heavy chain variable domains are coloured magenta and purple, respectively. N-linked glycans are rendered as blue spheres in **a**–**c**.

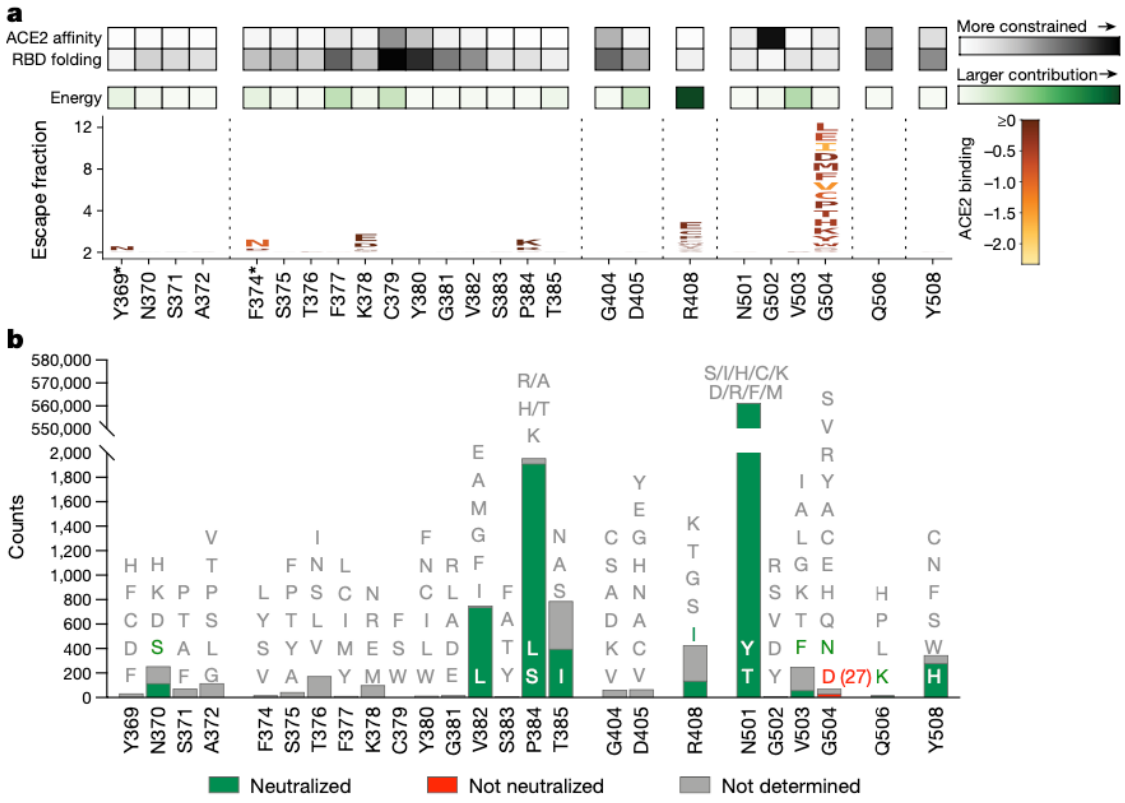


Figure 7.3: S2X259 is resilient to a broad spectrum of escape mutations.

a, Complete mapping of mutations reducing S2X259 binding using yeast-displayed RBD and deep-mutational scanning. Mean mutation effect on ACE2 affinity, RBD folding and contribution to S2X259 binding of substitutions at each position in the S2X259 epitope are reported. Asterisks denote mutations introducing N-linked glycosylation sites that may not be tolerated in full S. **b**, Frequency of mutants within the S2X259 epitope on the basis of SARS-CoV-2 genome sequences available on GISAID as of May 2021. S2X259 neutralizing activity against selected mutations is reported.

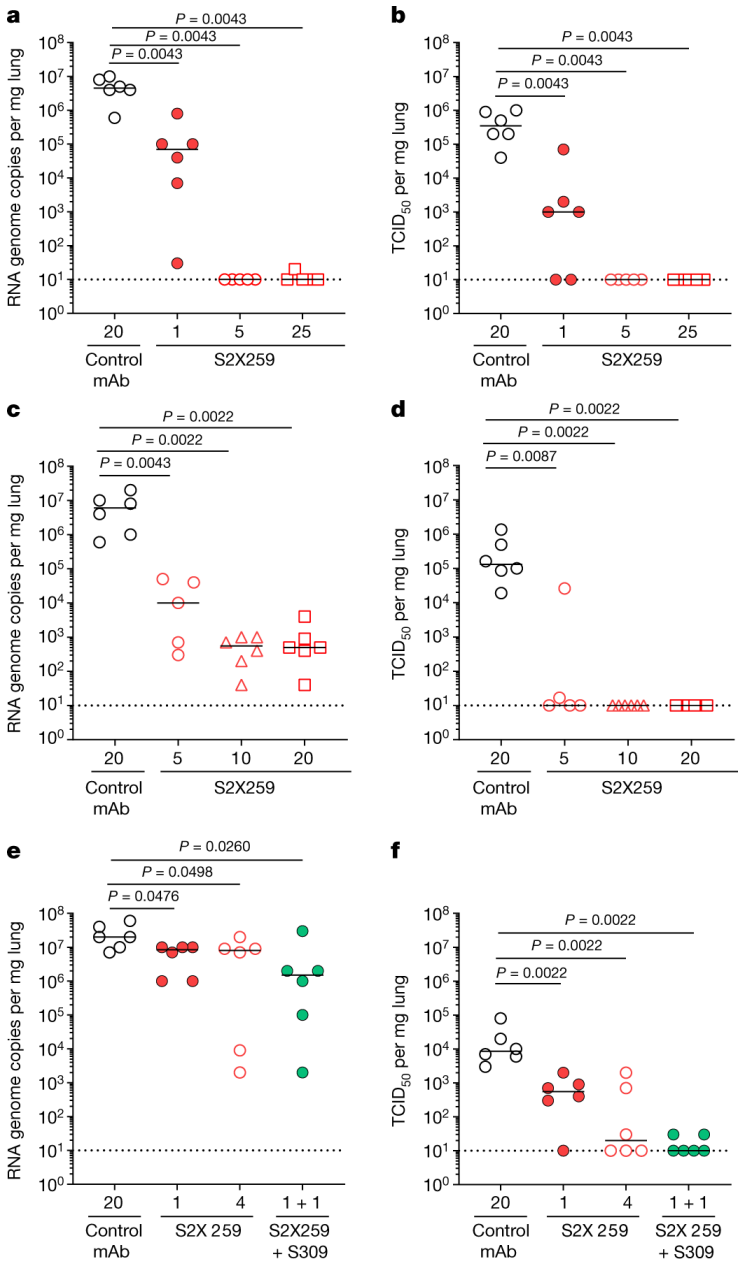
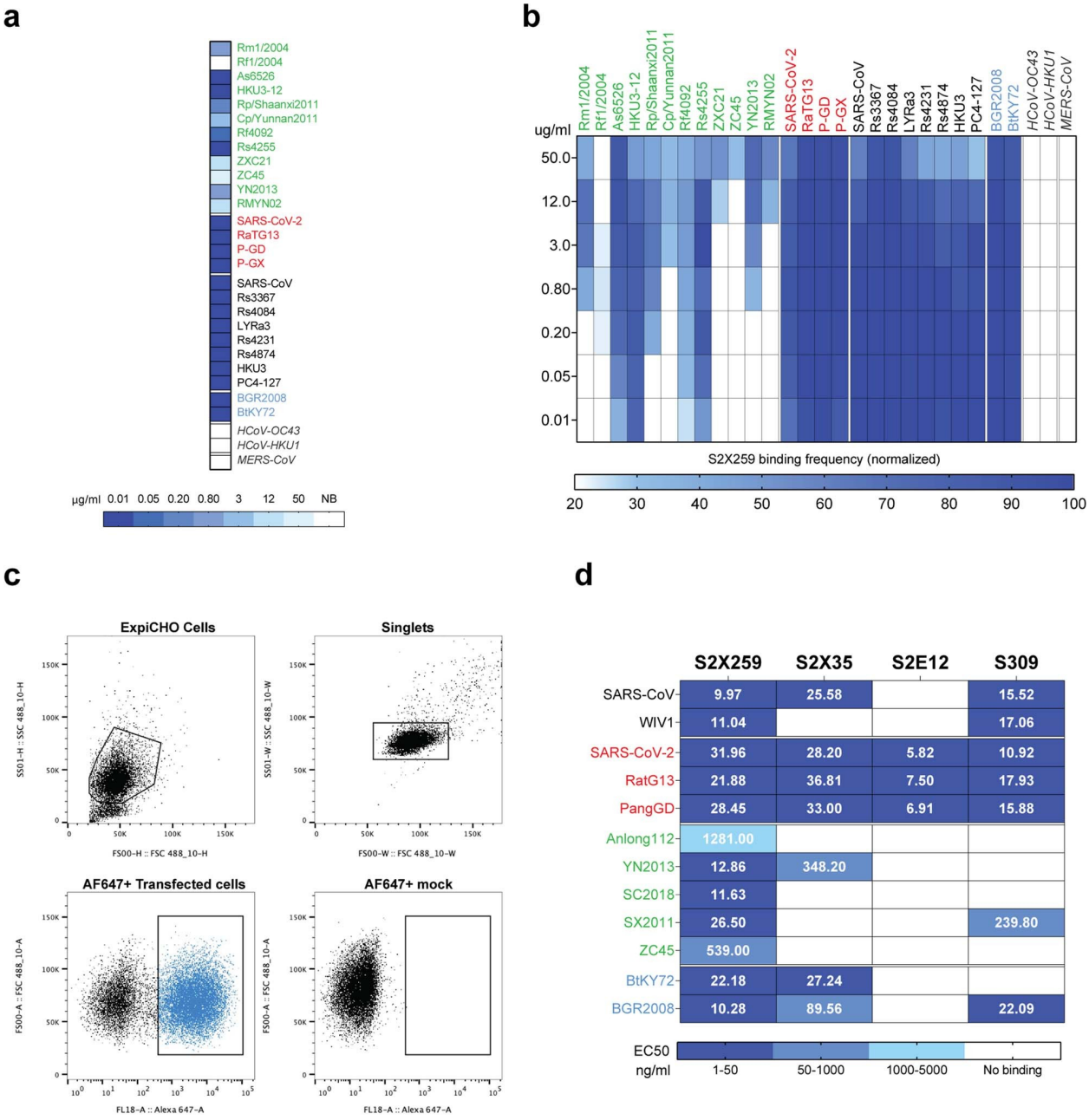


Figure 7.4: S2X259 protects hamsters challenged with prototypic and B.1.351 SARS-CoV-2.

a, **b**, Viral RNA loads (**a**) and replicating virus titres (50% tissue culture infectious dose ($TCID_{50}$)) (**b**) in the lungs of Syrian hamsters at 4 days after intranasal infection with prototypic SARS-CoV-2 (Wuhan-Hu-1-related) following prophylactic administration of S2X259 mAb at 1 (*n* = 6 hamsters), 5 (*n* = 5 hamsters) and 25 (*n* = 5 hamsters) mg kg^{-1}. **c**, **d**, Viral RNA loads (**c**) and replicating virus titres ($TCID_{50}$) (**d**) in the lungs of Syrian hamsters after therapeutic administration of S2X259 at 5 (*n* = 5 hamsters), 10 (*n* = 6 hamsters) and 20 (*n* = 6 hamsters) mg kg^{-1} 24 h after infection with SARS-CoV-2 (Wuhan-Hu-1-related). **e**, **f**, Quantification of viral RNA loads (**e**) and replicating virus titres ($TCID_{50}$) (**f**) in the lungs of Syrian hamsters 4 days after intranasal challenge with B.1.351 SARS-CoV-2 variant of concern following prophylactic administration of

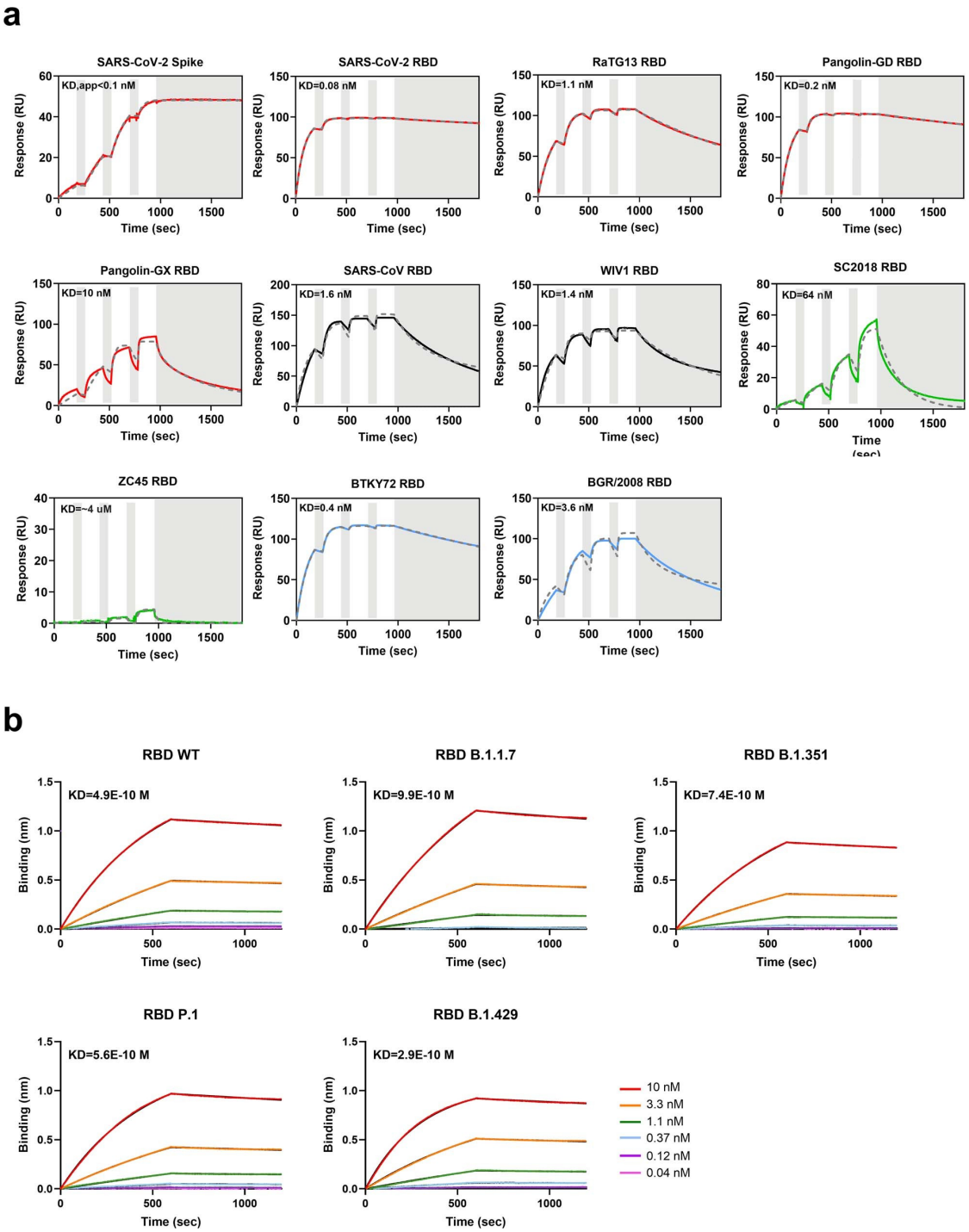
S2X259 at 1 mg kg^{-1} (n = 6 hamsters), 4 mg kg^{-1} (n = 6 hamsters) and in combination with S309 (1 + 1 mg kg^{-1}, n = 6 hamsters). Data from one independent experiment are presented. Two-tailed Mann–Whitney test was used for statistical analysis of significance.



Extended Data Figure 7-1: Breadth of S2X259 mAb across the Sarbecovirus subgenus.

a, Flow cytometry analysis of S2X259 cross-reactivity with a panel of 26 S glycoproteins representative of sarbecovirus clades 1a, 1b, 2 and 3. The colours represent the lowest concentration of mAb at which binding was observed. b, FACS binding of S2X259 to a panel of sarbecovirus S glycoproteins transiently expressed in ExpiCHO cells. Results

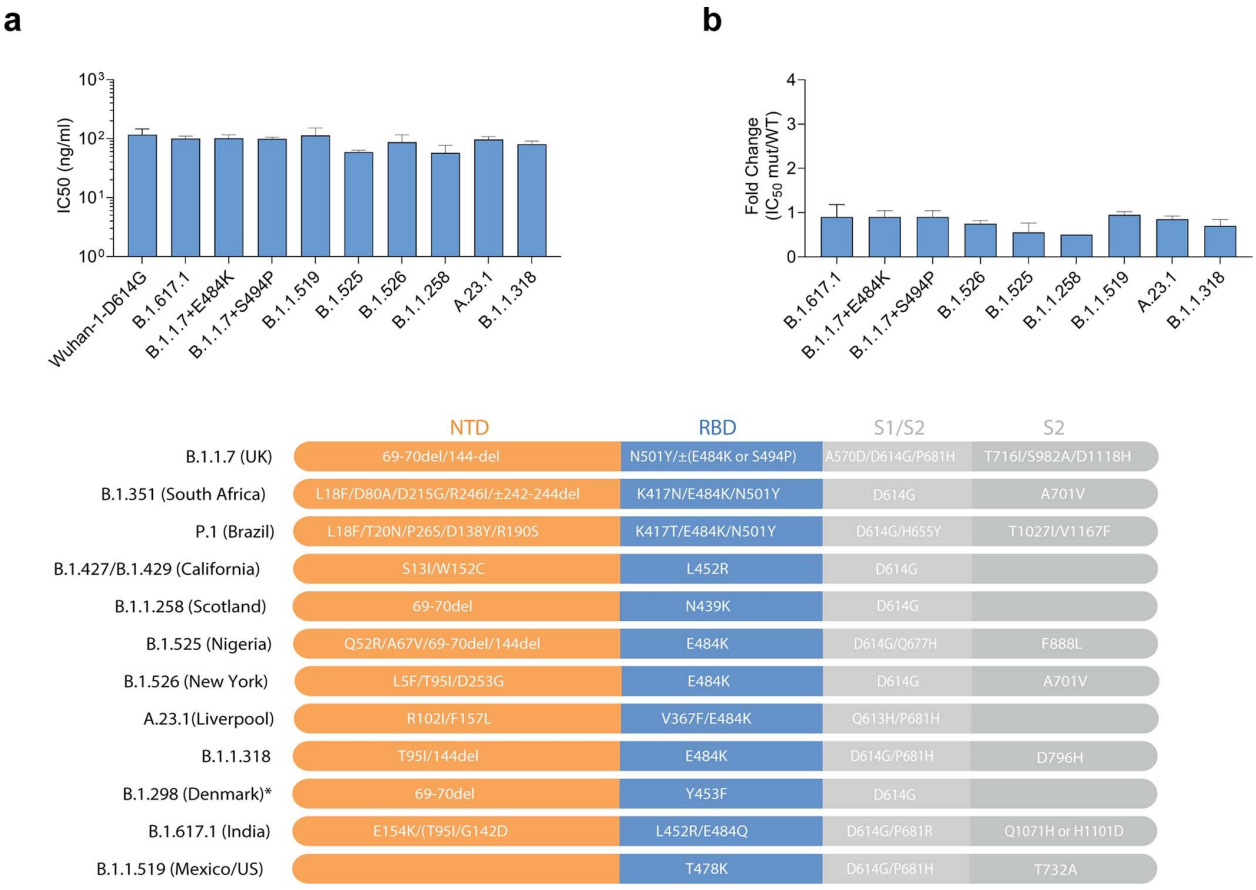
represent the binding frequency normalized using the ratio between mean fluorescence intensities for S2X259 and an expression control mAb (Pinto, 2021). c, FACS gating strategy used to assess binding of S2X259 or of the expression control mAb. d, ELISA binding of site-I-targeting S2E12(Tororici, 2020), site-II-targeting S2X259 and S2X3517, and site-IV-targeting S309(Pinto, 2020) mAbs to RBDs spanning the different clades of the *Sarbecovirus* subgenus. EC_{50}s are reported. *n* = 2 independent experiments.



Extended Data Figure 7-2: S2X259 Fab binding to recombinant sarbecovirus RBDs, prefusion SARS-CoV-2 S ectodomain trimer and RBD variants.

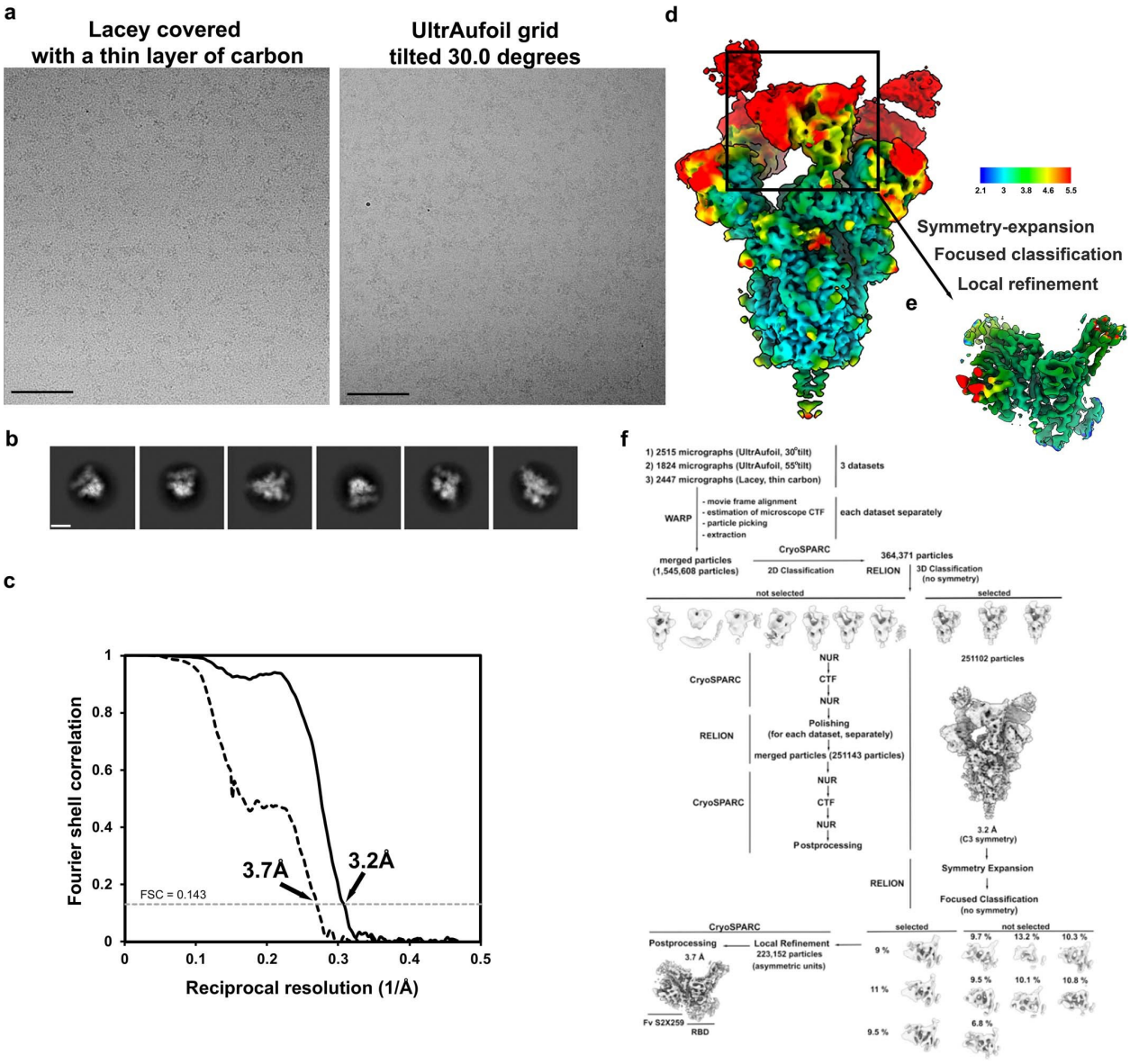
a, S or RBD antigens were captured on the sensor chip surface and binding to S2X259 Fab at 11, 33, 100 and 300 nM was monitored successively, in single-cycle kinetics format, by surface plasmon resonance. All data have been fit to a 1:1 binding model (black dashed line) and the equilibrium dissociation constant (K_D) is reported. For the S-

binding data, we report an apparent K_D ($K_{D,app}$) as kinetics are affected by conformational dynamics between open and closed RBD states. The colouring scheme matches the phylogenetic tree in Fig. 7.1a-b, Biolayer interferometry binding analysis of the S2X259 Fab to wild-type or variant-of-concern SARS-CoV-2 biotinylated RBDs immobilized at the surface of SA biosensors. The data are coloured according to the key and fits to a 1:1 binding model are shown as black lines. Equilibrium dissociation constants (K_D) are reported above each plot.



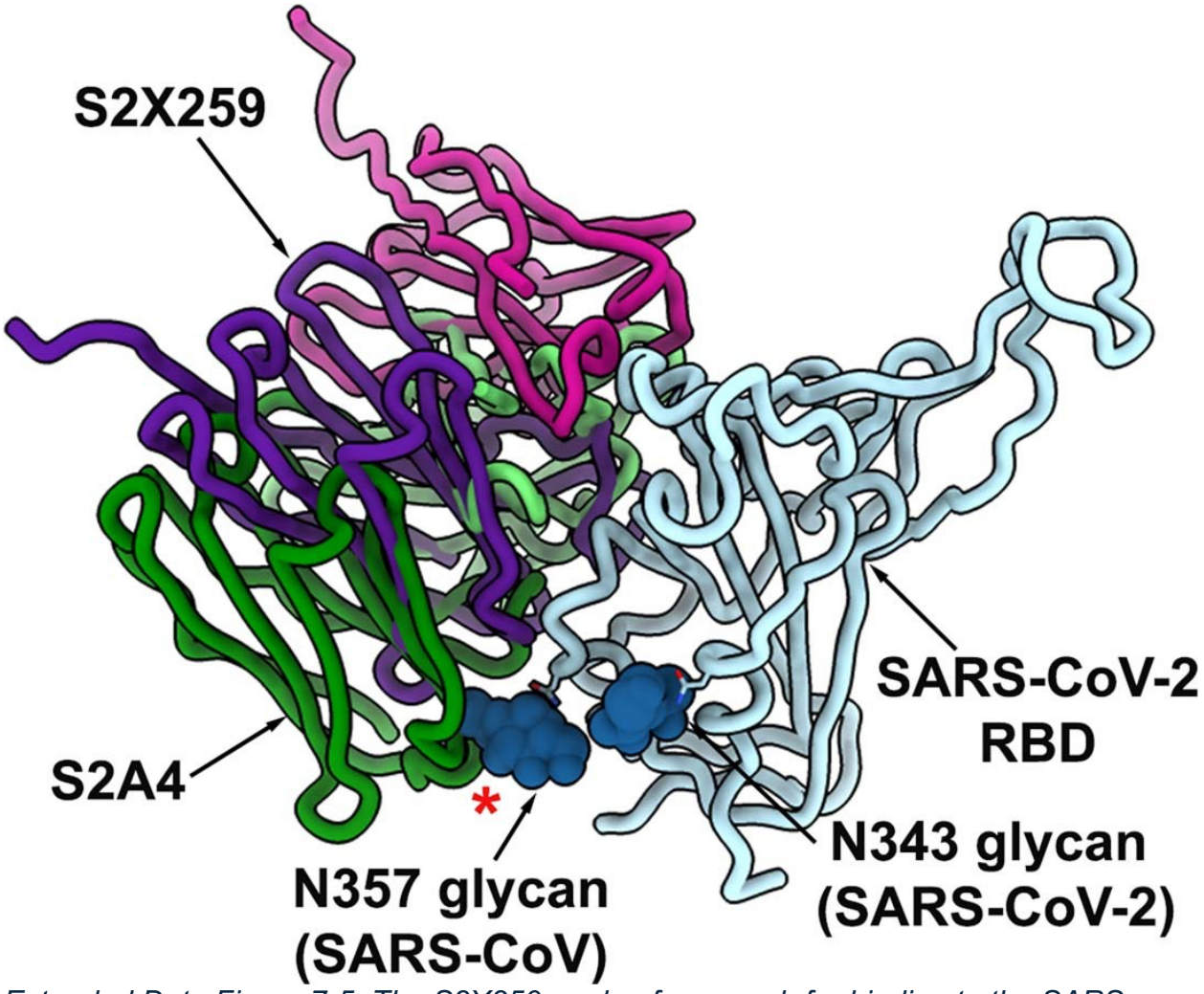
Extended Data Figure 7-3: S2X259 neutralizing activity is not affected by mutations present in circulating SARS-CoV-2 variant of concerns and recent lineages.

a, b, Neutralization by S2X259 of SARS-CoV-2 S VSV pseudotyped circulating variants shown as IC_{50} (a) and IC_{50} fold change as compared to Wuhan-Hu-1 virus with D614G substitution (b). Schematic of the SARS-CoV-2 S and the mutation landscape in each variant is shown beneath the panels. Del, deletion. Asterisks denotes virus that was tested as single RBD mutant in Fig. 7.1e.



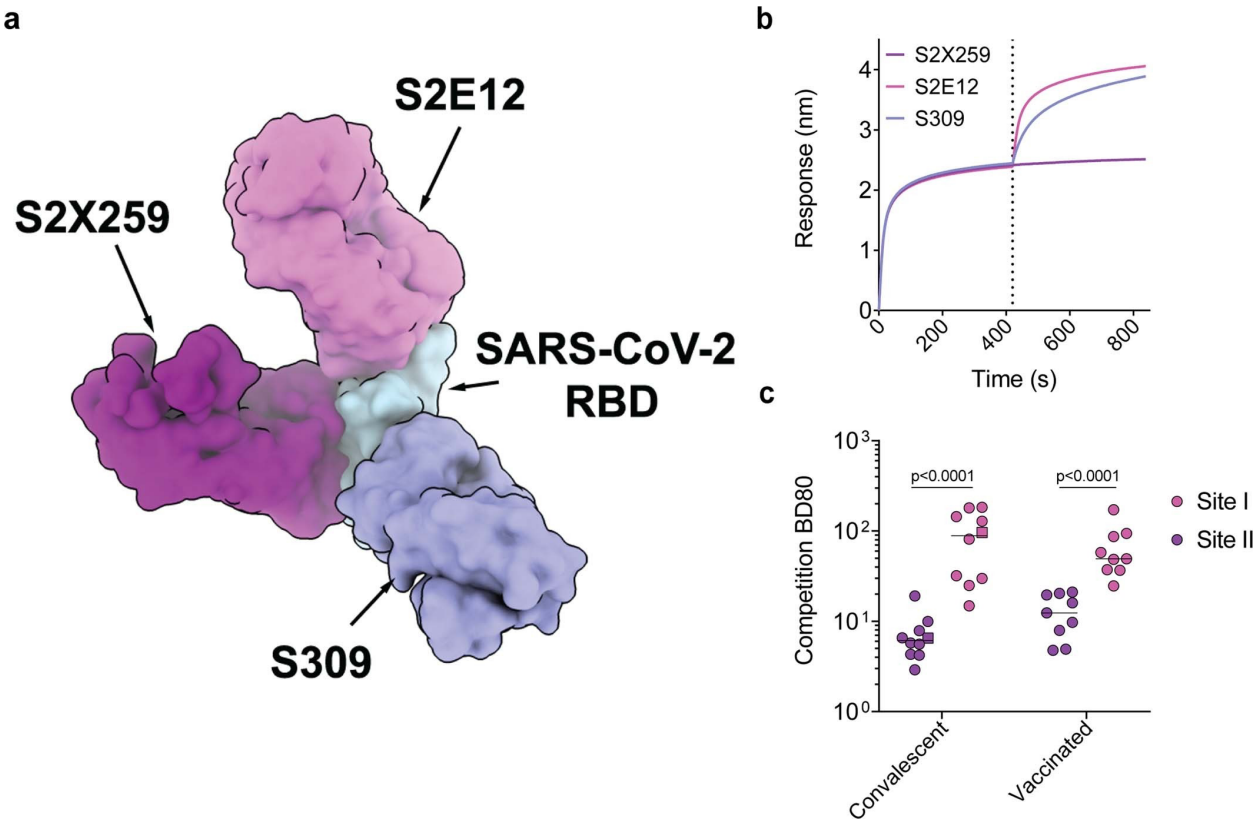
Extended Data Figure 7-4: Cryo-EM data processing and validation of S2X259-bound SARS-CoV-2 S.

a, **b**, Representative electron micrographs (**a**) and class averages (**b**) of SARS-CoV-2 S in complex with the S2X259 Fab. Scale bars, 500 Å (**a**), 100 Å (**b**). **c**, Gold-standard Fourier shell correlation curves for the S trimer bound to three S2X259 Fabs (solid black line) and the locally refined reconstruction of the RBD–S2X259 variable domains (dashed black line). The 0.143 cut-off is indicated by a horizontal dashed grey line. **d**, **e**, Local resolution map for the open S trimer bound to three S2X259 Fabs (**d**) and the locally refined reconstruction of the RBD–S2X259 variable domains (**e**). **f**, Cryo-EM data processing flowchart.



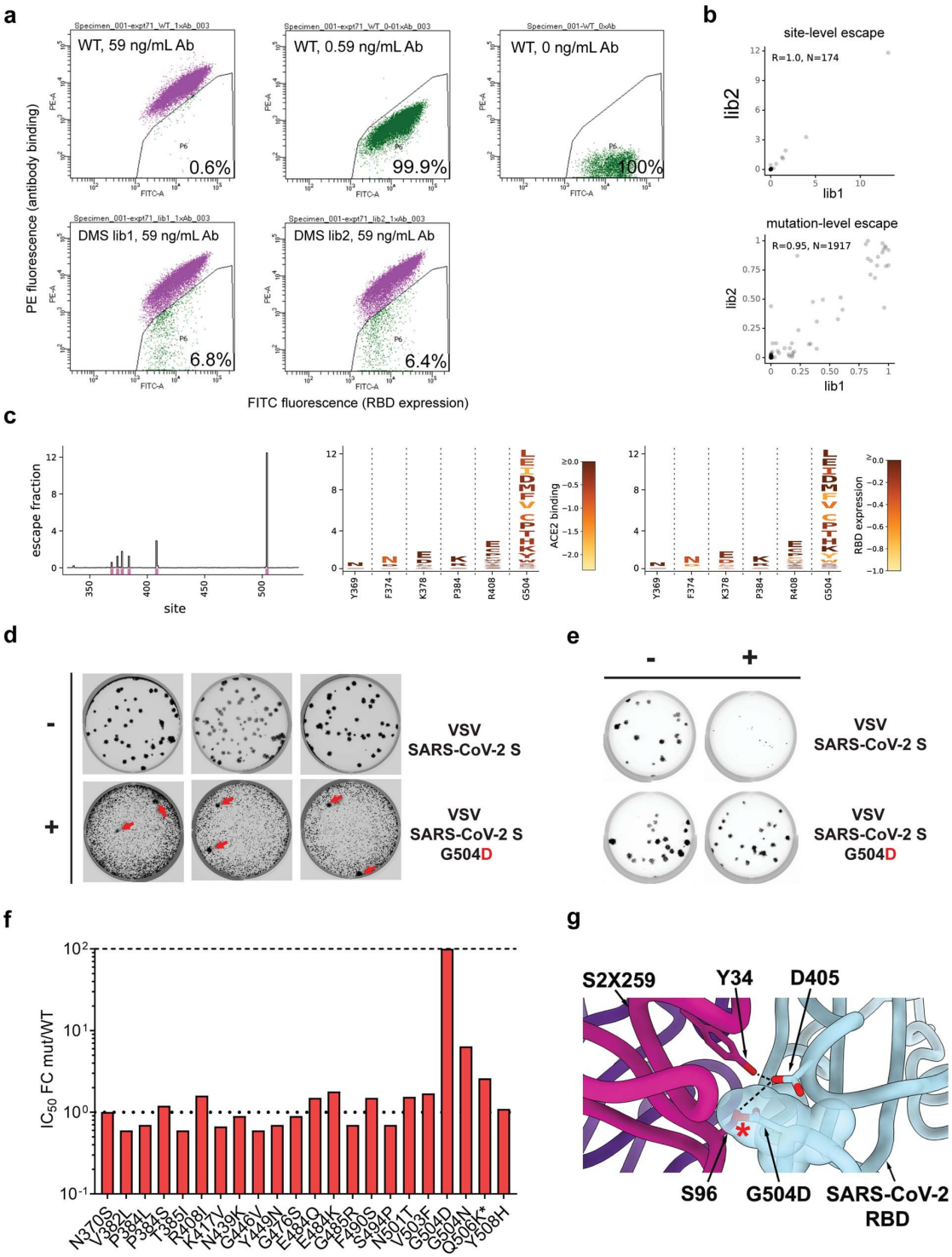
Extended Data Figure 7-5: The S2X259 angle of approach for binding to the SARS-CoV-2 RBD allows S2X259 to circumvent the SARS-CoV N357 glycan present in all sarbecovirus RBDs except SARS-CoV-2.

Ribbon diagram showing a superimposition of the S2X259-bound and S2A4-bound (PDB code 7JVA) SARS-CoV-2 RBD(Picolli, 2020). The SARS-CoV glycan at position N357 was modelled on the basis of the S230-bound SARS-CoV S structure (PDB code 6NB6(Walls,2019)) and is predicted to sterically hinder S2A4 binding (red star) but not S2X259 binding. The mAb light and heavy chains are coloured magenta and purple (S2X259) or light and dark green (S2A4), respectively. N-linked glycans are rendered as blue spheres.



Extended Data Figure 7-6: S2X259 can bind to the SARS-CoV-2 RBD in presence of site-I and -IV-targeting mAbs.

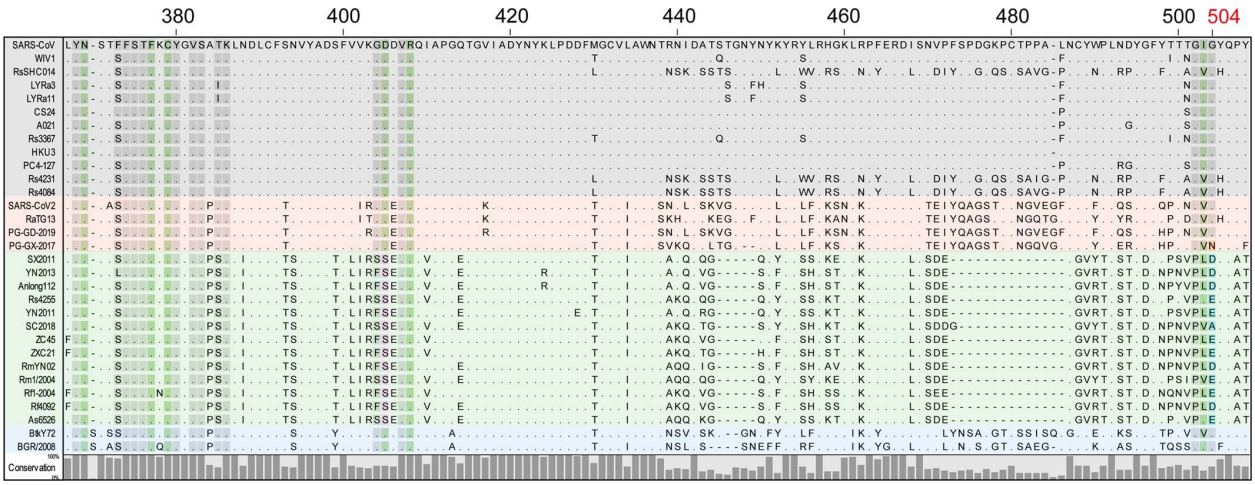
a, View of site-I-targeting S2E12(Tortorici, 2020) (pink), site-II-targeting S2X259 (magenta) and site-IV-targeting S309(Pinto, 2020) (purple) mAb bound to the SARS-CoV-2 RBD (light blue). **b**, Competition binding assays for S2X259 versus site-I-targeting S2E12 and site-IV-targeting S309 mAbs on SARS-CoV-2 RBD as measured by biolayer interferometry. One independent experiment out of two is shown. **c**, Competition ELISA (blockade-of-binding) between site-I-targeting S2H14(Piccolli, 2020) or site-II-targeting S2X259 and sera or plasma from individuals who had recovered from COVID-19 (*n* = 10 biological samples) and vaccinated individuals (*n* = 9 biological samples). Each plot shows the magnitude of inhibition of binding to immobilized RBD in the presence of each mAb, expressed as reciprocal sera or plasma dilution blocking 80% of the maximum binding response. The individual from whom S2X259 mAb was isolated is shown as a square. Statistical analysis was performed using two-tailed Mann–Whitney test.



Extended Data Figure 7-7: S2X259 has a high barrier for the emergence of resistance mutants.

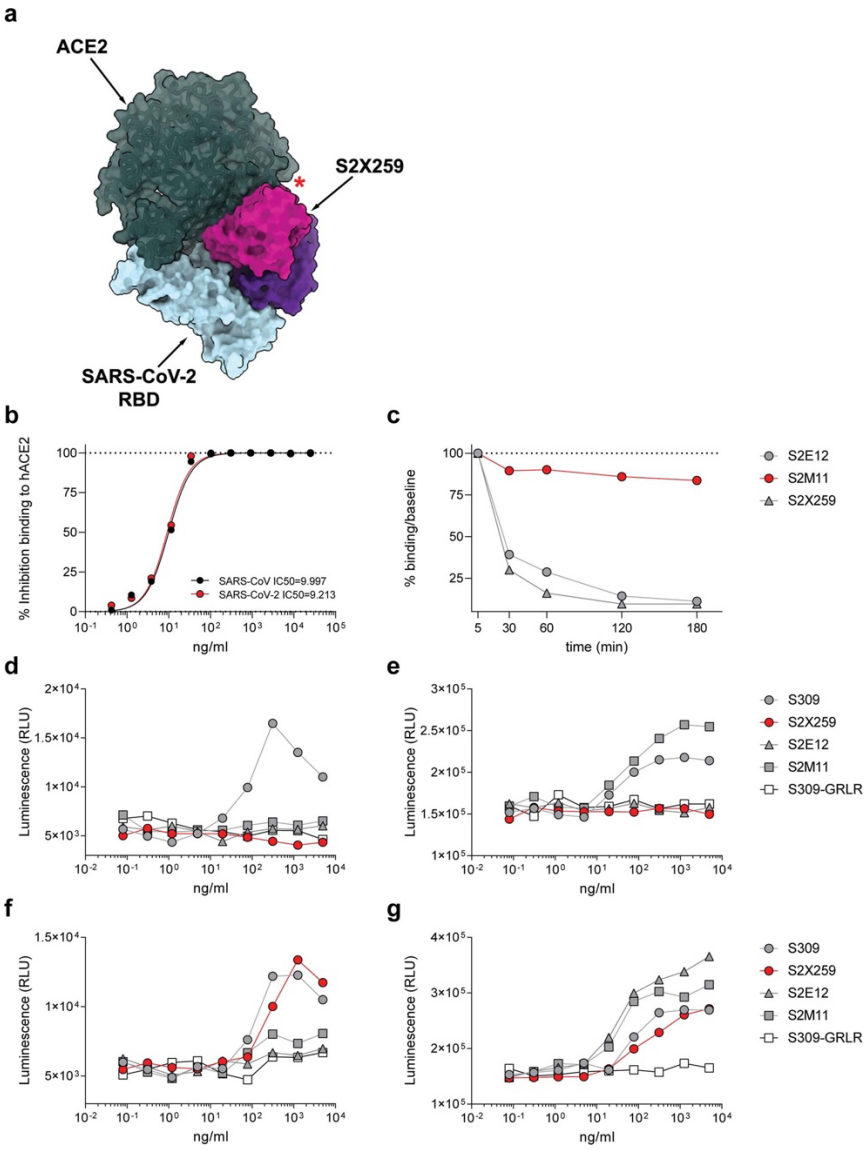
a, FACS gates used in deep-mutational scanning to select escape variants. Top, yeast controls expressing unmutated SARS-CoV-2 RBD labelled at relevant S2X259 concentrations for setting of selection gates. Bottom, fraction of cells in SARS-CoV-2 mutant libraries falling into the antibody-escape bin. **b**, Correlation in site-level (top, sum

of escape fractions for mutations at a site) and mutation-level (bottom) escape between independently generated and assayed RBD mutant libraries. **c**, Line plot of escape mutants along all positions in the SARS-CoV-2 RBD (left). Pink lines indicate sites that escape S2X259 binding illustrated at the mutation-level in logoplots (right). In logoplots, the height of a letter scales with its escape fraction. Letters are coloured according to their deleterious consequences for ACE2 binding (middle) or RBD expression (right) as determined in previous deep-mutational scanning experiments(Starr, 2020). **d**, Plaque assay using VSV-SARS-CoV-2 chimeric virus on Vero cells with no mAb (−) or S2X259 (+) in the overlay to isolate escape mutants (red arrow). Data are representative of three independent experiments. **e**, Plaque assays performed to validate the VSV-SARS-CoV-2 G504D mutant in Vero cells in the presence (+) or absence (−) of S2X259 in the overlay. Representative image of two independent experiments is shown. **f**, S2X259 in vitro neutralizing activity against SARS-CoV-2 S VSV pseudotyped mutants. For each mutant the fold change of the IC_{50}geometric mean versus SARS-CoV-2 S D614G is reported. *Q506K displayed a tenfold reduction in viral entry in comparison to the other mutants. Results from two independent experiments are reported. **g**, Zoomed-in view of the S2X259–RBD interface showing that the G504D substitution would disrupt mAb binding owing to steric hindrance (indicated with a red star).



Extended Data Figure 7-8: S2X259 epitope conservation across sarbecovirus clades.

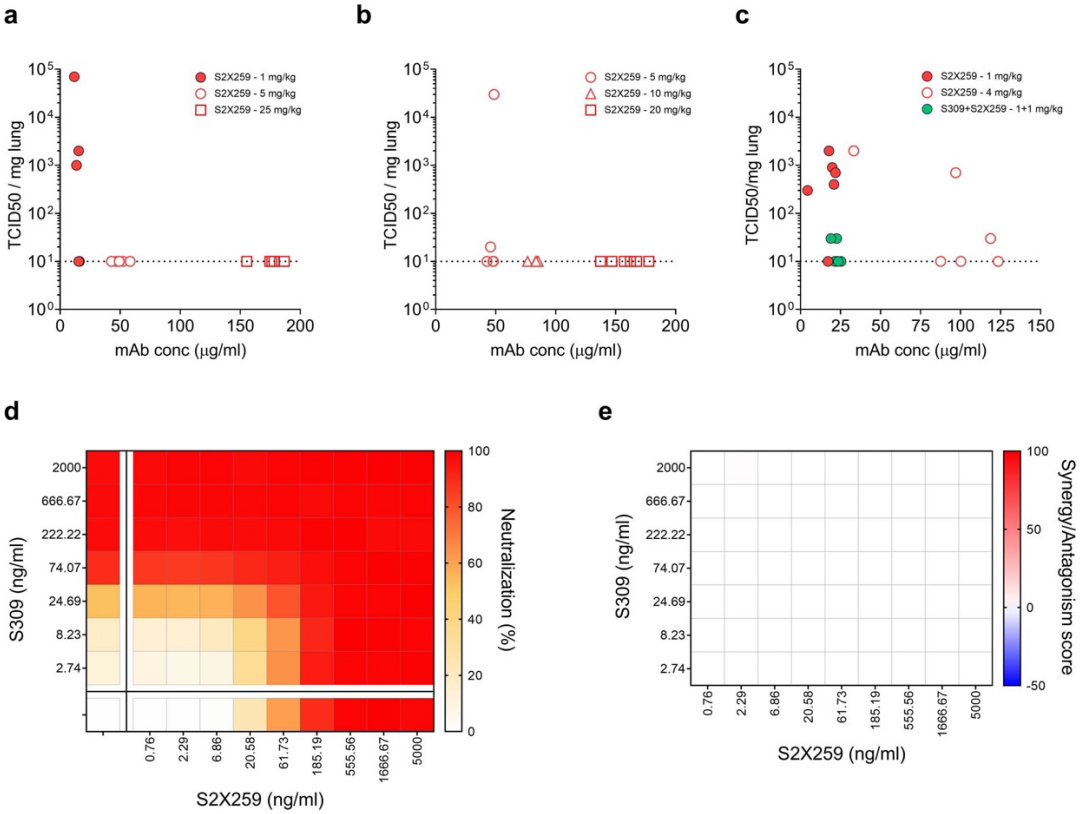
Protein sequence alignment of representative sarbecovirus RBDs with strictly conserved residues shown as dots. Overall conservation is represented as a bar plot at the bottom of the figure. Residue positions are based on SARS-CoV-2. Residues determined to be most important for S2X259 binding are denoted in dark green. Substitutions at positions D405 and G504 are indicated in pink and blue or orange, respectively. Additional residues representing the extended epitope are denoted in grey. Different clades within the *Sarbecovirus* subgenus are overlaid in grey (clade 1a), red (clade 1b), green (clade 2) and light blue (clade 3).



Extended Data Figure 7-9: Inhibition of ACE2 engagement, S1 subunit shedding and activation of FcγRIIa and FcγRIIIa in vitro.

a, S2X259 (purple and pink) and ACE2 (dark green) bind partially overlapping binding sites on the SARS-CoV-2 RBD (blue). **b**, Pre-incubation of serial dilutions of S2X259 with SARS-CoV-2 (red) or the SARS-CoV (black) RBDs prevents binding to immobilized human ACE2 (hACE2) ectodomain in ELISA. **c**, mAb-mediated S1 subunit shedding from cell-surface-expressed SARS-CoV-2 S as determined by flow cytometry. S2E12 was included as positive control and S2M11 was included as negative control. **d**, **e**, NFAT-driven luciferase signal induced in Jurkat cells stably expressing FcγRIIa H131 (**d**) variant or FcγRIIIa V158 (**e**) variant by S2X259 binding to full-length wild-type SARS-CoV-2 S on ExpiCHO target cells. **f**, **g**, NFAT-driven luciferase signal induced in Jurkat cells stably expressing FcγRIIa H131 (**f**) or FcγRIIIa V158 (**g**) variants by S2X259 binding to uncleavable full-length prefusion stabilized SARS-CoV-2 S (unable to release

the S1 subunit) transiently expressed in ExpiCHO cells. SE12, S2M11, S309 and S309-GRLR mAbs are included as controls.



Extended Data Figure 7-10: Correlation between mAb concentration and infectious virus in vivo and in vitro neutralizing activity of S2X259 + S309 antibody cocktail.

a–c, Infectious virus titres in the lungs at 4 days after infection plotted as a function of serum mAb concentrations before infection (day 0) with prototypic SARS-CoV-2 (a) and SARS-CoV-2 B.1.351 (c) in prophylactic setting or at 4 days after infection with prototypic SARS-CoV-2 (b) in therapeutic setting. d, e, SARS-CoV-2 B1.351 VSV-based pseudotype neutralization (d) and synergy score (e) measured combining S2X259 and S309 mAbs. Synergy was calculated using MacSynergyII (synergy score: 0.98, antagonism score: 0). Results represent four replicates of one independent experiment.

Table S 7.1: Characteristics of S2X259. VH and VL % identity refers to V gene identity compared to germline (IMGT).

Donor Gender	Male
Donor Age	52
Donor Status	Symptomatic
Days after symptoms onset	75
VH gene usage	VH1-69
VH % identity	94.1
HCDR3	ARGFNGNYYGWGDDDAFDI
VL gene usage	VL1-40
VL % identity	98.26
LCDR3	QSYDSSLSGPNWV

Table S 7.2: S2X259-mediated neutralization (IC50) of VSV pseudotypes harbouring SARS-CoV-2 S representing VOC, single mutants and different sarbecovirus clades.

Clade	Strain/VOC/RBD mutant	IC_{50} (ng/ml)
1a	SARS-CoV	90.8
	RsSHC014	7.3
	WIV1	581.1
	WIV16	196.5
1b	SARS-CoV-2	213.2
	B.1.1.7	205.0
	B.1.351	357.9
	B.1.429	114.2
	P.1	458.3
	N439K	281.3
	Y453F	221.5
	RaTG13	38.3
	P-GX	4857.0
	P-GD	106.1

Table S 7.3: CryoEM data collection and refinement statistics.

	SARS-CoV-2 S/S2X259 PDB 7RA8 EMD-24347	SARS-CoV-2 S/S2X259 (local refinement) PDB 7RAL EMD-24365
Data collection and processing		
Magnification	130,000	130,000
Voltage	300	300
Electron exposure ($e^-/Å^2$)	70	70
Defocus range (μm)	0.8-2.0	0.8-2.0
Pixel size (Å)	0.525	0.525
Symmetry imposed	C3	C1
Final particle images (no.)	251,102	223,152
Map resolution (Å)	3.2	3.6
FSC threshold	0.143	0.143
Validation		
MolProbity score	0.82	0.94
Clash score	0.56	0.56
Poor rotamers (%)	0.75	0.39
Ramachandran plot		
Favored (%)	97.29	96.25
Allowed (%)	2.62	3.22
Disallowed (%)	0.09	0.54
EMRinger score	3.45	3.85

Table S 7.4: X-ray data collection and refinement statistics.

	RBD/S2X259/S2H97 PDB ID: 7M7W
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	86.19, 66.40, 237.66
α, β, γ (°)	90.00, 94.34, 90.00
Resolution (Å)	63.94-2.65 (2.70-2.65)
R_{merge}	0.149 (2.494)
$I / \sigma I$	10.9 (0.8)
Completeness (%)	98.6 (98.3)
Redundancy	6.9 (6.8)
Refinement	
Resolution (Å)	2.65
No. reflections	73,189
R_{work} / R_{free}	0.221/0.271
No. atoms	
Protein	16,162
Ligand/ion	28
Water	95
B-factors	
Protein	75.86
Ligand	84.00
Water	50.09
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.817

Values for the highest-resolution shell are shown in parentheses.

Table S 7.5: Summary of nucleotide and amino acid mutations found in 18 neutralization-resistant VSV-SARS-CoV-2-S chimera plaques.

	S2X259		S2X35	
Plaque	**Nt mutation**	**Aa Mutant**	**Nt mutation**	**Aa Mutant**
1	G1511A	G504D	A1106C	Y369S
2	G1511A	G504D	G1511A	G504D
3	G1511A	G504D	T1508G	V503G
4	G1511A	G504D	G1213A	D405N
5	G1511A	G504D	G1511A	G504D
6	G1511A	G504D	A1132G	K378E
7	G1511A	G504D	G1511A	G504D
8	G1511A	G504D	G1213A	D405N
9	G1511A	G504D	T1508G	V503G
10	G1511A	G504D	T1508G	V503G
11	G1511A	G504D	G1511A	G504D
12	G1511A	G504D	G1511A	G504D
13	G1511A	G504D	G1511A	G504D
14	G1511A	G504D	T1508A	V503D
15	G1511A	G504D	G1213A	D405N
16	G1511A	G504D	T1508G	V503G
17	G1511A	G504D	G1511A	G504D
18	G1511A	G504D	G1511A	G504D
19			G1511A	G504D
20			G1510A	G504S
21			G1511A	G504D
22			G1511A	G504D
23			G1510A	G504S
24			A1132G	K378E
25			G1213A	D405N
26			G1510A	G504S
27			G1511A	G504D
28			G1511A	G504D
29			G1511A	G504D
30			A1132G	K378E
31			T1508G	V503G

CHAPTER 8. ANTIBODY-MEDIATED BROAD SARBECOVIRUS NEUTRALIZATION THROUGH ACE2 MOLECULAR MIMICRY.

In this story, I describe the isolation and characterization of a human monoclonal antibody, designated S2K146, that broadly neutralizes viruses belonging to SARS-CoV– and SARS-CoV-2–related sarbecovirus clades, which use angiotensin-converting enzyme 2 (ACE2) as an entry receptor. I detail our structural and functional studies showing that most of the virus residues that directly bind S2K146 are also involved in binding to ACE2. This allows the antibody to potently inhibit receptor attachment. S2K146 protects against SARS-CoV-2 Beta variant challenge in hamsters, and viral passaging experiments reveal a high barrier for emergence of escape mutants, making it a good candidate for clinical development. The conserved ACE2-binding residues present a site of vulnerability that might be leveraged for developing vaccines eliciting broad sarbecovirus immunity.

Adapted from: Park, Y.-J., De Marco, A., Starr, T.N., Liu, Z., Pinto, D., Walls, A.C., Zatta, F., **Zepeda, S.K.**, Bowen, J., Sprouse, K., Joshi, A., Giurdanella, M., Guarino, B., Noack, J., Abdelnabi, R. Foo, C. -L., Rosen, L. Lempp, F., Benigni, F., Snell,G., Neyts, J., Whelan, S.P., Virgin, H.W., Bloom, J.D., Corti, D., Pozzuto, M.S., Veesler, D. (2022) “Antibody-mediated broad sarbecovirus neutralization through ACE2 molecular mimicry." Science **0**(0): eabm8143.

8.1 Chapter Introduction

The zoonotic spillover of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in a global pandemic causing more than 266 million infections and more than 5.2 million fatalities as of December 2021. Continued SARS-CoV-2 evolution leads to the emergence of variants of concern (VOCs) that are characterized by higher transmissibility, immune evasion, and/or disease severity. For pandemic preparedness, we need pan-sarbecovirus countermeasures, such as vaccines and therapeutics that are effective against all SARS-CoV-2 variants and divergent zoonotic sarbecoviruses (Sanchez, 2021).

The coronavirus spike (S) glycoprotein promotes viral entry into host cells and is the main target of neutralizing antibodies elicited by infection or vaccination (Walls, 2016; Walls, 2017; Walls,2020; Wrapp, 2020, Totorici, 2019; Stamatatos, 2021). The S protein comprises an S_1 subunit, which recognizes host cell receptors, and an S_2 subunit that drives viral cell membrane fusion. The S_1 subunit includes the N-terminal domain and the receptor binding domain (RBD) and two additional domains designated C and D (*6*). For SARS-CoV and SARS-CoV-2, the RBD interacts with angiotensin-converting enzyme 2 (ACE2) to allow virus entry into host cells (Walls, 2020; Zhou, 2020; Letko, 2020; Hoffmann, 2020, Lim 2003; Li, 2005; Wang, 2020; Yan, 2020; Shang, 2020; Lan, 2020). The RBD is also the main target of serum neutralizing activity elicited by infection (Piccoli) and vaccination (Stamatatos,2021; Greaney, 2021) and exposes multiple antigenic sites that are recognized by broadly neutralizing sarbecovirus antibodies (Abs) (Totorici, 2021; Starr, 2021; Pinto, 2020; Jette, 2021; Martinez, 2021; Wec, 2020; Rappazzo, 2021) (Fig. S1). However, a large fraction of

Abs in polyclonal sera (Piccoli, 2020) and most monoclonal Abs (mAbs) selected for therapeutic development (Corti, 2021) target a subset of epitopes that overlap the ACE2-contact surface [designated the receptor binding motif (RBM)]. The marked genetic divergence and plasticity of the RBM among SARS-CoV-2 variants and sarbecoviruses have thus far limited the breadth of Abs recognizing this region, and they are readily escaped by mutations (Starr, 2021; McCallum, 2021; McCallum, 2021; Mlcochova, 2021; Starr, 2020, 2021).

8.2 Identification of broadly neutralizing sarbecovirus Antibody

To identify broadly neutralizing sarbecovirus Abs, we isolated SARS-CoV-2 S-specific [immunoglobulin G (IgG)] memory B cells from one symptomatic COVID-19 convalescent individual (who was not hospitalized) 35 days after symptoms onset. We identified one mAb, designated S2K146 [immunoglobulin heavy and light variable genes 3-43 and 1-44, respectively (IGHV3-43; IGL1-44)], which did not compete with S309 (site IV) (Pinto, 2020) or S2X259 (site II) (Tortorici, 2021) but competed with S2E12 (site I), a potent RBM mAb with neutralization breadth against SARS-CoV-2–related sarbecoviruses (Tortorici, 2020) (**Fig. 8.1A and Fig. S8.1**). Like S2E12, S2K146 bound to all SARS-CoV-2 VOC RBDs as well as all clade 1b sarbecovirus RBDs tested by enzyme-linked immunosorbent assay (ELISA) (**Fig. 8.1B and Fig. S8.2A**). In contrast to S2E12 and other site I–targeting Abs described so far, however, S2K146 also cross-reacted with the SARS-CoV and WIV-1 RBDs (clade 1a), which share 73 and 76% sequence identity with the SARS-CoV-2 RBD, respectively (**Fig. 8.1B and Fig. S8.2A**). S2K146 did not bind to clades 2 and 3 sarbecovirus RBDs, similarly to the broadly neutralizing sarbecovirus S309 mAb but in contrast to the S2X259 and S2H97 (site V)

mAbs (Tortorici, 2021; Starr, 2021). We observed S2K146 cross-reactivity with clades 1a and 1b sarbecoviruses using native S trimers transiently expressed on the surface of mammalian cells (**Fig. 8.1C**) and yeast surface–displayed RBDs (**Fig. 8.1D**), consistent with the ELISA results. On the basis of these results, we hypothesized that S2K146 recognizes a previously uncharacterized RBM epitope that is conserved among sarbecovirus clades 1a and 1b.

8.3 Neutralization Potency of S2K146 mAb

To evaluate the neutralization potency of the S2K146 mAb, we carried out dose-response inhibition assays using a vesicular stomatitis virus (VSV) pseudotyping platform. S2K146 efficiently blocked SARS-CoV and SARS-CoV-2 S-mediated entry into cells with median inhibitory concentration (IC_{50}) of 108 and 16 ng/ml, respectively (**Fig. S8.2B**). Moreover, S2K146 potently neutralized VSV pseudotypes harboring SARS-CoV-2 S glycoproteins from VOCs including Alpha, Beta, Gamma, Delta plus (AY.1/AY.2), Epsilon, and Lambda (**Fig. S8.2B**). S2K146 also weakly neutralized VSV pseudotyped with BtKY72 S (clade 3) harboring the K493Y/T498W mutations (SARS-CoV-2 numbering) (**Fig. S8.2C**), which enable human ACE2–mediated entry (Starr, 2022), whereas S2E12 did not recognize the wild-type or double-mutant BtKY72 RBD (**Fig. S8.3**). Finally, S2K146 neutralized authentic SARS-CoV-2 (isolate USA-WA1/2020, lineage A, IC_{50} = 10 ng/ml) and SARS-CoV-2 VOCs (Alpha, IC_{50} = 9 ng/ml; Beta, IC_{50} = 9 ng/ml; Delta, IC_{50} = 8 ng/ml; Kappa, IC_{50} = 30 ng/ml) with a potency approaching that observed with the ultrapotent S2E12 mAb (Tortorici, 2020) (Wuhan-1, IC_{50} = 3.5 ng/ml; Alpha, IC_{50} = 2.5 ng/ml; Beta, IC_{50} = 2 ng/ml; Delta, IC_{50} = 1.5 ng/ml; Kappa, IC_{50} = 4.5 ng/ml) in a side-by-side comparison (**Fig. 8.1E**).

To assess the role of somatic mutations for S2K146 binding and neutralization, we generated its inferred unmutated common ancestor (S2K146 UCA). Alignment with the UCA amino acid sequence reveals that S2K146 harbors seven and two somatic hypermutations in the heavy- and light-chain complementarity determining regions (CDR), respectively [V_H (variable region of immunoglobulin heavy chain) identity: 94.4%; V_L(variable region of immunoglobulin light chain) identity: 98.9%] (**Fig. S8.2D**). Except for WIV-1, S2K146 and S2K146 UCA showed no major cross-reactivity differences with a panel of RBDs representative of clade 1 sarbecoviruses, as determined by ELISA (**Fig. S8.2E**). Nevertheless, biolayer interferometry revealed that S2K146 bound to prefusion-stabilized SARS-CoV and SARS-CoV-2 S ectodomain trimers with enhanced avidities compared with S2K146 UCA (**Fig. S8.2F**). Accordingly, S2K146 UCA showed a marked loss in neutralizing activity against both SARS-CoV S and SARS-CoV-2 S VSV pseudotypes (**Fig. 8.1F**). Our results suggest that somatic hypermutations associated with S2K146 affinity maturation are especially important for enhancing mAb avidity and potency.

8.4 Structural Basis of S2K146 cross reactivity

To understand the sarbecovirus cross-reactivity of the RBM-specific S2K146 mAb, we determined a cryo–electron microscopy (cryo-EM) structure of the S2K146 Fab fragment in complex with the SARS-CoV-2 S ectodomain trimer at 3.2-Å resolution (**Fig. 8.2A, Fig. S8.4, and table S8.1**). Three-dimensional (3D) classification of the data led to the determination of a structure with three open RBDs, each bound to a S2K146 Fab, as well as a structure with two open RBDs and one closed RBD, with a Fab bound to each of them (**Fig. S8.4**). Our cryo-EM data show that the opening of two RBDs is

enough to allow three Fabs to bind to an S trimer, as the remaining closed RBD can engage an S2K146 Fab owing to its angle of approach.

To overcome the conformational heterogeneity of the S2K146-bound RBDs relative to the rest of the S trimer, we used focused 3D classification and local refinement of the S2K146 variable domains and RBD to obtain a reconstruction at 3.2-Å resolution enabling unambiguous model building and providing a detailed view of the binding interface (**Fig. 8.2B, Fig. S8.4, and table S8.1**). S2K146 recognizes an epitope in antigenic site I (Piccoli, 2020), which overlaps with the RBM and is partially masked when the three RBDs adopt a closed state leading to clashes between the mAb and a neighboring RBD (**Fig. 8.2, A and B, and Fig. S8.1**). The S2K146 paratope includes the heavy chain N terminus and CDRH1, CDRH2, and CDRH3, accounting for three-quarters of the surface buried upon binding, with light chain CDRL1, CDRL2, and CDRL3 making up the rest of the interface. A total of 1000 $Å^2$ of the paratope surface is buried at the interface with the RBM through electrostatic interactions and shape complementarity.

The S2K146 footprint on the SARS-CoV-2 RBD highly resembles that of the ACE2 receptor, with 18 of 24 epitope residues shared with the ACE2 binding site, including key ACE2 contact positions L455, F486, Q493, Q498, and N501 (**Fig. 8.2, C and D**). Moreover, electrostatic interactions formed between S2K146 and the SARS-CoV-2 RBD recapitulate some of the contacts involved in ACE2 binding, such as with residues K417, Y449, Y489, Q493, and G502 (**Fig. 8.2, E and F**). Although some S2K146 contact residues are mutated in several variants, such as K417 (Beta and Gamma), L452 (Delta, Epsilon, and Kappa), E484 (Beta, Gamma, and Kappa), and N501 (Alpha,

Beta, and Gamma), the retention of neutralization of these variants suggests that the binding interface is resilient to these residue substitutions (**Fig. 8.1E and Fig. S8.2B**). The cross-reactivity with and broad neutralization of SARS-CoV by S2K146 may be partially explained by the strict conservation or conservative substitution of nine and four epitope residues relative to SARS-CoV-2, respectively (**Fig. 8.2, G and H, and Fig. S8.5A**), consistent with the ability of both RBDs to bind human ACE2.

8.5 S2K146 overcomes mutational plasticity of RBM

S2K146 therefore overcomes the mutational plasticity of the RBM, which is implicated in immune evasion, by targeting residues required for binding to the ACE2 receptor. This is supported by S2K146 recognition of the reconstructed RBD ancestor of SARS-CoV and SARS-CoV-2 (**Fig. S8.5, A and B**), where human ACE2 binding first arose during sarbecovirus evolution (Starr, 2022), in line with the hypothesis that human ACE2 binding participates in conferring S2K146 susceptibility. As S2K146 does not compete with broadly neutralizing sarbecovirus mAbs targeting other antigenic sites, such as S309 (Pinto, 2020) and S2X259 (Tortorici, 2021) (**figs. S8.1 and S8.6**), they could be combined in a cocktail to enhance breadth further and set an even higher barrier for emergence of escape mutants.

To prospectively evaluate the impact of antigenic drift on S2K146 neutralization, we mapped RBD mutations that affect mAb binding using deep mutational scanning (DMS) of a yeast-displayed RBD mutant library covering all possible single residue substitutions in the Wuhan-Hu-1 RBD background (Starr, 2020). S2K146 binding was reduced by only a restricted number of amino acid substitutions compared to S2E12, which binds an overlapping but distinct epitope (**Fig. 8.3, A to D, and Fig. S8.7, A and**

B). All these mutations correspond to RBD residues buried upon ACE2 recognition (F456, A475, E484, F486, N487, and Y489) (**Fig. 8.3B**). Only one of these residue substitutions (Y489H) is accessible through a single-nucleotide change and could escape S2K146 recognition with a penalty on ACE2 binding affinity smaller than an order of magnitude, as determined by DMS data (Starr, 2020). None of the individual mutations present in the recently identified SARS-CoV-2 Omicron VOC affected S2K146 binding (**Fig. 8.3B**), although the effect of the full constellation of mutations remains to be evaluated. Conversely, DMS profiling of the S2K146 UCA revealed a greater number of binding-escape mutations, including some residue substitutions present in Omicron (e.g., Q493R or Q498R) (**Fig. 8.3B**). Therefore, the hotspot targeting of S2K146 on residues that are constrained in SARS-CoV-2 evolution appears to be a direct consequence of mAb affinity maturation.

To explore whether our escape map was consistent with in vitro viral evolution under mAb pressure, a replication competent VSV–SARS-CoV-2 S Wuhan-Hu-1/D614G chimera (Case, 2020) was passaged in cell culture in the presence of the S2K146 mAb. Consistent with the DMS data, Y489H was the sole mutation resulting from a single nucleotide substitution that was detected in all the 36 neutralization-resistant plaques sampled (**Fig. 8.3E, Fig. S8.8A, and table S8.2**). SARS-CoV-2 residue Y489 forms multiple interactions with S2K146 CDRH3 and accounts for ~10% of the total epitope buried surface area (**Fig. 8.3E**), in line with the major impact of the Y489H substitution on mAb neutralization. Of all the mutations at position 489 identified by DMS to reduce S2K146 binding (**Fig. 8.3B**), the Y-to-H substitution had the lowest

impact on ACE2 binding, which might explain why it was the sole neutralization escape mutant selected upon passaging.

To evaluate the fitness of the Y489H mutant, we carried out a competition assay in which replicating VSV chimeras harboring the Wuhan-Hu-1/D614G S with or without the Y489H substitution were mixed at equal titers and passaged together without mAb. Because of the fitness cost associated with the mutation, which dampens ~4.5-fold the 1:1 ACE2-binding affinity to the SARS-CoV-2 RBD (**Fig. S8.8B**), the Y489H S chimera was outcompeted by the Wuhan-Hu-1/D614G S chimera after only four rounds of passaging (**Fig. 8.3F**). Accordingly, only 29 out 2.9 million genomes were found to harbor the S Y489H mutation, underscoring the rarity of and the fitness cost imposed by this residue substitution (**Fig. 8.3G**). Collectively, these data illustrate the high barrier for emergence of escape mutants imposed by the S2K146 mAb, making it a good candidate for clinical development.

S2K146 targets antigenic site Ia, which overlaps with the RBM, indicating that mAb binding would compete with ACE2 attachment to the RBD via steric hindrance (**Fig. 8.4A**). Indeed, S2K146 inhibited binding of the SARS-CoV-2 and SARS-CoV RBDs to human ACE2 in a concentration-dependent manner, as measured by competition ELISA (**Fig. 8.4B**). As S2K146 conformationally selects for open RBDs, we assessed whether the mAb could promote shedding of the S_1 subunit from cell surface–expressed full-length SARS-CoV-2 S, similar to some other RBD-specific mAbs (Piccoli, 2020; Tortorici, 2021; Wec, 2020; Tortorici, 2020). S2K146 induced shedding of the S_1 subunit as efficiently as the RBM-targeting S2E12 mAb, whereas the control mAb S2M11, which locks S in the prefusion closed state, did not (Tortorici) (**Fig. 8.4C**). Furthermore,

S2K146 Fab triggered fusogenic rearrangement of a wild-type–like S ectodomain trimer, as previously described for several SARS-CoV and SARS-CoV-2 neutralizing mAbs (Starr, 2021; Walls, 2019; Lempp, 2021; Huo, 2020) (**Fig. 8.4D**). Thus S2K146-mediated sarbecovirus neutralization relies on competitively blocking viral attachment to the ACE2 receptor and putative inactivation of S trimers at the surface of virions before encountering host cells.

The efficient S2K146-induced S_1 shedding could explain the lack of FcγRIIa and FcγRIIIa activation, which we used as a proxy for Ab-dependent cellular phagocytosis and Ab-dependent cellular cytotoxicity, respectively (**Fig. S8.9, A and** B). However, S2K146 also did not activate FcγRIIa and triggered FcγRIIIa only weakly when target cells expressed an uncleavable prefusion-stabilized SARS-CoV-2 S protein (**Fig. S8.9, C and D**). The greater efficiency of S2E12 for activating FcγRIIIa, relative to S2K146, might be explained by the different angles of approach at which these two mAbs bind to the RBD (**Fig. S8.9, E and F**).

Next, we evaluated the therapeutic activity of S2K146 against challenge with the SARS-CoV-2 Beta VOC in a Syrian hamster model of infection (Abdelnabi, 2021; Boudewijns, 2020). S2K146 was administered in doses of 1, 5, and 10 mg/kg of body weight via intraperitoneal injection 24 hours after intranasal challenge with SARS-CoV-2, and the lungs of the animals were collected 3 days later for the quantification of viral RNA and replicating virus. In parallel, six animals were administered 1 mg/kg of the ultrapotent S2E12 mAb for benchmarking (Tortorici,2020). Viral RNA loads in the lungs were reduced by ~1, 4, and 3 orders of magnitude after receiving 1, 5, and 10 mg/kg of S2K146, respectively (**Fig. 8.4E**). Viral replication in the lungs was completely

abrogated for the 5 and 10 mg/kg groups and reduced by greater than 2.5 orders of magnitude for the 1 mg/kg group (**Fig. 8.4F**). Overall serum mAb concentrations measured at day 4 after infection inversely correlated with viral RNA loads and infectious virus in the lungs (Fig. S10, A and B). S2K146 therefore effectively protects against SARS-CoV-2 challenge in vivo in a stringent therapeutic setting.

8.6 Chapter Discussion

The SARS-CoV-2 RBD accounts for most serum neutralizing activity in both COVID-19 convalescent (Piccoli, 2020; Dejnirattisai, 2021) and vaccinated individuals (Statatomos, 2021; Greaney, 2021), and a subset of antigenic sites are targeted by broadly neutralizing sarbecovirus Abs (Tortorici, 2021; Starr, 2021; Pinto, 2020; Jette, 2021; Martinez, 2021; Wec, 2020; Rappazzo, 2021). RBD-based subunit vaccines and mRNA vaccines based on chimeric S glycoproteins elicit broadly neutralizing sarbecovirus Abs and heterotypic protection in vivo (Arunachalam, 2021; Walls, 2020; Walls, 2021; Saunders, 2021; Martinez, 2021). Most of the Abs with broad neutralizing activity are expected to target conserved RBD epitopes, owing to their much greater potency and protection efficacy compared to Abs that target the conserved fusion machinery (Pinto, 2021; Sauer, 2021; Song, 2021; Zhou, 2021; Wang, 2021; Hsieh, 2021). The discovery of a functionally constrained and conserved RBM epitope associated with broad sarbecovirus neutralization is consistent with the strong cross-reactivity with the SARS-CoV RBM observed with polyclonal Abs elicited by a clinical stage SARS-CoV-2 vaccine in nonhuman primates (Walls, 2021) and will guide the development of next-generation pan-sarbecovirus vaccines to protect from future zoonotic transmission events.

The broadly neutralizing sarbecovirus mAb S309 was isolated from a survivor of a 2003 SARS-CoV infection, and its derivative (sotrovimab) has received emergency use authorizations in several countries around the world for the early treatment of mild-to-moderate COVID-19 in adults and some pediatric patients who test positive for SARS-CoV-2 by direct viral testing and who are at high risk for progression to severe COVID-19, including hospitalization or death (Tortorici, 2021; McCallum, 2021; MIcochova, 2021; Cathcart). S309 has proven resilient to the emergence of SARS-CoV-2 variants in preclinical studies, possibly owing to targeting of a conserved RBD epitope with very limited mutational tolerance (Starr, 2021; Cathcart, 2021). The mechanism of S2K146-mediated ACE2 molecular mimicry also provides a high barrier for emergence of escape mutants in spite of the known mutational plasticity of the SARS-CoV-2 RBM (Starr, 2020). Therefore, the discovery of the S2K146 mAb might be a milestone for future treatment of COVID-19 patients and for pandemic preparedness against divergent sarbecoviruses.

8.7 Materials and Methods

Cell lines

Cell lines used in this study were obtained from ATCC (Vero-E6) or ThermoFisher Scientific (Expi CHO cells and Expi293F™ cells) or were generated via lentiviral transduction (Expi CHO-S, HEK293T-ACE2, Vero-TMPRSS2) (*37*).

Recombinant protein production

RBDs from different sarbecoviruses used in ELISA and BLI experiments were

expressed with N-terminal signal peptide and C-terminal thrombin cleavage siteTwinStrep-8xHis-tag in Expi293F cells at 37°C and 8% CO2. Cells were transfected using PEI MAX (Polysciences) at a DNA:PEI ratio of 1:3.75. Transfected cells were supplemented three days after transfection with 3 g/L glucose (Bioconcept) and 5 g/L soy hydrolysate (Sigma-Aldrich Chemie GmbH). Cell culture supernatant (423 mL) was collected seven days after transfection and supplemented with 47 mL 10x binding buffer (1 M Tris-HCl, 1.5 M NaCl, 20 mM EDTA, pH 8.0) and 25 mL BioLock (IBA GmbH) and incubated on ice for 30 min. Proteins were purified using a 5 mL Strep-Tactin XT Superflow high capacity cartridge (IBA GmbH) followed by buffer exchange to PBS using HiPrep 26/10 desalting columns (Cytiva).

SARS-CoV-2 S hexapro (*54*), used for cryo-EM single particle studies, was expressed and purified as described before (*33*). The SARS-CoV-2 S 'wildtype' ectodomain trimer used for refolding experiments followed by negative stain EM was engineered as follows and recombinantly expressed as previously described (*37*). The SARS-CoV-2 S D614G 'wildtype' has a mu-phosphatase signal peptide ending in ETGT, begins at Q14, a mutated S1/S2 cleavage site (SGAR), ends at residue K1211 and is followed by a TEV cleavage site, fold-on trimerization motif, and an hexa-histidine tag in the pCMV vector.

The SARS-CoV-2 HexaPro construct was previously described (*54*) and placed into CMVR with an octa-histidine tag.

The SARS-CoV-2 Wuhan-Hu-1 RBD (*4*) was synthesized by GenScript into pcDNA3.1 and contains an N-terminal mu-phosphatase signal peptide, a C-terminal

octa-histidine tag, flexible linker, and avi tag (GHHHHHHHHGGSSGLNDIFEAQKIEWHE). The boundaries of the construct are 328RFPN331 and 528KKST531.

The AncAsia-RBD sequence was taken from Starr et al. (*34*). The AncAsia-RBD-Avi construct was synthesized by GenScript into CMVR with an N-terminal mu-phosphatase signal peptide and a C-terminal octa-histidine tag, flexible linker, and avi tag (GHHHHHHHHGGSSGLNDIFEAQKIEWHE). The boundaries of the construct are 328RFPN331 and 528KLST531 (SARS-CoV-2 numbering).

SARS-CoV-2 S-D614G used for refolding experiments contains a mu-phosphatase signal peptide beginning at Q14, a mutated S1/S2 cleavage site (SGAR), ends at residue K1211 followed by a TEV cleavage site, foldon trimerization motif, and an octa-histidine tag in the pCMV vector.

SARS-CoV-2 S HexaPro, SARS-CoV-2 S D614G, SARS-CoV-2 Wuhan-Hu-1 RBD and AncAsia RBD were produced in Expi293F Cells (ThermoFisher Scientific) grown in suspension using Expi293 Expression Medium (ThermoFisher Scientific) at 37°C in a humidified 8% CO2 incubator rotating at 130 rpm. Cells grown to a density of 3 million cells per mL were transfected using the ExpiFectamine 293 Transfection Kit (ThermoFisher Scientific) and cultivated for 3-5 days. Proteins were purified from clarified supernatants using a nickel HisTrap HP affinity column (Cytiva) and washed with ten column volumes of 20 mM imidazole, 25 mM sodium phosphate pH 8.0, and 300 mM NaCl before elution on a gradient to 500 mM imidazole. Proteins were buffer exchanged into 20 mM sodium phosphate pH 8 and 100 mM NaCl and concentrated using centrifugal filters (Amicon Ultra) before being flash frozen.

Recombinant BtKY72 RBD production

As previously described, BtKY72 RBD construct (BtKY72 S residues 318-520) was synthesized by GenScript into a CMVR plasmid with a N-terminal mu-phosphatase signal peptide and a C-terminal hexa-histidine tag (-HHHHHHHH) joined by a short linker (-GGSS) to an Avi tag (-GLNDIFEAQKIEWHE) (*Starr, 2022*). BtKY72 mutant construct K493Y/T498W (BtKY72 S residue 482/487) were subcloned by GenScript from the BtKY72 RBD construct. BtKY72 and BtKY72 mutant RBD constructs were produced in Expi293F cells in Gibco Expi293 Expression Medium at 37°C in a humidified 8% CO2 incubator rotating at 130 rpm. The cultures were transfected using PEI-25K with cells grown to a density of 3 million cells per mL and cultivated for 3-5 days. Proteins were purified from clarified supernatants using a 1mL HisTrap HP affinity column (Cytiva), concentrated, and then biotinylated with a commercial BirA kit (Avidity). Proteins were then purified from the BirA enzyme by affinity purification using a 1 mL HisTrap HP affinity column (Cytiva), concentrated, and flash frozen in 1x PBS, pH 7.4 as previously described (*34*).

Biolayer Interferometry (BLI) analysis of mAb binding to BtKY72 RBDs

Experiments were carried out on a Octet Red96 (Forte Bio) and all samples were prepared in 10x Kinetics Buffer (Forte Bio). To assess S2K146, S2X259, S2E12 association with wildtype BtKY72, BtKY72 K493Y/T498W, SARS-CoV-2 RBD, RBDs were immobilized on Ni-NTA biosensors (Forte Bio) that were hydrated in water for 10 min. The biosensors were equilibrated to the kinetics buffer for 60 seconds prior to loading of the RBDs. Loading occurred in 10-20 µg/mL RBD and normalized to a 1 nm shift, then a baseline signal was established in 10x Kinetics Buffer for 120 seconds.

Association with S2K146, S2X259, and S2E12 was performed in 200 ug/mL of IgG, and dissociation was performed in 10x Kinetics buffer. Reference subtraction was performed with no load controls used to assess non-specific association between IgG solutions and the biosensor. Figures were plotted in Prism.

Antibody isolation and recombinant production

Antigen specific IgG+ memory B cells were isolated and cloned from PBMC of SARS-CoV-2 convalescent individuals. Briefly, CD19+ B cells were enriched by staining with CD19 PE-Cy7 and anti-PE microbeads (Milteniy), followed by positive selection using LS columns. Enriched B cells were stained with anti-IgD, anti-IgM, anti-IgA, anti-CD14,all PE labelled and prefusion SARS-CoV-2 S-Avi tag conjugated with streptavidin Alexa-Fluor 647 (Life Technologies). SARS-CoV-2-specific IgG+ memory B cells were sorted4 and seeded on MSC (mesenchymal stromal cells) at 0.5 cell/well in the presence of CpG2006, IL-2, IL6, IL-10 and IL-21, as previously described (*55*). After 7 days, B cell supernatants were screened by ELISA for binding to a panel of RBDs representative of different sarbecovirus clades as well as by neutralization using high-throughput VSV SARS-CoV-2 S-abesed microneutralization. Abs VH and VL sequences were obtained by reverse transcription PCR (RT-PCR) and mAbs were expressed as recombinant human IgG1, carrying the half-life extending M428L/N434S (LS) mutation in the Fc region fragment. ExpiCHO cells were transiently transfected with heavy and light chain expression vectors as previously described (*Tortorici, 2021*). Using the Database IMGT (http://www.imgt.org), the VH and VL gene family and the number of somatic mutations were determined by analyzing the homology of the VH and VL sequences to known human V, D and J genes. UCA sequences of heavy and light

variable regions were constructed using IMGT/V-QUEST. MAbs affinity purification was performed on ÄKTA Xpress FPLC (Cytiva) operated by UNICORN software version 5.11 (Build 407) using HiTrap Protein A columns (Cytiva) for full length human and hamster mAbs and CaptureSelect CH1-XL MiniChrom columns (ThermoFisher Scientific) for Fab fragments, using PBS as mobile phase. Buffer exchange to the appropriate formulation buffer was performed with a HiTrap Fast desalting column (Cytiva). The final products were sterilized by filtration through 0.22µm filters and stored at 4ºC

Enzyme-linked immunosorbent assay

96 half area well-plates (Corning, 3690) were coated over-night at 4°C with 25 µl of sarbecoviruses RBD proteins WIV1 (AGZ48831.1), Anlong-112 (ARI44804.1), YN2013(AIA62330.1), SC2018 (QDF43815.1), ZC45 (AVP78031.1), Rp/Shaanxi2011 (AGC74165.1), BM48-31/BGR/2008 (YP_003858584.1), RaTG13 (QHR63300.2), SARS-CoV2 (YP_009724390.1), SARS-CoV Urbani (AAP13441.1), BtKY72 (APO40579.1), Pangolin-Guangdong-2019 (EPI_ISL_410721), Pangolin_Guanxi-2017 (EPI_ISL_410539) and SARS-CoV-2 RBD mutants, prepared at 5 µg/ml in PBS pH 7.2. After a blocking step of 60 min at room temperature with PBS 1% BSA (Sigma-Aldrich, A3059), plates were incubated with mAb serial dilutions for 60 min at room temperature. After 4 washing steps with PBS 0.05% Tween 20 (PBS-T) (Sigma-Aldrich, 93773), goat anti-human IgG secondary antibody (Southern Biotech, 2040-04) was added and incubated for 45 min at room temperature. Plates were then washed 4 times with PBS-T and 4-NitroPhenyl phosphate (pNPP, Sigma-Aldrich, 71768) substrate was added. After 45 min incubation, absorbance at 405 nm was measured by a plate reader (Biotek) and

data plotted using Prism GraphPad.

30 µL of 5 µg/mL of SARS-CoV-2 AncAsia RBD diluted in PBS was incubated on a 384-well Nunc Maxisorp plate (ThermoFisher 464718) overnight at 4˚C. Plates were slapped dry before addition of 60 µL blocker Casein in PBS (ThermoFisher) and incubation for one hour at 37˚C. Plates were slapped dry and a 1:4 serial dilution of mAbs beginning at 300 µg/mL was added and incubated for one hour at 37˚C. Plates were slapped dry and washed 4x with TBST using a BioTek plate washer followed by5 addition of Invitrogen anti-Human IgG (ThermoFisher A18817) and one hour incubation at 37˚C. Plates were once again slapped dry and washed 4x with TBST before addition of room temperature TMB Microwell Peroxidase (Seracare 5120-0083). The reaction was quenched after 1-2 minutes with 1 N HCl and the A450 of each well was read using a BioTek plate reader. Prism (GraphPad) nonlinear regression with “Sigmoidal, 4PL, X is concentration” was used to determine the ED50 of each sample.

Transient Expression of sarbecovirus S protein in ExpiCHO-S Cells

ExpiCHO cells were seeded at 6 × 106 cells/ml into 50 ml bioreactor tubes in 5 ml culture medium. Spike coding plasmids (5 µg) were diluted in OptiPRO SFM, mixed with ExpiFectamine CHO Reagent (Life Technologies) and added to the cells. After transfection, cells were incubated at 37°C with 8% CO2 with an orbital shaking speed of 120 rpm (orbital diameter of 25 mm) for 48 hours.

Binding to cell surface expressed sarbecovirus S proteins by flow cytometry

Transiently transfected ExpiCHO cells were harvested and washed in wash buffer (PBS 2% FBS, 2 mM EDTA). Cells were counted, distributed into round bottom 96-well plates (Corning) and incubated with serially diluted antibodies in wash buffer (starting

concentration: 10 µg/ml, 8 points of dilution 1:4). Alexa Fluor647-labeled Goat Anti-Human IgG secondary Ab (Jackson Immunoresearch) was prepared at 2 µg/mL added onto cells after two washing steps. Cells were then washed twice and resuspended in wash buffer for data acquisition at ZE5 cytometer (Biorad).

Competition assay and affinity determination by Biolayer Interferometry (BLI)

BLI experiments were carried out using an Octet Red96 (ForteBio) and all reagents were prepared in Kinetics buffer (KB) (PBS 0.01% BSA). To assess S2K146 competition with S2X259, S309 and S2E12, His-tagged SARS-CoV-2 RBD was prepared at 8 µg/ml in Kinetics buffer (PBS 0.01% BSA) and loaded on pre-hydrated anti-penta-HIS biosensors (Sartorius) for 2.5 min. Biosensors were then moved into a solution containing 20 µg/ml S2K146 mAb and association recorded for 5 min. A second association step was subsequently performed into S2X259, S309 and S2E12 mAbs solutions at 20 µg/ml and recorded for 5 min. Response values were exported and plotted using GraphPad Prism (version 9.1.1). To assess binding affinities, S2K146 and respective UCA Ab were prepared at 3 µg/ml and immobilized on pre-hydrated protein-A biosensors (Sartorius) for 75 sec. After a 30 sec stabilization step in KB, biosensors were moved in SARS-CoV or SARS-CoV-2 :2 dilution series (starting concentration: 18.5 nM) for the 600 sec association step, and then moved back in KB to record dissociation signals for 540 sec. The data were baseline subtracted, results fitted using the Pall FortéBio/Sartorius analysis software (version 12.0)
and plotted using GraphPad Prism (version 9.1.1)

VSV-based pseudovirus production and neutralization assay

VSV psudoviruses were produced using the following constructs: SARS-CoV S, SARS-CoV-2 S, WIV-1 S, RaTG13 S, PG-GD S, PG-GX S, the VOC Alfa, Beta, Gamma, Epsilon, the VOI B.1.1.519, B.1.1.218 and SARS-CoV2 S bearing the single mutations K417V, E484K. Pseudotyped viruses were prepared using Lenti-X 293 cells seeded in 15-cm dishes. Briefly, cells in culture medium (DMEM supplemented with 10% heat-inactivated FBS, 1% PenStrep) were transfected with 25 µg of plasmid encoding for the corresponding S glycoprotein using TransIT-Lenti (Mirus) as transfectant reagent. One day post-transfection, cells were infected with VSV (G*ΔG-luciferase) for 1 h, washed 3 times in PBS with Ca2+/Mg2+ (Thermo Fisher) before adding 25 ml of culture medium/dish.

Particles were harvested after 18-24 h, clarified from cellular debris by centrifugation at 2,000 x g for 20 min at 4°C, aliquoted and stored at −80°C until use in neutralization experiments. For neutralization experiments, Vero E6 cells were seeded at 20,000 cells/well in culturemedium into white 96-well plates (PerkinElmer, 6005688) and cultured overnight at 37°C with 5% CO2. Ten-point 3-fold mAb serial dilutions were prepared in culture medium and mixed 1:1 with pseudotyped VSV prepared in culture medium in order to infect cells with the desired MOI. After 60 min incubation at 37 °C, cell culture medium was aspirated and 50µl of PVs/mAb mixture was added onto cells and incubated 60 min at 37°C 5% CO2. After 60 min, 100 µl of culture medium was added to the cells and incubation at 37°C 5% CO2 followed in the next 16-24 h. At the end of the incubation time, culture medium was removed from the cells and 50 µl/well of Steadylite (PerkinElmer) diluted 1:2 with PBS with Ca2+Mg2+ was added to the cells and incubated in the dark for 10 min. Luminescence signals were read using a Synergy

H1 Hybrid Multi-Mode plate reader (Biotek). Measurements were done in duplicate and at least six wells per plate contained untreatedinfected cells (defining the 0% of neutralization, “MAX RLU” value) and infected cells in the presence of S2E12 and S2X259 at 25 µg/ml each (defining the 100% of neutralization,“MIN RLU” value). Average of Relative light units (RLUs) of untreated infected wells (MAX RLUave) was subtracted by the average of MIN RLU (MIN RLUave) and used to normalize percentage of neutralization of individual RLU values of experimental data according to the following formula: (1-(RLUx - MIN RLUave) / (MAX RLUave – MIN RLUave)) x 100. Data were analyzed and visualized with Prism (Version 9.1.1). IC50 values were calculated from the interpolated value from the log(inhibitor) versus response, using variable slope (four parameters) nonlinear regression with an upper constraint of ≤100, and a lower constrain equal to 0.

BtKY72 (K493Y/T498W) S pseudovirus production and neutralization assay

VSV pseudovirus harboring BtKY72 K493Y/T498W S (based on SARS-CoV-2 numbering) with a native signal peptide and C-terminal 21 residue deletion synthesized by GenScript were prepared as previously described (*Starr, 2022*). Briefly, HEK-293T cells seeded in poly-D-lysine coated 100 mm dishes at ~75 % confluency were washed five times with Opti-MEM and co-transfected with Lipofectamine 2000 (Life Technologies) with 24 µg of the S glycoprotein plasmids. After 5 h at 37°C, media supplemented with 20% FBS and 2% PenStrep was added. After 20 hours, cells were washed five times with DMEM and cells were transduced with VSVΔG-luc (*56*) and incubated at 37°C. After 2 h, infected cells were washed an additional five times with DMEM prior to adding media supplemented with anti-VSV-G antibody (I1-mouse

hybridoma supernatant diluted 1:25, from CRL-2700, ATCC) to reduce parental background. After 18-24 h, the supernatant was harvested and clarified by low-speed centrifugation at 2,500 g for 10 min. The supernatant was then filtered (0.45 μm) and concentrated 10 times using a 30 kDa cut off membrane. The pseudotypes were then aliquoted and frozen at -80 °C.

For neutralization experiments, HEK-293T cells expressing hACE2 (Crawford et al. 2020) in DMEM supplemented with 10% FBS and 1% PenStrep were seeded at 20,000 cells per well into clear bottom, white manually poly-D-lysine coated 96 well plates and incubated at 37°C. The following day, an additional half-area, 96-well plate was prepared with twelve 3-fold serial dilutions of mAb of either S2K146 or S2X259 starting at 250ug/mL or 25 ug/mL respectively. An equal volume of diluted pseudovirus was added and incubated for 30 minutes at room temperature. Excess media was removed from cells and the mAb-pseudovirus mixture was transferred to the cells for 2 hours at 37C. After the 2 hour incubation, an equal volume of DMEM-20%FBS-2%PenStrep was added for overnight incubation. The following day, One-Glo-EX substrate (Promega) was added and incubated in the dark for 5 minutes. The plates were immediately read on a Biotek plate reader. Relative luciferase units were plotted and normalized in Prism with cells alone without pseudovirus defining 100% neutralization and cells with pseudovirus only defining 0% neutralization. Data were analyzed and visualized with Prism (Version 9.1.1). IC50 values were calculated from the interpolated value from the log(inhibitor) versus response, using variable slope (four parameters) nonlinear regression with an upper constraint of ≤100, and a lower constrain equal to 0.

Authentic SARS-CoV-2 isolates

SARS-CoV-2 strains used in this study were obtained from BEI (WT: Lineage A, BEI ref. NR-5228; Alpha: Lineage B.1.1.7, BEI ref. NR-54000; Beta: Lineage B.1.351, BEI ref. NR-54009; Kappa: Lineage B.1.617.1, BEI ref. NR-55486; Delta: Lineage B.1.617.2, BEI ref. NR-55611).

Neutralization of authentic SARS-CoV-2 viruses

Vero-TMPRSS2 cells were seeded into black-walled, clear-bottom 96-well plates at 2 × 104 cells/well and cultured overnight at 37°C. The next day, 9-point 5-fold serial dilutions of mAbs were prepared in infection media (DMEM + 10% FBS). The different SARS-CoV-2 strains were diluted in infection media at a final MOI of 0.01 PFU/cell, added to the mAb dilutions and incubated for 30 min at 37°C. Media was removed from the cells, mAb-virus complexes were added and incubated at 37°C for 18 h. Cells were fixed with 4% PFA (Electron Microscopy Sciences, #15714S), permeabilized with Triton X-100 (SIGMA, #X100-500ML) and stained with an antibody against the viral nucleocapsid protein (Sino Biologicals, #40143-R001) followed by a staining with the nuclear dye Hoechst 33342 (Fisher Scientific, # H1399) and a goat anti-rabbit Alexa Fluor 647 antibody (Invitrogen, #A-21245). Cells were imaged with an automated multimode plate reader (Biotek, Cytation 5).

Blockade of SARS-CoV and SARS-CoV-2 binding to ACE2

SARS-CoV and SARS-CoV-2 mouse/rabbit Fc-tagged RBDs (final concentration 20 ng/ml) were incubated with serially diluted recombinant mAbs (from 25 µg/ml) and incubated for 1 h 37°C. The complex RBD:mAbs was then added to a pre-coated hACE28 (2 µg/ml in PBS) 96-well plate MaxiSorp (Nunc) and incubated 1 hour at room temperature. Subsequently, the plates were washed and a goat anti-mouse/rabbit IgG

(Southern Biotech) coupled to alkaline phosphatase (Jackson Immunoresearch) added to detect mouse Fc-tagged RBDs binding. After further washing, the substrate (p-NPP, Sigma) was added, and plates read at 405 nm using a microplate reader (Biotek). The percentage of inhibition was calculated as follow: (1-((OD sample-OD neg ctr)/(OD pos. ctr-OD neg. ctr))*100

Cell-surface mAb-mediated S1 shedding

CHO cells stably expressing the prototypic SARS-CoV-2 Spike protein were harvested, washed in wash buffer (PBS 1% BSA 2 mM EDTA) and resuspended in PBS. Cells were then counted and 90'000 cells/well were dispensed into a round-bottom 96 well plate (Corning) to be treated with 10 ug/ml TPCK-Trypsin (Worthington Biochem) for 30 min at 37°C. After a washing step, cells were incubated with 15 ug/ml mAbs solution for 180, 120, 60, 30 or 5 min at 37°C. After the incubation for the allotted time, cells were washed with ice-cold wash buffer and stained with 1.5 ug/ml Alexa Fluor647-labeled Goat Anti-Human IgG secondary Ab (Jackson Immunoresearch) for 30 min on ice in the dark. Cells were then washed twice with cold wash buffer and analyzed using a ZE5 cytometer (Biorad) with acquisition chamber T= 4°C. Binding at each time point (MFI) was determined normalizing to the MFI at 5 minutes time point and data plotted using GraphPad Prism v. 9.1.1

Evaluation of escape mutants via deep mutational scanning

A previously described deep mutational scanning approach (*57*) was used to

identify RBD mutations that escape S2K146 binding exactly as described in (*20*). Briefly,duplicate libraries containing virtually all possible amino acid changes compatible withACE2 binding and RBD folding within the Wuhan-Hu-1 SARS-CoV-2 RBD sequence were expressed on the surface of yeast. Libraries were labelled at 63 ng/mL S2K146 or 83 ng/mL S2K146 UCA antibody (the EC90 for binding to yeast-displayed SARS-CoV-2 RBD determined in isogenic pilot binding experiments), and fluorescence-activated cell sorting (FACS) was used to select RBD+ cells that exhibit reduced antibody binding as previously described. Libraries were sequenced before and after selection to determine per-mutation escape fractions as previously described. Experiments were performed in duplicate with independently generated mutant libraries, and we report the average mutant escape fraction across the duplicates. Data for S2E12 exactly as described in (*20*) are included for comparison to S2K146. Complete computational pipeline for deep mutational scanning data analysis is available on GitHub: https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_S2K146.

Evaluation of sarbecovirus cross-reactivity via high-throughput yeast-display binding assays

Evaluation of mAb cross-reactivity with a panel of yeast-displayed sarbecovirus RBDs was performed as described with modifications (*20*). Whereas we previously assessed mAb binding to the sarbecovirus RBD panel at a single concentration (analogous to DMS selections), binding was here determined via FACS-seq for a mAb dilution series (10,000 to 0.01 ng/mL in 10-fold dilutions, plus 0 ng/mL mAb). The ‘escape9 fraction’ of each sarbecovirus RBD at each mAb concentration was fit to a sigmoid binding curve to determine the quantitative EC50 for antibody binding to each

RBD. In addition to S2K146 breadth, S2E12 was profiled using this new method in this paper to enable direct comparison. Complete computational pipeline for the analysis of breadth of mAb binding is available on GitHub: https://github.com/jbloomlab/SARSr-CoV_RBD_MAP.

Selection of SARS-CoV-2 monoclonal antibody escape mutants

A VSV-SARS-CoV-2 Wuhan-Hu-1 D614G S chimera was used to select for mAb resistant mutants, as previously described(*Case, 2020*). Briefly, mutants were recovered by plaque isolation on Vero E6 cells with the indicated mAb in the overlay. The concentration of mAb in the overlay was determined by neutralization assays at a multiplicity of infection (MOI) of 100. Escape clones were plaque-purified on Vero E6 cells in the presence of mAb, and plaques in agarose plugs were amplified on MA104 cells with the mAb present in the medium. Viral stocks were amplified on MA104 cells at an MOI of 0.01 in Medium 199 containing 2% FBS and 20 mM HEPES pH 7.7 (Millipore Sigma) at 34°C. Viral supernatants were harvested upon extensive cytopathic effect and clarified of cell debrisby centrifugation at 1,000 x g for 5 min. Aliquots were maintained at - 80°C. Viral RNA was extracted from VSV-SARS-CoV-2 S mutant viruses using RNeasy Mini kit (Qiagen), and the S gene was amplified using OneStep RT-PCR Kit (Qiagen). The mutations were identified by Sanger sequencing (GENEWIZ). Their resistance was verified by subsequent virus infection in the presence or absence of mAb. Vero E6 cells were seeded into 12 well plates overnight. The virus was serially diluted using DMEM and cells were infected at 37°C for 1 h. Cells were cultured with an agarose overlay in the presence or absence of mAb at 34°C for 2 days. Plates were

scanned on a biomolecular imager and expression of eGFP monitored at 48 hours post-infection.

Determination of binding affinity to Wuhan-hu-1 and Yusing surface plasmon resonance

Measurements were performed using a Biacore T200 instrument. A CM5 chip with covalently immobilized StrepTactin XT was used for surface capture of StrepTag-containing RBDs. Running buffer was HBS-EP+ pH 7.4 (Cytiva) and measurements were performed at 25℃. Experiments were performed with a 3-fold dilution series of monomeric hACE2: 300, 100, 33, 11 nM and were run as single-cycle kinetics. Data were double reference-subtracted and fit to a binding model using Biacore Evaluation software. The 1:1 binding model was used except for S2K146 Fab binding to Y489H which was fit to a Heterogeneous Ligand model (a low-amplitude kinetic phase with very slow dissociation was assumed to be an artifact; the higher amplitude, faster-dissociation kinetic phase was interpreted as the affinity of the Fab:RBD interaction; the resulting KD is reported as approximate).

Viral replication fitness assays

Vero E6 cells (ATCC, CRL-1586) were seeded at 1×106 cells per well in 6-well plates. Cells were infected with multiplicity of infection (MOI) of 0.02, with VSV chimeras harboring SARS-CoV-2 Wuhan-Hu-1 D614G S and SARS-CoV-2 Wuhan-Hu-1 Y489H/D614G S mixed at equal titers. Following 1 h incubation, cell monolayers were washed three times with Hanks' Balanced Salt Solution (HBBS) and cultures were incubated for 72 h in humidified incubators at 34°C. To passage the progeny viruses, virus mixture was continuously passaged four times in Vero E6 cells at MOI of 0.02.

Cellular RNA samples from each passage were extracted using RNeasy Mini kit (QIAGEN) and subjected to next-generation sequencing as described previously to confirm the introduction and frequency of substitutions (*35*).

S2K146-induced S refolding

10 μM native-like SARS-CoV-2 S(*Lempp, 2021*) was incubated with 13 μM S2K146 Fab for 12 hours at room temperature. Samples were diluted to 0.01 mg/mL immediately prior to adsorption to glow-discharged carbon-coated copper grids for ~30 sec prior to a 2% uranyl formate staining. Micrographs were recorded using the Leginon software on a 120 kV FEI Tecnai G2 Spirit with a Gatan Ultrascan 4000 4k x 4k CCD camera at 67,000 nominal magnification. The defocus ranged from -1.0 to -2.0 μm and the pixel size was 1.6 Å.

CryoEM sample preparation, data collection and data processing

Recombinantly expressed and purified S2K146 Fab and SARS-CoV-2 S hexapro(*54*) were incubated at 1 mg/ml (for UltraAuFoil grids) or 0.1mg/ml (for lacey thin carbon grids) with a 1.2 molar excess of Fab at 4°C for 1 hr. Three microliters of the complex mixture were loaded onto freshly glow discharged R 2/2 UltrAuFoil grids (200 mesh) or lacey grids covered with a thin layer of manually added carbon, prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 6-6.5 sec blot time (for the UltrAuFoil grids) or with a blot force of −1 and 2.5 sec blot time (for the lacey thin carbon grids) at 100 % humidity and 22°C. Data were acquired using an FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K3 direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using

Leginon (*58*) at a nominal magnification of 105,000x with a pixel size of 0.843 Å. The dose rate was adjusted to 15 counts/pixel/s, and each movie was acquired in super-resolution mode fractionated in 75 frames of 40 ms. 7,289 micrographs were collected with a defocus range comprised between -0.5 and -2.5 µm.

Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking, and extraction were carried out using Warp (*59*). Two rounds of reference-free 2D classification were performed using CryoSPARC (*60*) to select well-defined particle images. These selected particles were subjected to two rounds of 3D classification with 50 iterations each (angular sampling 7.5˚ for 25 iterations and 1.8˚ with local search for 25 iterations), using our previously reported closed SARS-CoV-2 S structure as initial model (PDB 6VXX) (*4*) using Relion (*61*). 3D refinements were carried out using non-uniform refinement along with per-particle defocus refinement in CryoSPARC (*62*). Selected particle images were subjected to the Bayesian polishing procedure (*63*) implemented in Relion3.0 before performing another round of non-uniform refinement in cryoSPARC followed by per-particle defocus refinement and again non-uniform refinement. To improve the density of the S/S2K146 interface, the particles from the class with 2 RBDs opened were subjected to focus 3D classification without refining angles and shifts using a soft mask on the closed RBD and bound S2K146 variable domains with a tau value of 60 in Relion. Particles belonging to classes with the best resolved local density were selected and subject to local refinement using CryoSPARC. Local resolution estimation, filtering, and sharpening were carried out using CryoSPARC. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion and Fourier shell

correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (*Chen*, 2013; Rosenthal, 2003).

Model building and refinement

UCSF Chimera (*Pettersen, 2004*) and Coot (*Emsley, 2010*) were used to fit atomic models into the cryoEM maps. Spike-RBD/S2K146 model was refined and relaxed using Rosetta using sharpened and unsharpened maps (*Frenz, 2019; Wang, 2016*). Model validation and analysis used
MolProbity (*Chen, 2010*), EMringer (*Barad, 2015*), Phenix (*Liebschner, 2019*) and Privateer (*Agirre, 2015*). Figures were generated using UCSF ChimeraX (*Goddard, 2018*).

Measurement of Fc-effector functions

S2K146-dependent activation of human FcγRIIa and IIIa was performed with a bioluminescent reporter assay. ExpiCHO cells stably expressing full-length wild-type SARS-CoV-2 S (target cells) or full-length prefusion stabilized SARS-CoV-2 S, which harbours the 2P mutation and S1/S2 furin cleavage site mutation (RRARS to SGAG) as previously described (*Walls, 2020*), were incubated with different amounts of mAbs. After a 15-minute incubation, Jurkat cells stably expressing FcγRIIIa receptor (V158 variant) or FcγRIIa receptor (H131 variant) and NFAT-driven luciferase gene (effector cells) were added at an effector to target ratio of 6:1 for FcγRIIIa and 5:1 for FcγRIIa. Signaling was quantified by the luciferase signal produced as a result of NFAT pathway activation. Luminescence was measured after 20 hours of incubation at 37˚C with 5% CO2 with a luminometer using the Bio-Glo-TM Luciferase Assay Reagent according to the manufacturer's instructions (Promega).

Hamster challenge experiment.

The hamster infection model of SARS-CoV-2 including the associated analytical procedures, have been described before (*Abdelnabi, 2021; Boudewijns, 2020*). In brief, female Syrian hamsters (Mesocricetus auratus) of 6-8 weeks old were anesthetized with ketamine/xylazine/atropine and inoculated intranasally with 50 µL containing 1×104 TCID50 Beta B.1.351 (derived from hCoV-19/Belgium/rega-1920/2021; EPI_ISL_896474, 2021-01-11). This variant was originally isolated in house from nasopharyngeal swabs taken from travelers returning to Belgium (baseline surveillance) and were subjected to sequencing on a MinION platform (Oxford Nanopore) directly from the nasopharyngeal swabs (*Abdelnabi, 2021*). Animals were treated once by intraperitoneal injection 24h post SARS-CoV-2 challenge (i.e. therapeutic administration) with S2K146 mAb (at 10, 5 and 1 mg/Kg). S2E12 mAb was administrated as control (at 1 mg/kg). Isotype control was administered at 10 mg/kg. Hamsters were monitored for appearance, behavior and weight. At day 4 pi, hamsters were euthanized by i.p. injection of 500 µL Dolethal (200mg/ml sodium pentobarbital, Vétoquinol SA). Lungs were collected for viral RNA and infectious virus quantification by RT-qPCR and end-point virus titration, respectively. Serum samples were collected at day 4 pi for analysis of Ab levels.

SARS-CoV-2 RT-qPCR

Hamster lung tissues were collected after sacrifice and were homogenized using bead disruption (Precellys) in 350 µL TRK lysis buffer (E.Z.N.A.® Total RNA Kit, Omega Bio-tek) and centrifuged (10.000 rpm, 5 min) to pellet the cell debris. RNA was extracted according to the manufacturer's instructions. RT-qPCR was performed on a

LightCycler96 platform (Roche) using the iTaq Universal Probes One-Step RT-qPCR kit (BioRad) with N2 primers and probes targeting the nucleocapsid (*Abdelnabi, 2021*). Standards of SARS-CoV-2 cDNA (IDT) were used to express viral genome copies per mg tissue.

End-point virus titrations

Lung tissues were homogenized using bead disruption (Precellys) in 350 µL minimal essential medium and centrifuged (10,000 rpm, 5min, 4°C) to pellet the cell debris. To quantify infectious SARS-CoV-2 particles, endpoint titrations were performed on confluent Vero E6 cells in 96- well plates. Viral titers were calculated by the Reed and Muench method (*75*) using the Lindenbach calculator and were expressed as 50% tissue culture infectious dose (TCID50) per mg tissue.

8.9 Figures and Tables

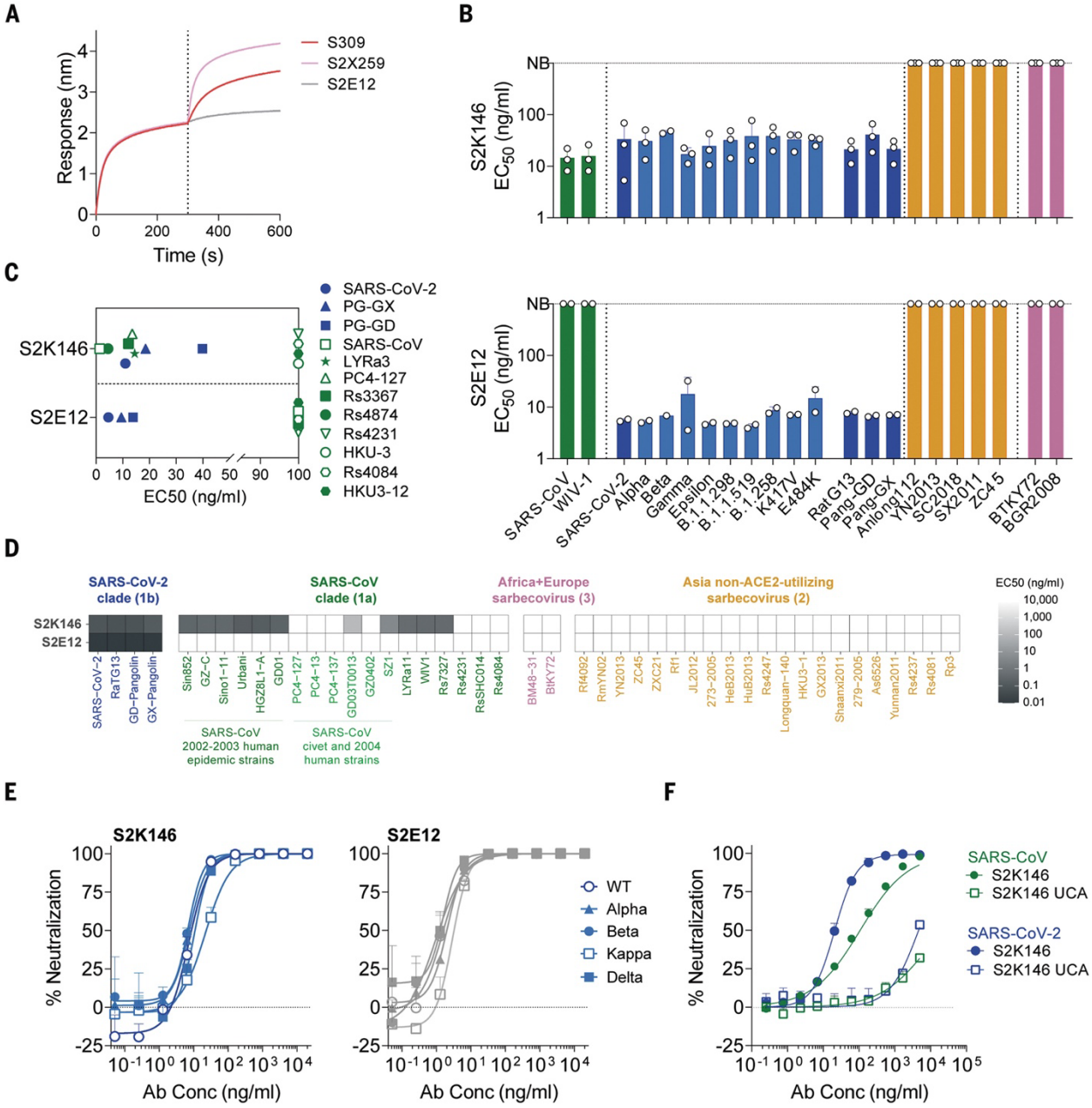


Figure 8.1: . Identification of the S2K146 cross-reactive and broadly neutralizing sarbecovirus mAb.

(**A**) Binding of site I–targeting S2E12 (Tortorici, 2020), site IV–targeting S309 (Tortorici, 2021), and site II–targeting S2X259 (Tortorici, 2021) recombinant IgG1 (second phase) after association of S2K146 mAb (first phase) to His-tagged SARS-CoV-2 RBD immobilized on anti-His sensors, as measured by biolayer interferometry. (**B**) Cross-reactivity of S2K146 (upper panel) and S2E12 (lower panel) with 22 sarbecovirus RBDs from four sarbecovirus clades and SARS-CoV-2 variants analyzed by ELISA. Median

effective concentration (EC_{50}) of at least two independent experiments is shown. Error bars indicate standard deviation between experimental repeats. (**C**) Flow cytometry analysis of S2K146 cross-reactivity with a panel of 12 S glycoproteins representative of sarbecovirus clades 1a and 1b transiently expressed on the surface of mammalian cells. (**D**) S2K146 cross-reactivity with sarbecovirus RBDs displayed at the surface of yeast. (**E**) S2K146- and S2E12-mediated neutralization of replication-competent SARS-CoV-2 (USA-WA1/2020) and SARS-CoV-2 VOC viruses. WT; wild type; Ab Conc, antibody concentration. (**F**) S2K146- and S2K146 UCA–mediated neutralization of VSV pseudotypes harboring SARS-CoV-2 S or SARS-CoV S. Error bars indicating standard deviation between replicates are represented only in one direction.

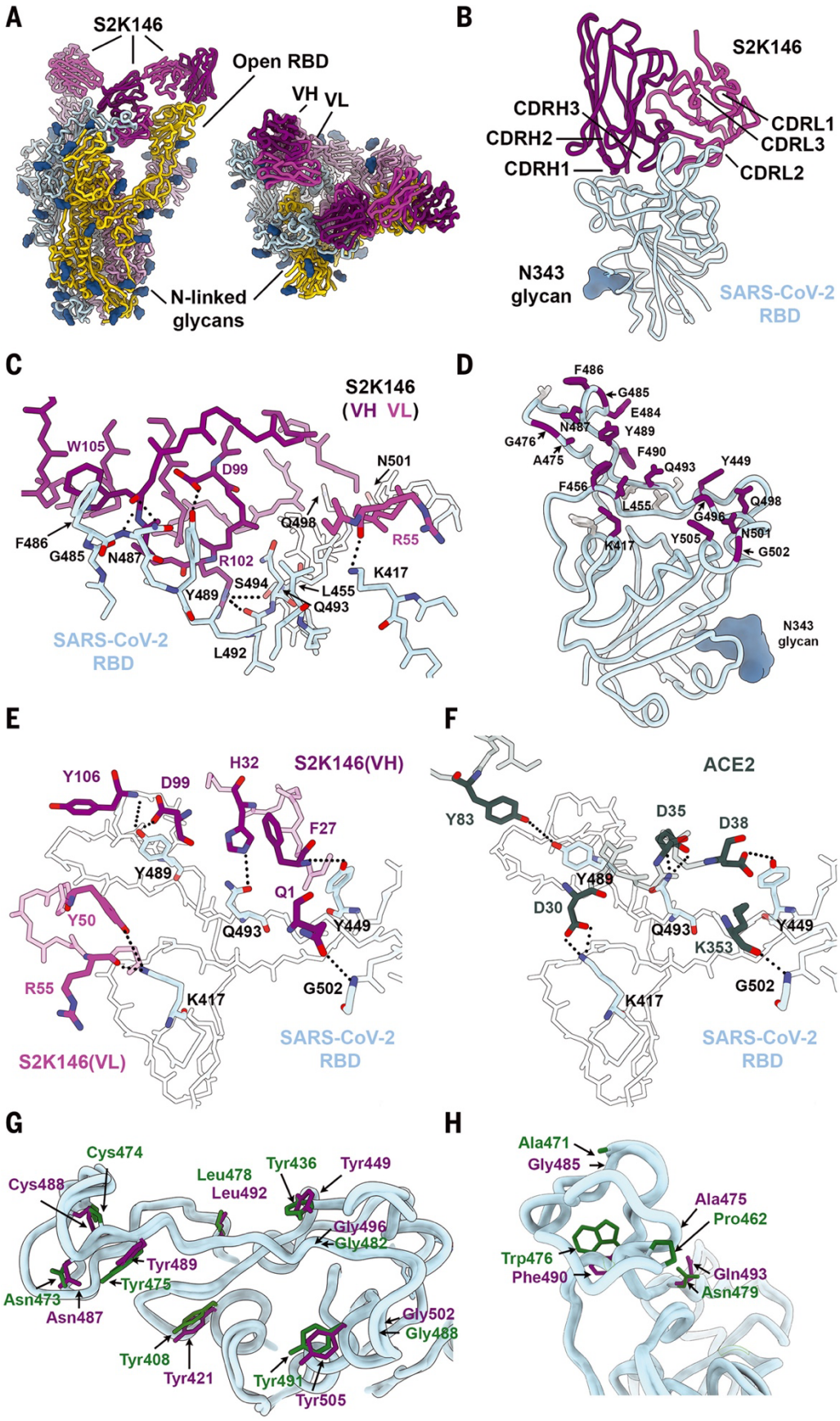

A
S2K146
Open RBD
VH VL
N-linked glycans
B
S2K146
CDRH3
CDRH2
CDRH1
CDRL1
CDRL3
CDRL2
N343 glycan
SARS-CoV-2 RBD
C
S2K146 (VH VL)
W105
D99
N501
Q498
R55
F486
G485
N487
R102
S494
K417
Y489
L455
Q493
SARS-CoV-2 RBD
L492
D
F486
G485
E484
N487
Y489
G476
A475
F490
Q493
Y449
F456
L455
Q498
G496
N501
K417
Y505
G502
N343 glycan
E
Y106
D99
H32
S2K146(VH)
F27
Y489
Q1
Y50
Q493
Y449
R55
K417
G502
S2K146(VL)
SARS-CoV-2 RBD
F
ACE2
Y83
D35
D38
Y489
D30
Q493
Y449
K353
G502
K417
SARS-CoV-2 RBD
G
Cys474
Cys488
Leu478
Leu492
Tyr436
Tyr449
Gly496
Gly482
Tyr489
Tyr475
Asn473
Asn487
Gly502
Gly488
Tyr408
Tyr421
Tyr491
Tyr505
H
Ala471
Gly485
Ala475
Pro462
Trp476
Phe490
Gln493
Asn479

Figure 8.2: The S2K146 broadly neutralizing mAb recognizes RBD antigenic site I.

(**A**) Cryo-EM structure viewed along two orthogonal orientations of the prefusion SARS-CoV-2 S ectodomain trimer with three S2K146 Fab fragments bound to two open RBDs and one partially closed RBD. SARS-CoV-2 S protomers are colored cyan, pink, and gold. S2K146 heavy chain and light chain variable domains are colored purple and magenta, respectively. Glycans are rendered as blue spheres. (**B**) Ribbon diagram of the S2K146-bound SARS-CoV-2 RBD. (**C**) Zoomed-in view of the contacts formed between S2K146 and the SARS-CoV-2 RBD. Selected epitope residues are shown as sticks, and electrostatic interactions are indicated with dotted lines. S2K146 heavy chain and light chain variable domains are colored as in (A). Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**D**) S2K146 epitope residues shown as sticks and colored purple (labeled) if they are involved in ACE2 binding or colored gray if not (unlabeled). (**E** and **F**) Similar electrostatic interactions (dotted lines) formed between S2K146 (E) or ACE2 (F) and the SARS-CoV-2 RBD. (**G**) The side chains of the nine S2K146 epitope residues conserved between the SARS-CoV-2 (purple) and SARS-CoV [green, Protein Data Bank (PDB) ID 2AJF (Li, 2005)] RBDs are shown as sticks. (**H**) The side chains of the four S2K146 epitope residues conservatively substituted between the SARS-CoV-2 (purple) and SARS-CoV (green) RBDs are shown as sticks.

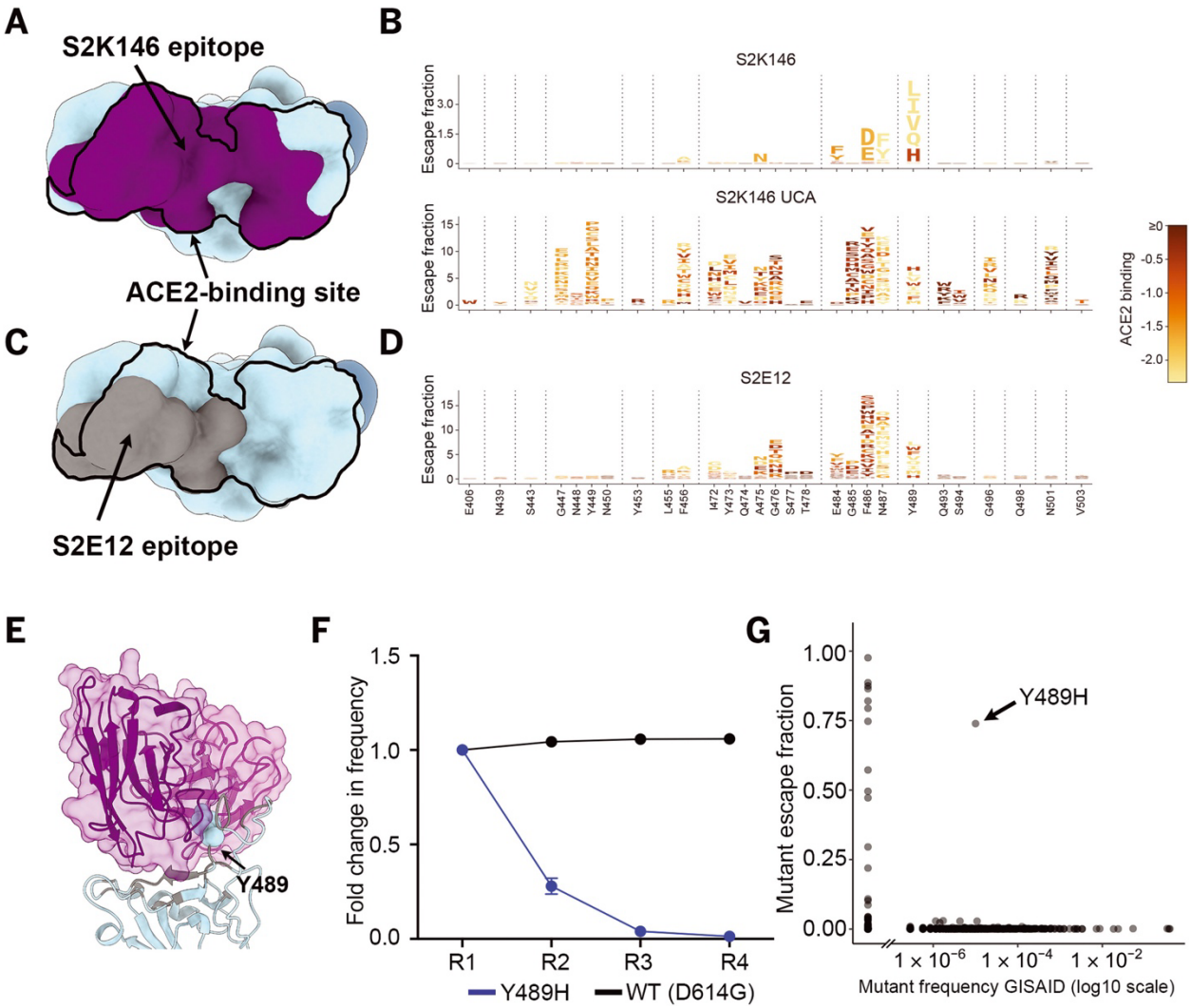


Figure 8.3: S2K146 is resilient to a broad spectrum of escape mutations.

(**A**) Molecular surface representation of the SARS-CoV-2 RBD with the S2K146 epitope colored purple and the ACE2 footprint indicated as a black outline. (**B**) Mapping of RBD mutations reducing S2K146 (top) or S2K146 UCA (bottom) binding using DMS of the yeast-displayed SARS-CoV-2 RBD. Sites of strong escape (pink underlines in Fig. S7A) are shown in logo plot. Letters are colored according to how mutations affect the ACE2 binding affinity of the SARS-CoV-2 RBD, as measured via yeast display by Starr *et al*. (Starr, 2020). (**C**) Molecular surface representation of the SARS-CoV-2 RBD with the S2E12 epitope colored gray and the ACE2 footprint indicated as a black outline. The N343 glycan is rendered as blue spheres in (A) and (C). (**D**) Mapping of RBD mutations reducing S2E12 binding using DMS of the yeast-displayed SARS-CoV-2 RBD. Sites of strong escape (purple underlines in Fig. S7A) are shown in logo plot, as measured previously by Starr *et al.* (Starr, 2021). (**E**) Zoomed-in view of the S2K146-bound SARS-CoV-2 RBD (blue) highlighting the Y489H neutralization escape mutation. The S2K146 heavy and light chain variable domains are shown as ribbons within transparent purple and magenta surfaces, respectively. (**F**) Viral replication competition between VSV

chimeras harboring the SARS-CoV-2 Wuhan-Hu-1/D614G S with or without the Y489H substitution using VeroE6 cells. (**G**) Mutations reducing binding of S2K146 to the RBD on the basis of DMS (escape score) are plotted versus their frequencies among the human-derived SARS-CoV-2 sequences on GISAID as of 27 September 2021. The large escape mutant (>5× global median escape fraction) with nonzero frequency is indicated.

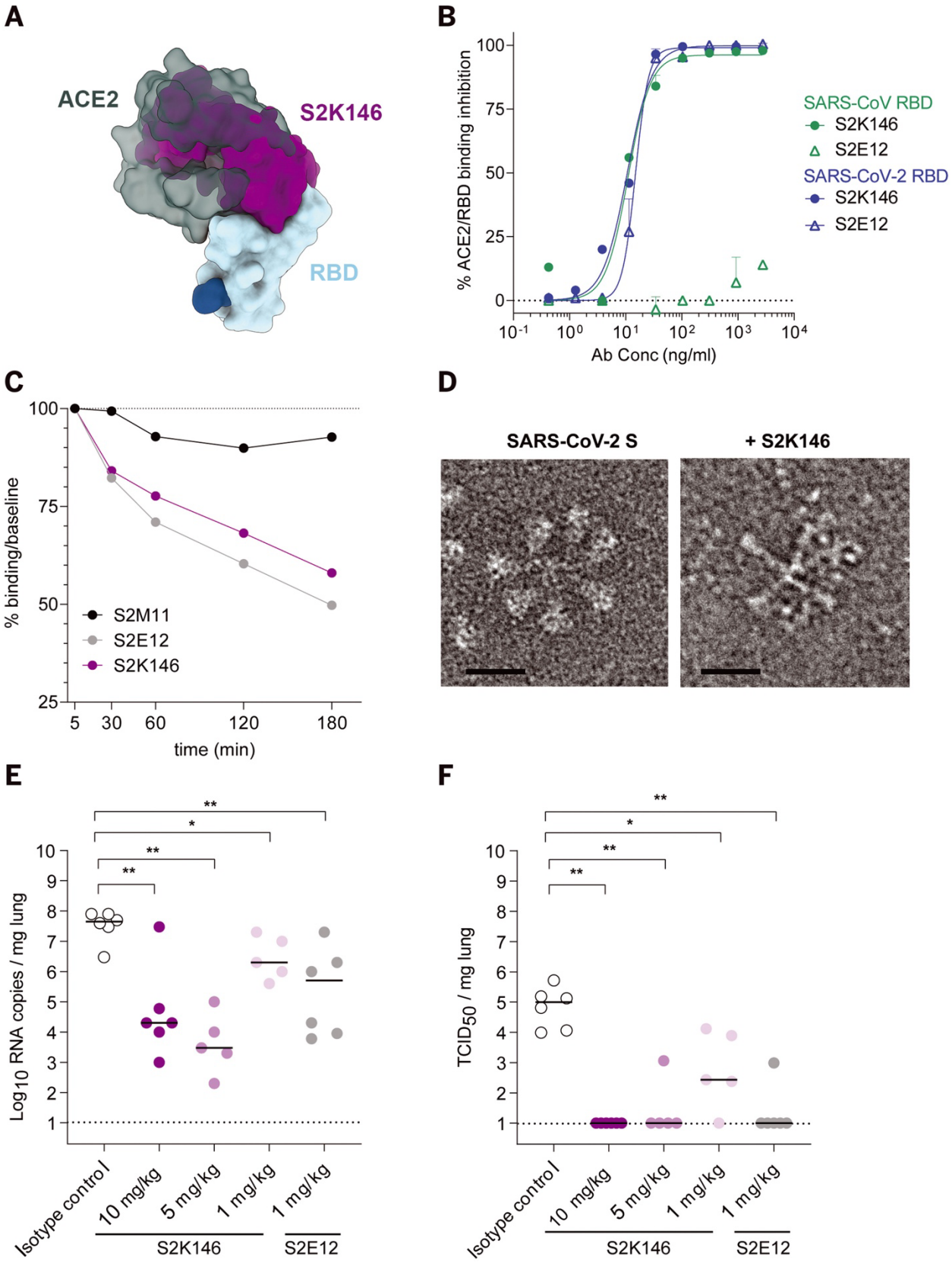
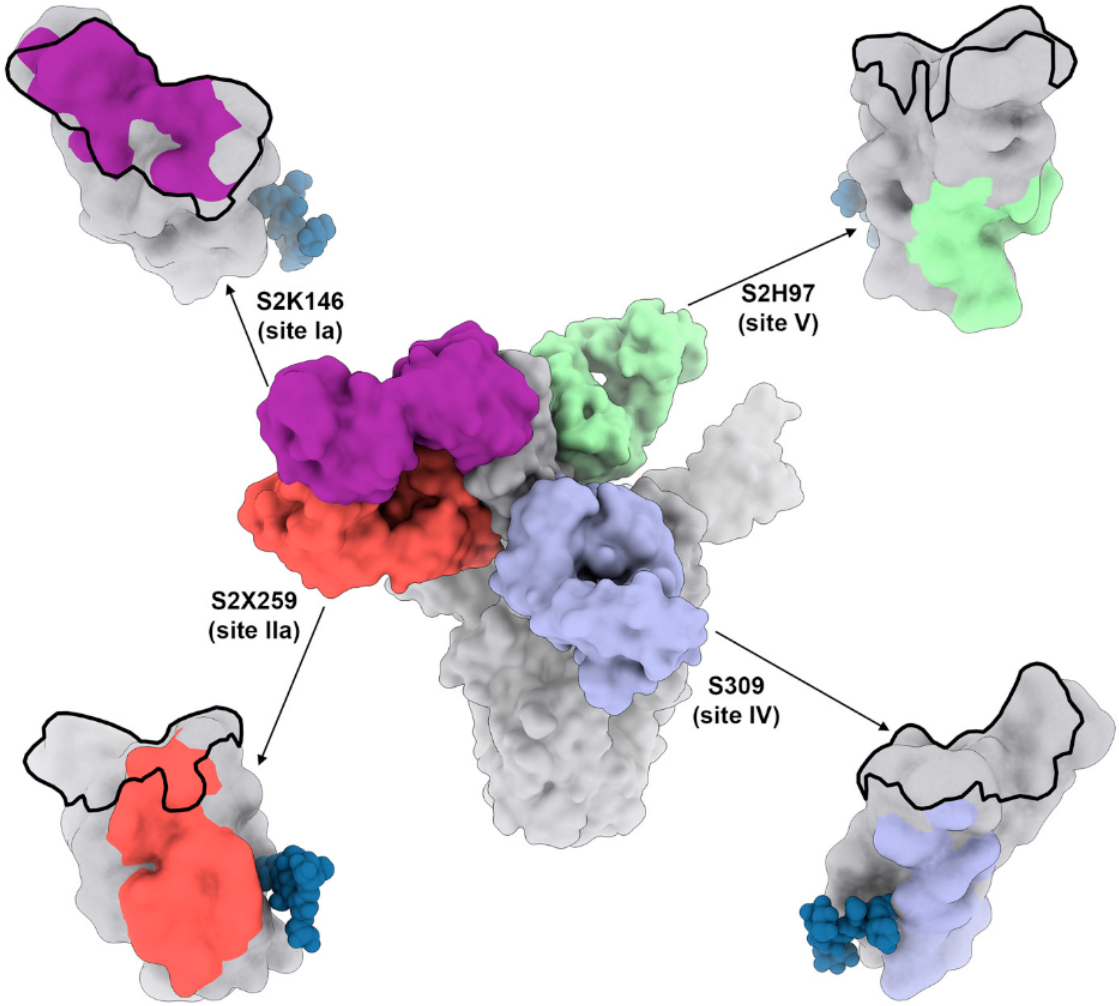

A
ACE2
S2K146
RBD
B
% ACE2/RBD binding inhibition
Ab Conc (ng/ml)
SARS-CoV RBD
S2K146
S2E12
SARS-CoV-2 RBD
S2K146
S2E12
C
% binding/baseline
time (min)
S2M11
S2E12
S2K146
D
SARS-CoV-2 S
+ S2K146
E
Log10 RNA copies / mg lung
Isotype control
10 mg/kg
5 mg/kg
1 mg/kg
1 mg/kg
S2K146
S2E12
F
TCID50 / mg lung
Isotype control
10 mg/kg
5 mg/kg
1 mg/kg
1 mg/kg
S2K146
S2E12

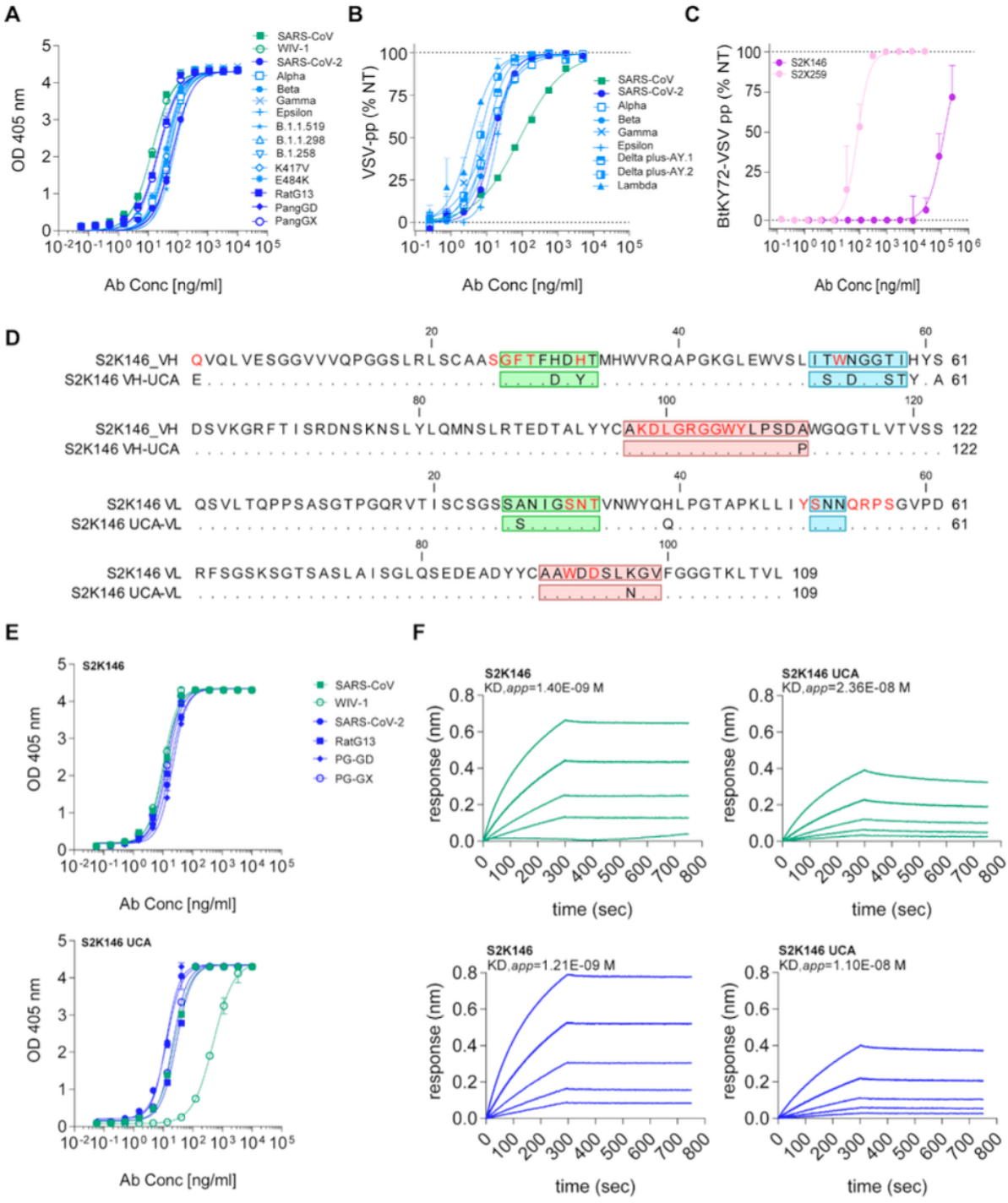
Figure 8.4: S2K146 blocks receptor attachment, triggers premature S refolding, and protects against SARS-CoV-2 challenge therapeutically.

(**A**) Superimposition of the S2K146-bound (purple) and ACE2-bound [dark gray, PDB ID 6M0J (Lan, 2020)] SARS-CoV-2 RBD (light blue) structures showing steric overlap. The N343 glycan is rendered as blue spheres. (**B**) Preincubation of serial dilutions of S2K146 with the SARS-CoV-2 RBD prevents binding to immobilized human ACE2 (hACE2) ectodomain in ELISA. Error bars indicate standard deviation between replicates. (**C**) S2K146-mediated S_1-shedding from cell surface–expressed SARS-CoV-2 S as determined by flow cytometry. S2E12 mAb was used as positive control, whereas S2M11 was used as a negative control. (**D**) Cropped electron micrographs of negatively stained SARS-CoV-2 S trimer before (left, prefusion state) or after (right, postfusion state) incubation with S2K146. One representative micrograph for each dataset is shown out of 93 (SARS-CoV-2 S alone) and 225 (SARS-CoV-2 S with S2K146) micrographs. Scale bars, 200 Å. (**E** and **F**) Quantification of viral RNA (E) and replicating virus titers [50% tissue culture infectious dose ($TCID_{50}$)] (F) in the lung of Syrian hamster 4 days after intranasal infection with SARS-CoV-2 Beta VOC followed by therapeutic administration of S2K146 mAb at three different doses: 10, 5, or 1 mg/kg of body weight (n = 5 or 6 animals per group). S2E12 mAb was administered as control (n = 6 animals). Isotype control was administered at 10 mg/kg (n = 6 animals). *P < 0.5, **P < 0.01, as determined by Mann Whitney two-tailed test.



Supplemental Figure 8.1: Overview of SARS-CoV-2 RBD cross-neutralizing antigenic sites.

Composite model of the SARS-CoV-2 S-protein trimer (gray) with four distinct mAbs (S2K146 in purple, S2X259 in red (*Tortorici, 2021*), S2H97 in green (*Starr, 2021*), and S309 in blue (*Pinto, 2020*)) bound to one RBD in the open conformation. The magnified panels show the footprint of each mAb on the SARS-CoV-2 RBD with matching color and the ACE2-binding site is outlined in black. The glycan at position N343, which is conserved across the sarbecovirus subgenus, is rendered as dark blue spheres.



Supplemental Figure 8.2: In vitro characterization of the S2K146 mAb

A) S2K146 mAb binding to clade 1a and 1b sarbecovirus RBDs analyzed by ELISA. **B)** S2K146-mediated neutralization of VSV pseudotypes harboring Wuhan-Hu-1 SARS-

CoV-2 S, B.1.351 (Beta) S, B.1.1.7 (Alpha) S, P.1 (Gamma) S, B.1.429 (Epsilon) S, C.37

(Lambda) S, AY.1/.2 (Delta+) S or SARS-CoV S. **C)** S2K146-mediated neutralization of

VSV pseudotypes harboring BtKY72 S (K493Y/T498W, SARS-CoV-2 residue numbering

(*34*)) using HEK293T cells stably expressing human ACE2 (*Crawford, 2020*). The S2X259 mAb was included as a positive control. **D**) Alignment of the amino acid sequence of the variable

regions of heavy and light chains of S2K146 and S2K146 UCA. Heavy and light chain

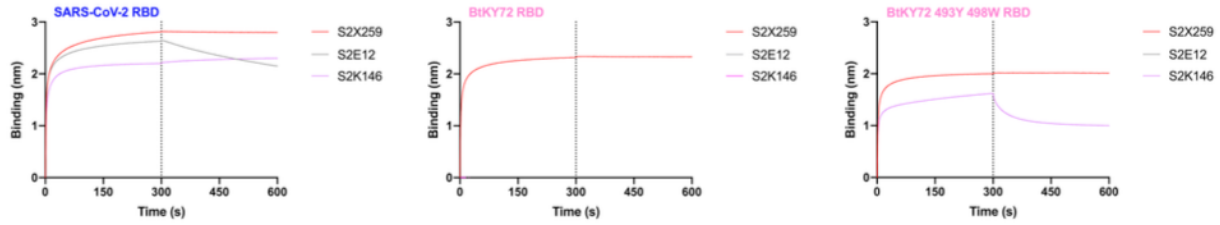
CDR1, CDR2, and CDR3 (IMGT definition) are indicated by green, blue, and red boxes,

respectively. Residues that make contact with the SARS-CoV-2 RBD are indicated in red.

E) ELISA binding analysis of S2K146 and S2K146 UCA to sarbecovirus clade 1a and

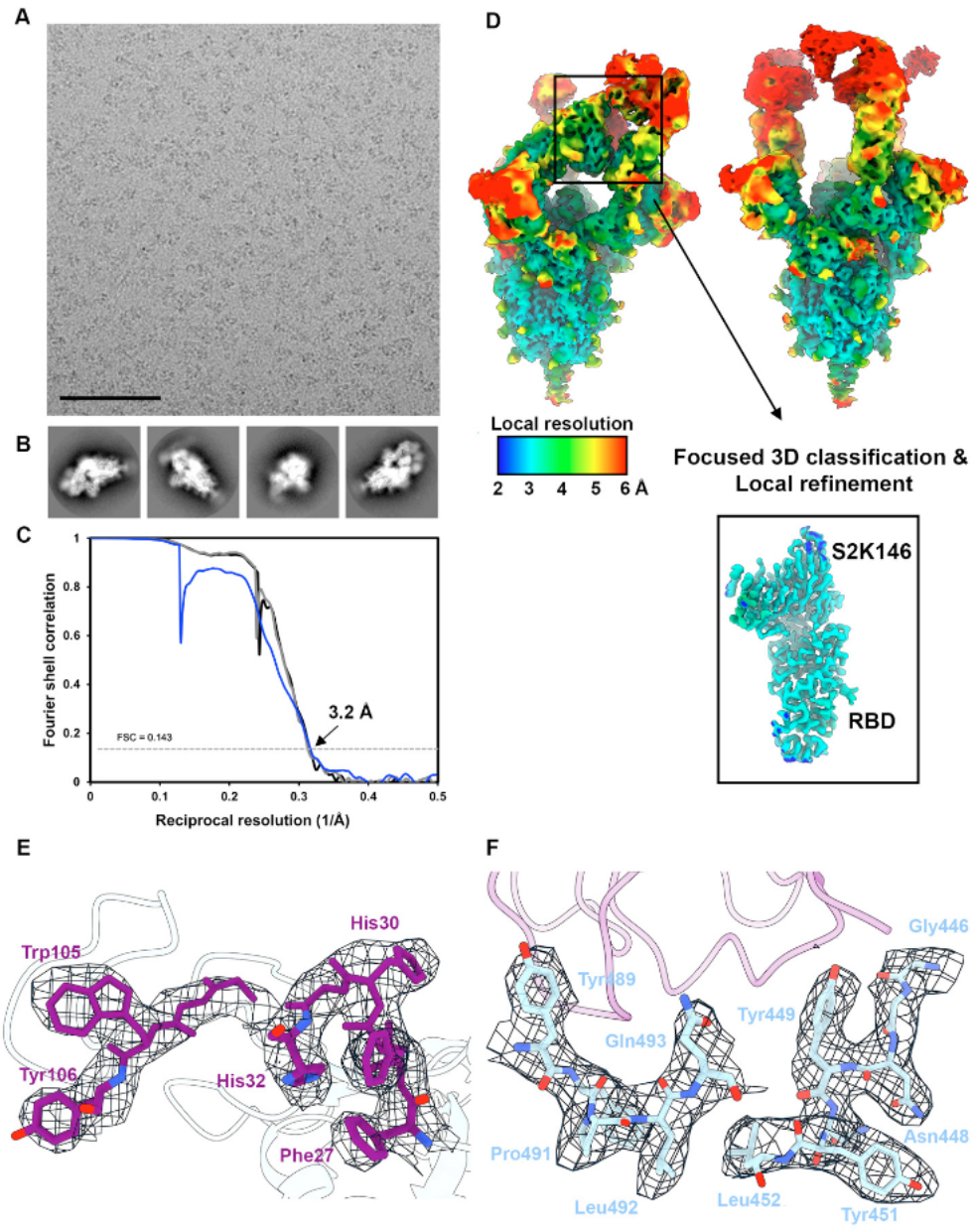
clade1b RBDs. **F)** Biolayer interferometry binding analysis of S2K146 and S2K146 UCA

IgG1 to prefusion SARS-CoV S (green) and SARS-CoV-2 S (blue). Apparent KD (KD,*app)* are reported.



Supplemental Figure 8.3: Evaluation of mAb binding to the BtKY72 RBD.

Binding of S2K146, S2E12 or S2X259 IgGs at 200 µg/mL to the SARS-CoV-2 Wuhan-Hu-1 RBD, the wildtype BtKY72 RBD or the BtKY72 K493Y/T498W RBD (defined based on SARS-CoV-2 numbering) immobilized at the surface of Ni-NTA biolayer interferometry biosensors.



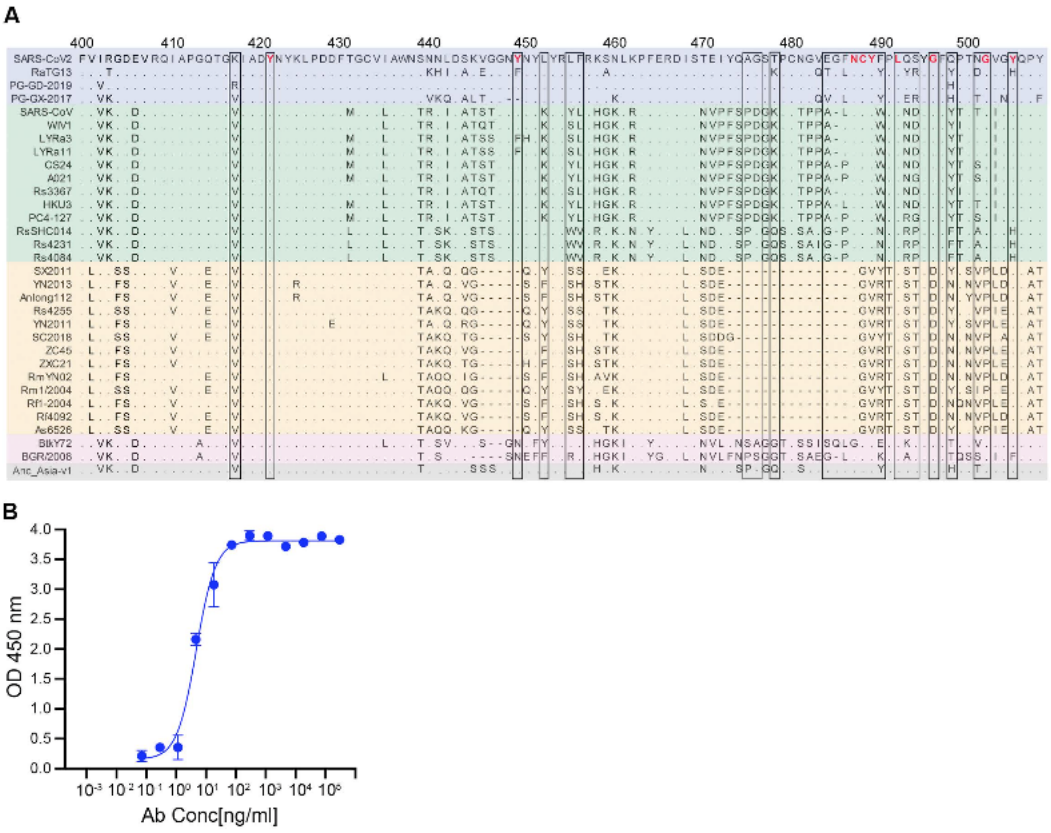
Supplemental Figure 8.4: CryoEM data processing of the S2K146-bound SARS-CoV-2 S dataset.

A-B) Representative electron micrograph (A) and 2D class averages (B) of SARS-CoV-2

S in complex with the S2K146 Fab embedded in vitreous ice. The scale bar represents 100 nm. **C)** Gold-standard Fourier shell correlation curves for the S2K146-bound SARS-CoV-2 S maps with three RBDs open (black line), two RBDs open (grey line) and locally

refined RBD/S2K146 variable domain (blue line). The 0.143 cutoff is indicated by a horizontal dashed line. **D)** Local resolution maps calculated using cryoSPARC for the whole SARS-CoV-2 S trimer reconstruction with two RBDs in the open state (left) and three RBDs in the open state (right) as well as for the locally refined RBD/S2K146 variable

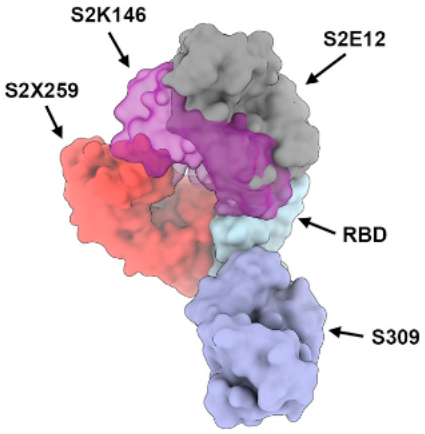
domain region (inset). **E-F)** CryoEM density (mesh representation) of the paratope (E) and epitope (F) with the corresponding model fit. The S2K146 variable heavy and light chains are shown in purple and magenta, respectively, whereas the SARS-CoV-2 RBD is light blue.



Supplemental Figure 8.5: S2K146 epitope conservation across sarbecoviruses.

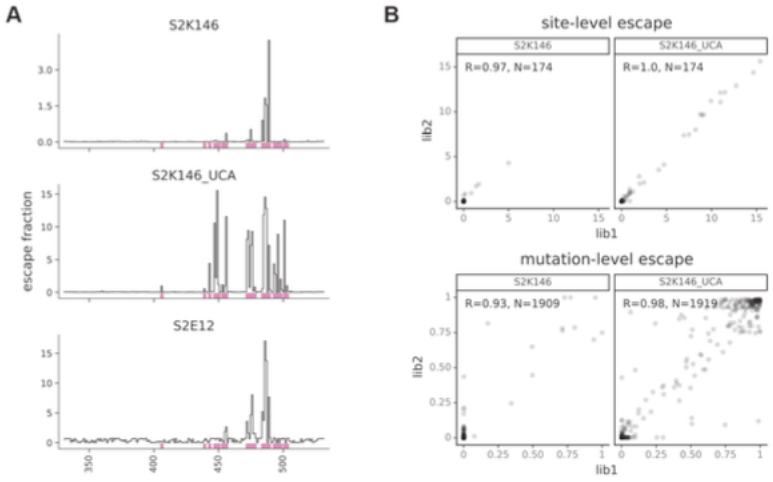
A) Protein sequence alignment of RBDs representative of sarbecovirus clade 1a (green), clade 1b (blue), clade 2 (orange) and clade 3 (pink) along with the reconstructed ancestor of all Asian sarbecoviruses (AncAsia) (*Starr, 2022*). Residue numbering is based on SARS-CoV-2 sequence with identical residues indicated as dots. Residues buried upon S2K146 binding are highlighted in boxes and epitope residues conserved between the SARS- CoV-2 and SARS-CoV RBDs are denoted in red.

B) Binding of S2K146 to the AncAsia RBD evaluated by ELISA.



Supplemental Figure 8.6: SARS-CoV-2 broadly neutralizing sarbecovirus mAbs.

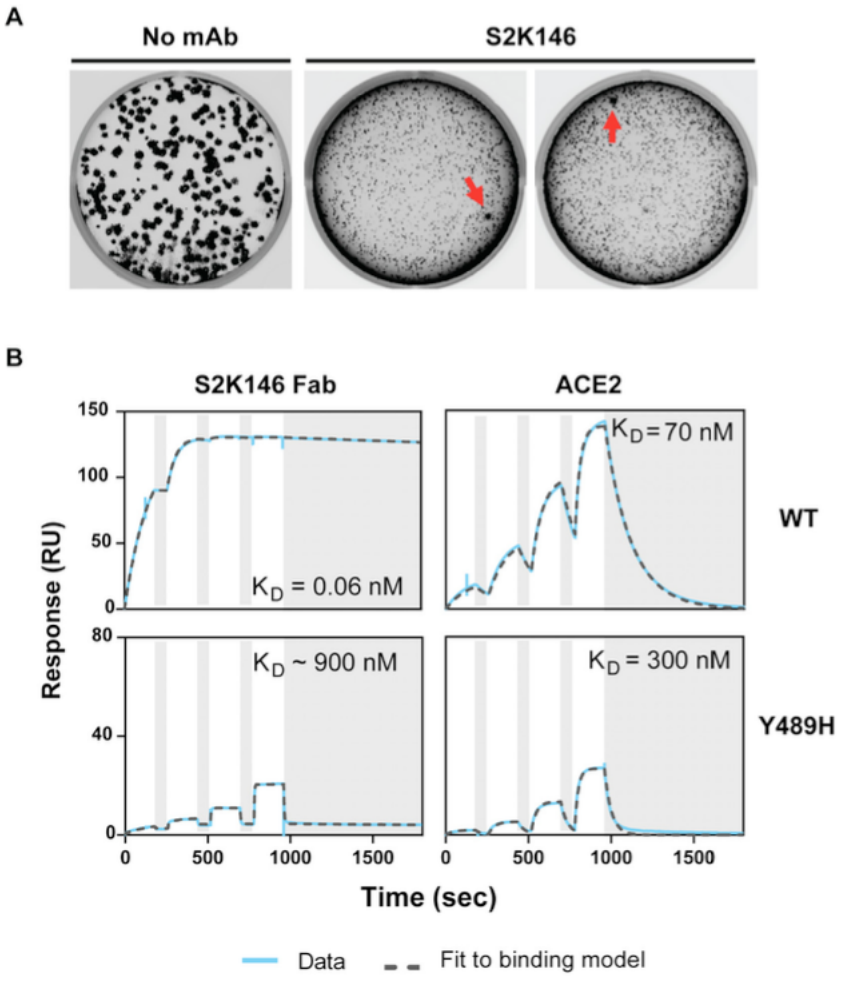
Structural superimposition of S2K146 (purple), site I-targeting S2E12 (grey) (*Tortorici, 2020*), site II-targeting S2X259 (red) (*Tortorici, 2021*), and site IV-targeting S309 (blue) (*Pinto, 2020*) Fabs bound to the SARS-CoV-2 RBD (light blue). S2K146 is depicted as a semi-transparent surface to show that S2E12 and S2K146 bind largely overlapping binding sites on the RBD. Although the constant domains of S2K146 and S2X259 slightly overlap, the flexibility around the Fab elbow allows binding of both mAbs as shown in Fig 1A.



Supplemental Figure 8.7: DMS analysis of the yeast-displayed SARS-CoV-2 RBD.

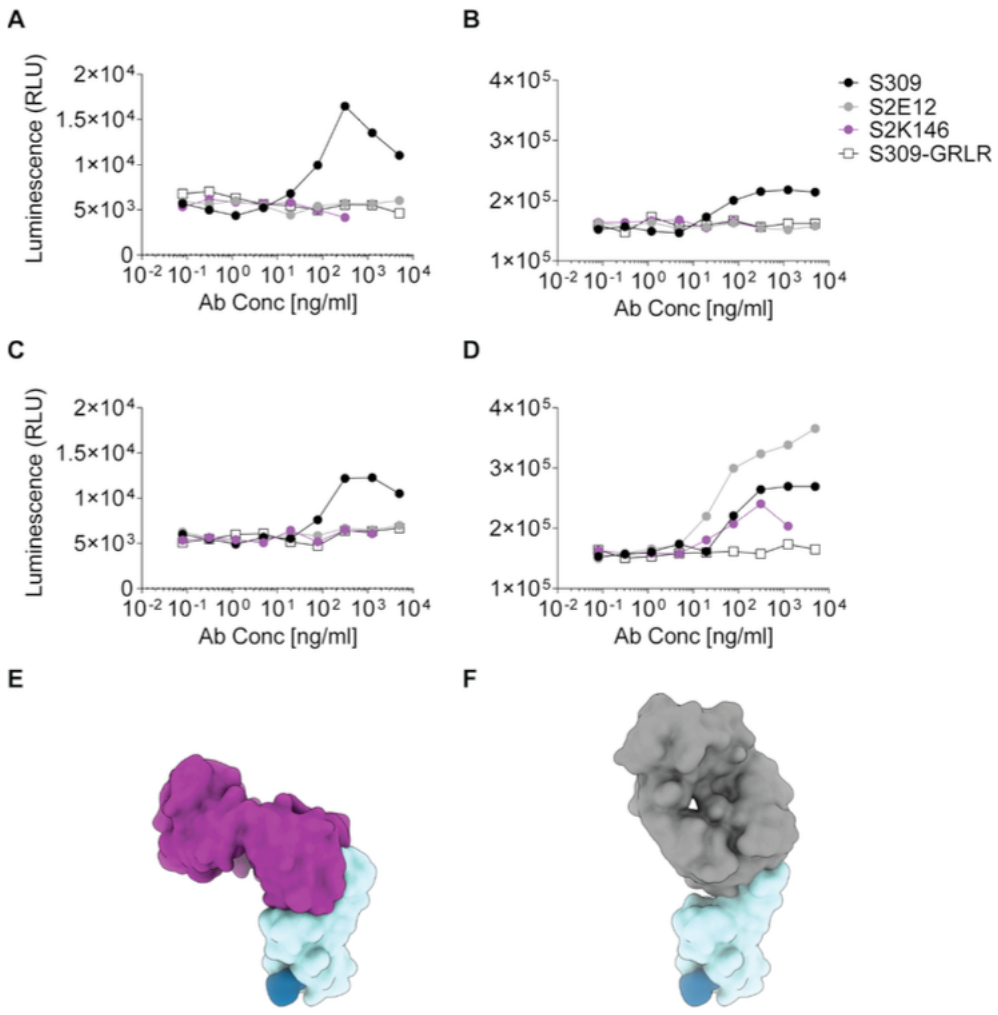
A) Mapping of escape mutations reducing S2K146, S2K146 UCA or SE12 binding using yeast-displayed RBD DMS. Line plots show escape at each site in the RBD (summed

effects of all mutations at each site). Sites of strong escape are indicated with purple lines. **B**) Correlation in per-mutation (down) and sum-per-site (up) escape fraction for replicate DMS library experiments.



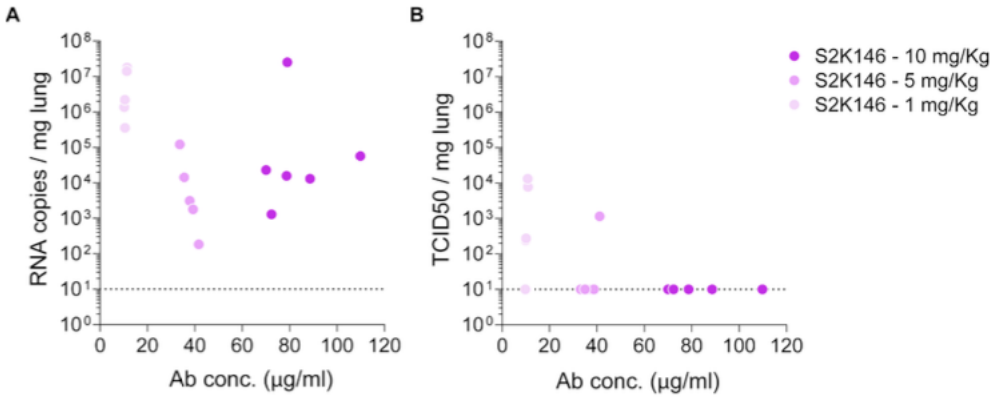
Supplemental Figure 8.8: S2K146 escape clone selection by plaque assays using VSV-SARS-2.

A) Plaque assays performed using VSV-SARS-CoV-2 Wuhan-Hu-1 D614G S chimeric virus on Vero cells with or without S2K146 in the overlay to isolate escape mutants (red arrow). Data are representative of two sets of experiments. **B)** Binding of the S2K146 Fab or monomeric human recombinant ACE2 ectodomain to the SARS-CoV-2 Wuhan-Hu-1 or Y489H RBD immobilized at the surface of SPR chips. The y axis is scaled based on RBD surface density. Experiments were performed with a 3-fold dilution series of Fab or monomeric hACE2: 300, 100, 33, 11 nM and were run as single-cycle kinetics.



Supplemental Figure 8.9: Activation of FcγγRIIa and FcγγRIIIa in vitro

A-B) NFAT-driven luciferase signal induced in Jurkat cells stably expressing Fc**γγ**RIIa H131 variant (A) or Fc**γγ**RIIIa V148 variant (B) upon S2K146 mAb binding to full-length wildtype SARS-CoV-2 S on CHO target cells. **C-D)** NFAT-driven luciferase signal induced in Jurkat cells stably expressing Fc**γγ**RIIa H131 variant (C) or Fc**γγ**RIIIa V148 variant (D) upon S2K146 mAb binding to uncleavable full-length wildtype SARS-CoV-2 S on CHO target cells. **E)** Surface rendering of site-I-targeting S2K146 (purple) mAb bound to the SARS-CoV-2 RBD (light blue). **F)** Surface rendering of site-I-targeting S2E12 (grey) mAb (*Tortorici, 2020*) bound to the SARS-CoV-2 RBD (light blue). The N343 glycan is rendered as a blue surface in E and F.



Supplemental Figure 8.10: Correlation between mAb concentration and viral loads in the lungs of

(A-B) Viral RNA copy number (A) and replicating virus titers (B)

of SARS-CoV-2 Beta VOC in the lung of Syrian hamsters at 4 days post-infection plotted as a function of S2K146 mAb serum concentrations at day 1 post infection. (S2K146 10mg/kg n = 6; S2K146 5 mg/kg n = 5; S2K146 1 mg/kg n = 5) was administered 1 day post infection.

Table S 8.1: CryoEM data collection and refinement statistics.

	SARS-CoV-2 S/S2K146 (3RBDs open) EMD-25785	SARS-CoV-2 S/S2K146 (2RBDs open) PDB 7TAT EMD-25784	SARS-CoV-2 S RBD/S2K146 (Local refinement) PDB 7TAS EMD-25783
Data collection and processing			
Magnification	105,000	105,000	105,000
Voltage (kV)	300	300	300
Electron exposure (e^{-}/Å^2)	60	60	60
Defocus range (μm)	0.5-2.5	0.5-2.5	0.5-2.5
Pixel size (Å)	0.843	0.843	0.843
Symmetry imposed	C1	C1	C1
Final particle images (no.)	212,054	348,047	97,818
Map resolution (Å)	3.2	3.2	3.2
FSC threshold	0.143	0.143	0.143
Map sharpening *B* factor (Å^2)	-96	-99	-52
Validation			
MolProbity score		0.88	1.58
Clashscore		0.53	3.47
Poor rotamers (%)		0.18	1.69
Ramachandran plot			
Favored (%)		96.78	96.09
Allowed (%)		2.82	3.91
Disallowed (%)		0.4	0
EMRinger score		4.94	4.31

Table S 8.2: Summary of nucleotide and amino acid mutations found in 36 neutralization-resistant VSV-SARS-CoV-2-S D614G chimera plaques.

Clones	Nucleotide mutant	Amino acid Mutant
#1	T1465C/ T1602G	Y489H/ V534V
#2	T1465C	Y489H
#3	T1465C	Y489H
#4	T1465C	Y489H
#5	T1465C	Y489H
#6	T1465C	Y489H
#7	T1465C	Y489H
#8	T1465C	Y489H
#9	T1465C	Y489H
#10	T1465C	Y489H
#11	T1465C	Y489H
#12	T1465C	Y489H
#13	T1465C	Y489H
#14	T1465C	Y489H
#15	T1465C	Y489H
#16	T1465C	Y489H
#17	T1465C	Y489H
#18	T1465C	Y489H
#19	T1465C	Y489H
#20	T1465C	Y489H
#21	T1465C	Y489H
#22	T1465C	Y489H
#23	T1465C	Y489H
#24	T1465C	Y489H
#25	T1465C	Y489H
#26	T1465C	Y489H
#27	T1465C	Y489H
#28	T1464C/ T1465C	C488C/ Y489H
#29	T1465C	Y489H
#30	T1263C/ T1465C	Y421Y/ Y489H
#31	T1465C/ A1671G	Y489H/ K557K
#32	T1465C	Y489H
#33	T1465C	Y489H
#34	T1465C	Y489H
#35	T1465C	Y489H
#36	T1465C	Y489H

CHAPTER 9. ELICITATION OF BROADLY PROTECTIVE SARBECOVIRUS IMMUNITY BY RECEPTOR-BINDING DOMAIN NANOPARTICLE VACCINES.

This next story took place early in the COVID19 pandemic when supplying the global populace with vaccines was made challenging by the enormous demand and the breakdown of supply chains; as a result designing vaccines with high effectiveness and minimal dosing requirements was goal to ensure maximum distribution of vaccines. In this next story, I describe our efforts to design a vaccine utilizing sarbecoviruses RBDs displayed with high valency on the surface of a nanoparticle. I detail our findings that two ultra-low dose immunizations or a single immunization with RBD-NP produces potent neutralizing Ab responses in mice; conferring protection against SARS-CoV-2 MA10 challenge, suggesting that the nanoparticle platform could enable dose-sparing regimens to achieve global vaccination.

Adapted from: Walls, A. C., et al. (2021). "Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines." Cell **184**(21): 5432-5447.e5416.

9.1 Chapter Introduction

The emergence of SARS-CoV-2 in late 2019 resulted in the COVID-19 pandemic that brought the world to a standstill. Moreover, the recurrent spillovers of coronaviruses in humans along with detection of SARS-CoV-2-, SARS-CoV-, and MERS-CoV-related coronaviruses in bats suggest that future zoonotic transmission events may continue to occur (Menachery et al., 2015, 2016; Zhou et al., 2020a). SARS-CoV-2 infects host cells through the attachment of the viral transmembrane spike (S) glycoprotein to angiotensin-converting enzyme 2 (ACE2), followed by fusion of the viral and host membranes (Hoffmann et al., 2020; Lan et al., 2020; Letko et al., 2020; Shang et al., 2020; Walls et al., 2020a; Wang et al., 2020; Wrapp et al., 2020; Yan et al., 2020; Zhou et al., 2020b). The SARS-CoV-2 S protein is the primary target of neutralizing antibodies (Abs), and the immunodominant receptor-binding domain (RBD) accounts for more than 90% of the neutralizing activity in COVID-19 convalescent sera and vaccinated individuals (Greaney et al., 2021a, 2021b; Piccoli et al., 2020). Numerous monoclonal Abs (mAbs) recognizing distinct antigenic sites on the RBD were isolated and shown to neutralize viral entry and protect small animals and non-human primates (NHPs) from SARS-CoV-2 challenge (Barnes et al., 2020; Baum et al., 2020a; Brouwer et al., 2020; Piccoli et al., 2020; Pinto et al., 2020; Rogers et al., 2020; Starr et al., 2021; Tortorici et al., 2020; Wec et al., 2020; Zost et al., 2020). As a result, SARS-CoV-2 S is the focus of nucleic acid, vectored, and protein subunit vaccines currently being developed and deployed (Corbett et al., 2020a, 2020b; Jackson et al., 2020; Mercado et al., 2020; Polack et al., 2020; Tostanoski et al., 2020; Yu et al., 2020).

Worldwide sequencing of SARS-CoV-2 clinical isolates has led to the identification of numerous mutations in the >1,500,000 genome sequences available to date (https://www.gisaid.org/). The SARS-CoV-2 S D614G mutation has become globally dominant and is associated with enhanced viral transmission and replication but does not significantly affect Ab-mediated neutralization (Hou et al., 2020; Korber et al., 2020; Plante et al., 2020; Yurkovetskiy et al., 2020). Conversely, some mutations found in circulating SARS-CoV-2 isolates were shown to promote escape from mAbs and to reduce neutralization by immune sera (Baum et al., 2020b; Collier et al., 2021; Li et al., 2020; McCallum et al., 2021a, 2021b; Wang et al., 2021a; Weisblum et al., 2020; Wibmer et al., 2021). As a result, formulation of mAb cocktails or the use of mAbs targeting conserved epitopes and neutralizing a broader spectrum of circulating SARS-CoV-2 variants emerged as a promising strategy to overcome this issue (Baum et al., 2020b; Cathcart et al., 2021; Dong et al., 2021; Greaney et al., 2020; Jette et al., 2021; Martinez et al., 2021a; Pinto et al., 2020; Tortorici et al., 2020, 2021). The recent emergence of several variants and variants of concern (VOCs) with numerous S mutations is especially worrisome, including B.1.1.7 (alpha), B.1.351 (beta), B.1.427/B.1.429 (epsilon), P.1 (gamma), and B.1.617.2 (delta) that originated in the UK, South Africa, the USA, Brazil, and India, respectively (Davies et al., 2020; Deng et al., 2021; Faria et al., 2021; McCallum et al., 2021b; Tegally et al., 2020). Some of these mutations lead to significant reductions in the neutralization potency of N-terminal domain (NTD)- and RBD-specific mAbs, convalescent sera, and Pfizer/BioNTech BNT162b2- or Moderna mRNA-1273-elicited sera (Collier et al., 2021; McCallum et al., 2021a, 2021b; Wang et al., 2021a).

We recently described a multivalent subunit vaccine displaying the SARS-CoV-2 RBD nanoparticle (RBD-NP) in a highly immunogenic array using a computationally designed self-assembling protein nanoparticle (Bale et al., 2016; Walls et al., 2020b). Vaccination with RBD-NP resulted in 10-fold higher neutralizing Ab titers in mice than the prefusion-stabilized S2P trimer (which is used in most current vaccines) despite a 5-fold lower dose and protected mice against mouse-adapted SARS-CoV-2 (SARS-CoV-2-MA) challenge (Dinnon et al., 2020; Walls et al., 2020b). We demonstrated that the RBD-NP vaccine elicited robust neutralizing Ab and CD4 T cell responses in NHPs when formulated with several clinic-ready adjuvants and conferred protection against SARS-CoV-2 infection in the nose, pharynges, and bronchioles (Arunachalam et al., 2021). The RBD-NP vaccine is currently being evaluated in two phase I/II clinical trials (ClinicalTrials.gov: NCT04742738 and NCT04750343) and recently received funding from the coalition for epidemic preparedness innovations for phase III clinical trials.

Although the S fusion machinery (S_2 subunit) has higher sequence conservation than the RBD (Pinto et al., 2021; Sauer et al., 2021; Tortorici and Veesler, 2019; Walls et al., 2016, 2020a), the breadth of neutralization and protection provided by RBD-based vaccines remains unknown. The isolation of RBD-specific cross-reactive mAbs neutralizing SARS-CoV-2 and SARS-CoV suggests that RBD-based vaccines could, in principle, elicit Abs that neutralize distantly related sarbecoviruses, which have future pandemic potential (Cathcart et al., 2021; Jette et al., 2021; Martinez et al., 2021a; Pinto et al., 2020; Rappazzo et al., 2021; Starr et al., 2021; Tortorici et al., 2021; Wec et al., 2020). RBD-based vaccines are also unaffected by S mutations outside of the RBD, especially in the highly variable NTD that is heavily mutated in many VOCs (Andreano

et al., 2020; Avanzato et al., 2020; Choi et al., 2020; Collier et al., 2021; Wang et al., 2021a; McCallum et al., 2021a, 2021b; McCarthy et al., 2021). Here, we explored dose-sparing strategies for the RBD-NP vaccine and evaluated the impact of genetic diversity among SARS-CoV-2 clinical isolates and sarbecoviruses on vaccine-elicited Abs. We further designed mosaic and cocktail sarbecovirus RBD-NPs that elicit broad and protective Ab responses against heterologous sarbecovirus challenge, which could represent the next generation of pan-sarbecovirus vaccines.

9.2 Dose-sparing RBD-NP vaccination protects mice against SARS-CoV-2 challenge

Considering the unprecedented need for rapid global distribution of SARS-CoV-2 vaccines, we investigated the ability of RBD-NP to induce neutralizing and protective Ab titers at lower doses than previously evaluated. Specifically, we set out to test whether vaccine-elicited neutralizing Ab titers are altered by lowering the immunogen dose 10-fold (two 0.1 μg immunizations, RBD antigen dose) or reducing the number of doses (a single 1 μg immunization, RBD antigen dose) compared to our initial work (Walls et al., 2020b). To test this, BALB/c cByJ mice were immunized intramuscularly at week 0 with either 0.1 μg (ten mice) or 1 μg (two groups of eight mice or sixteen mice total) of AddaVax-adjuvanted RBD-NP. Two weeks post-prime, serum binding and neutralizing Ab titers were roughly comparable for the two vaccine doses (neutralization geometric mean titer [GMT] ~2–4 × 10^2) measured using a mouse leukemia virus (MLV) pseudotyping system (Millet and Whittaker, 2016; Walls et al., 2020a) (**Figs. 9.1A and 9.1B; Data S9.1**). Three weeks post prime, we boosted the mice that were immunized with 0.1 μg of the vaccine and eight of the sixteen mice immunized with 1 μg of the

vaccine with the same respective dose. Two weeks post-boost, serum binding and neutralizing Ab titers were comparable for the 0.1 μg and 1 μg groups (neutralization GMT ~2 × 10^4) (**Figs. 9.1C and 9.1D; Data S9.1**), suggesting that a 10-fold reduction in the RBD-NP dose does not affect serum neutralizing responses. Furthermore, we observed that the magnitude of binding and neutralizing Ab titers increased over time for the mice that received a single immunization (neutralization GMT 2 × 10^3). These results suggest that lowering the vaccine dose or reducing the number of immunizations represent possible dose-sparing strategies that did not compromise elicitation of high levels of neutralizing Ab titers using RBD-NPs.

To further evaluate vaccine efficacy, eight mice that received one dose and eight mice that received two doses of the 1-μg RBD-NP vaccine along with six unvaccinated mice were challenged ten weeks post-prime with 10^5 plaque-forming units (PFUs) of mouse-adapted SARS-CoV-2 MA10 (Leist et al., 2020). The SARS-CoV-2 MA10 model causes weight loss and lung pathology when 10^5PFUs (the highest titer tested [Leist et al., 2020]) are used in young BALB/c mice. The mice were followed for four days to assess protection from disease. All mice in the RBD-NP-vaccinated groups were protected from weight loss throughout the duration of the experiment regardless of the number of doses, whereas control mice lost ~10% of their weight by day four (**Fig. 9.1E**). Analysis of lung pathology and viral titers in lung tissues indicated that the vaccinated mice were not affected by SARS-CoV-2 MA10 challenge, whereas the control mice showed lung discoloration four days post-infection and a high viral load (~1 × 10^4 PFUs) (**Figs. 9.1F and 9.1G**). These results indicate that one or two immunizations with 1 μg RBD-NPs results in protection against SARS-CoV-2 MA10-

induced disease. Furthermore, mice vaccinated twice with a 10-fold lower dose of RBD-antigen (0.1 μg) had higher serum neutralizing Ab titers than mice receiving a single 1-μg dose, suggesting that they would likely also be protected since neutralizing Ab titers are a correlate of protection against SARS-CoV-2 (Arunachalam et al., 2021; McMahan et al., 2021). These data suggest that multiple RBD-NP dose-sparing regimens may be possible.

9.3 RBD-NP vaccination induces Abs targeting diverse antigenic sites in NHPs

To characterize the epitopes targeted by RBD-NP-elicited polyclonal Abs, we used deep mutational scanning (DMS) with sera obtained 56 days post-prime from one of two pigtail macaques immunized at days 0 and 28 with a dose of AddaVax-adjuvanted RBD-NP containing 88 μg of the SARS-CoV-2 RBD (with a linker of twelve glycine/serine (GS) residues between the RBD and the NP) (Walls et al., 2020b) (**Figs. 9.2A and 9.2B**). These experiments rely on yeast surface display of RBD libraries covering nearly all possible amino acid mutations coupled with fluorescence-activated cell sorting (FACS) to identify RBD mutants with attenuated Ab binding compared to the wild-type (Wuhan-Hu-1) SARS-CoV-2 RBD (Greaney et al., 2020; Starr et al., 2020a). No single mutation had more than a marginal effect on serum Ab (IgG/IgM/IgA) binding, indicating broad targeting of distinct RBD epitopes, whereas several previously described COVID-19 human convalescent plasmas (HCPs) analyzed by DMS displayed greater sensitivity to individual mutations and are shown for comparison (Greaney et al., 2021b) (**Figs. 9.2B and S1; Table S9.1**). These results show that RBD-NP vaccination elicits highly diverse polyclonal Ab responses that target multiple distinct antigenic sites

and are more resilient to escape mutations reducing polyclonal Ab binding than HCP (Barnes et al., 2020; Greaney et al., 2021b; Piccoli et al., 2020; Starr et al., 2021).

To measure the magnitude of vaccine-elicited Abs against distinct RBD antigenic sites, we used quantitative competition ELISAs with human ACE2-Fc (hACE2-Fc), which binds to the receptor-binding motif (RBM) corresponding to antigenic sites Ia and Ib, as well as with structurally characterized mAbs recognizing antigenic sites II (CR3022) and IV (S309) (Huo et al., 2020; Piccoli et al., 2020; Pinto et al., 2020; Yuan et al., 2020) (**Figs. 2C–2E; Data S1**). These experiments used sera from the aforementioned pigtail macaques (RBD GMT EC_{50} 5×10^4) followed for 168 days and sera from 5 rhesus macaques immunized twice on days 0 and 21 with AS03-adjuvanted RBD-NPs containing 25 μg of the SARS-CoV-2 RBD (with a 16-GS linker between the RBD and the NP; RBD GMT EC_{50} 1×10^4) and followed for 98 days (Arunachalam et al., 2021; Walls et al., 2020b; **Figs. 9.2A and S9.2A**). These two sets of RBD-NP sera were compared to sera from 5 rhesus macaques immunized twice on days 0 and 21 with 12.5 μg of AS03-adjuvanted prefusion-stabilized HexaPro S trimer (RBD EC_{50} 1×10^4) and followed for 42 days (Arunachalam et al., 2021; Hsieh et al., 2020), as well as a panel of COVID-19 HCP (**Table S9.2**) (RBD EC_{50} 2×10^4) (**Figs. 9.2A and S9.2A**). All NHP sera post boost had high titers of Abs targeting sites Ia and Ib in the immunodominant RBM (day 42 hACE2 competition RBD-NP GMT 3×10^3 and HexaPro GMT 7×10^2) (**Fig. 9.2C; Data S9.1**), a correlate of neutralization potency (Piccoli et al., 2020), in agreement with the potent immunogenicity and protective efficacy of RBD-NP and HexaPro S in NHPs (Arunachalam et al., 2021). We also observed strong Ab responses against antigenic sites II (RBD-NP day 42 GMT 4 ×

10^2 and HexaPro GMT 4 × 10^1) with CR3022 competition and IV (RBD-NP day 42 GMT 2.5 × 10^3 and HexaPro GMT 6.5 × 10^1) (**Figs. 9.2D and 9.2E; Data S9.1**), which comprise conserved sarbecovirus epitopes recognized by broadly neutralizing sarbecovirus mAbs such as S2X259 (site II) and S309 (site IV) (Piccoli et al., 2020; Pinto et al., 2020; Tortorici et al., 2021). Ab responses were durable against all 3 antigenic sites regardless of dose or adjuvant, with a decrease of competition titers by roughly half an order to one order of magnitude 168 days (pigtail macaque) or 98 days (rhesus macaque) post-prime. RBD-NP elicited much higher peak binding titers toward all RBD antigenic sites evaluated compared to HexaPro S and HCP, showcasing the potency, durability, and multi-specificity of Ab responses induced by multivalent display of the RBD, possibly as a result of enhanced accessibility of the antigen.

9.4 RBD-NP vaccine elicits potent neutralizing Ab responses in NHPs against a panel of SARS-CoV-2 S variants

To assess the neutralization breadth of RBD-NP- and HexaPro S-elicited Abs in NHPs at peak titer (42 days post-prime), we evaluated serum neutralizing activity against a panel of pseudotyped viruses comprising wild-type (D614G) SARS-CoV-2 S and twelve single-residue SARS-CoV-2 RBD mutants detected in clinical isolates (N439K, L452R, Y453F, L455F, S477N, E484A/K/Q, F486L, S494P, N501Y, and G504D) as well as the B.1.1.7 (H69-V70 deletion, Y144 deletion, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H); B.1.1.7 + E484K (H69-V70 deletion, Y144 deletion, E484K, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H); B.1.351 (L18F, D80A, D215G, L242-L244 deletion, R246I, K417N, E484K, N501Y, D614G, A701V); and P.1 (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G,

H655Y, T1027I). Several single-residue RBD mutations did not affect neutralizing titers compared to wild-type D614G SARS-CoV-2 S, for which GMTs of 1.8×10^3 and 2.4×10^3 were determined with RBD-NP- and HexaPro S-elicited sera, respectively, using a vesicular stomatitis virus (VSV) pseudotyped virus (**Fig. 9.3A; Data S9.1**). For example, the N439K- (GMT 2.5×10^3 for RBD-NP and 2.7×10^3 for HexaPro S) and mink-associated Y453F (GMT 4×10^3 for RBD-NP and 5×10^3 for HexaPro S) mutation did not dampen the neutralization potency of either RBD-NP- or HexaPro S-elicited sera (**Fig. 9.3A; Data S9.1**). However, the N501Y substitution (present in the B.1.1.7, P.1, and B.1.351 lineages) reduced the neutralization potency of RBD-NP- (GMT 5×10^2) and HexaPro S (GMT 8×10^2)-elicited sera 3-to-4-fold and the L452R mutation (present in the B.1.427/B.1.429, B.1.617.1, and B.1.617.22 variant) led to a 3.4-to-4.4-fold drop in neutralization potency (GMT 4×10^2 for RBD-NP and 7×10^2 for HexaPro S). These substitutions have been associated with the loss of binding and neutralization for vaccine-elicited Abs and mAbs (Collier et al., 2021; McCallum et al., 2021b; Wang et al., 2021b; Thomson et al., 2021; Wang et al., 2021c). The E484A mutation reduced serum neutralization activity up to 2-fold (GMT $1.1x10^3$ for RBD-NP and $1.1x10^3$ for HexaPro S) whereas E484K, present in B.1.351 and P1 (GMT 2×10^2 for RBD-NP and 2×10^2 for HexaPro S), and E484Q, found in B.1.617.1 (GMT 5×10^2 for RBD-NP and 3.6×10^2 for HexaPro S), decreased it 9-fold (E484K) and 4-fold (E484Q) for RBD-NP and 12-fold (484K) and 7-fold (484Q) for HexaPro S compared to D614G SARS-CoV-2 S. These experiments suggest that a substantial fraction of the neutralizing activity elicited by both RBD-NP and HexaPro S is focused on the RBM, especially near position 484,

as further supported by the ~5-fold decrease in neutralization resulting from the F486L (GMT 3.7 × $10^2$for RBD-NP and 5 × 10^2 for HexaPro S) substitution.

To understand the impact of the full constellation of mutations present in the S proteins of the aforementioned VOCs, we evaluated RBD-NP- and HexaPro S-elicited serum neutralizing activity against corresponding HIV and VSV pseudotyped variants (D614G GMT RBD-NP HIV 5 × 10^3 and VSV RBD-NP 6 × 10^2, respectively, and D614G GMT HexaPro S HIV 3.5 × 10^3 and VSV 1.2 × 10^3) (**Figs. 9.3B and 9.3C**). Although we did not observe major reductions in neutralization titers (up to to 2-fold) toward the B.1.1.7 VOC (GMT RBD-NP HIV 2.5 × 10^3, VSV 4.7 × 10^2 and HexaPro S HIV 1.8 × 10^3, VSV 8.8 × 10^2) (**Figs. 9.3B and 9.3C**), addition of the E484K mutation to the B.1.1.7 background reduced neutralization 4-to-6-fold for RBD-NP- (GMT HIV 8.8 × 10^2) and HexaPro S-elicited sera (GMT HIV 8.6 × 10^2), confirming the importance of the 484 position (**Fig. 9.3B**). Neutralization of B.1.351 was 10-fold lower using HIV pseudovirus (GMT RBD-NP 5 × 10^2 and HexaPro S 3.6x10^2) (**Fig. 9.3B**) and ~3-fold lower using VSV pseudovirus (GMT RBD-NP 2 × 10^2 and HexaPro S 4x10^2) (**Fig. 9.3C**) for both RBD-NP- and HexaPro-elicited sera, in agreement with authentic virus neutralization data (Arunachalam et al., 2021). Interestingly, RBD-specific binding Ab titers did not differ significantly between the B.1.351 RBD (GMT RBD-NP 6.5 × 10^3 and HexaPro S 3 × 10^3) and Wuhan-Hu-1 RBD (GMT RBD-NP 1.4 × 10^4 and HexaPro S 1.7 × 10^4) (**Fig. S9.2A**), indicating that neutralizing activity is accounted for by a portion of total binding Abs. Reductions in neutralizing Ab titers were also observed against P.1 S VSV pseudoviruses, with a 6-fold drop for RBD-NP (GMT 1 × 10^2) and 8-fold dampening for HexaPro S (GMT 1.5 × 10^2) (**Fig. 9.3C**). Furthermore, we observed a 10-fold reduction

in plasma neutralizing potency for individuals who received two doses of the Pfizer-BioNTech BNT162b2 mRNA vaccine against B.1.351 S pseudotyped virus (GMT 6 × 10^1) relative to the wild-type (D614G) SARS-CoV-2 S (GMT 6 × 10^2) pseudotyped virus (**Fig. 9.3B; Data S9.1**). These findings show that, as is the case for many convalescent individuals (Greaney et al., 2021b; Liu et al., 2021; Piccoli et al., 2020), an important fraction of vaccine-elicited neutralizing Abs in NHPs and humans is focused on the RBM (specifically around position 484), independently of the immunogen (RBD-NP, HexaPro S, or 2P-stabilized S [Pallesen et al., 2017]) or the vaccine modality (protein subunit or mRNA). However, we note that neutralization assays may underestimate the contribution of NTD-specific or non-RBM RBD-targeted neutralizing Abs (Lempp et al., 2021; McCallum et al., 2021a; Suryadevara et al., 2021), and protection from challenge will be the ultimate readout.

To further investigate the relationships between neutralizing Ab titers and emerging SARS-CoV-2 variants, we immunized the two pigtail macaques a third time with RBD-NP, formulated with an oil-in-water (O/W) emulsion-based adjuvant, 168 days after the primary immunization (**Fig. 9.2A**). This boost induced potent serum neutralizing activity against wild-type (D614G) SARS-CoV-2 (GMT 2 × 10^4) (**Fig. 9.3D**) as well as the B.1.1.7 (GMT 2 × 10^4) and B.1.351 (GMT 8 × 10^3) VOC pseudoviruses, suggesting that an overall increase in neutralizing Ab titers may be a suitable strategy to cope with emerging variants. These results are consistent with recent studies demonstrating that boosting COVID-19 convalescent individuals with a single mRNA vaccination elicited high (neutralizing) Ab titers, including against the B.1.351 variant (Abu Jabal et al., 2021; Krammer et al., 2021; Stamatatos et al., 2021), and suggest that a third

vaccination of naive individuals could be a suitable strategy to limit the impact of emerging variants. This strategy is currently being evaluated for immunocompromised patients in several countries.

9.5 RBD-NP vaccine elicits cross-reactive sarbecovirus polyclonal Abs in NHPs

As RBD-NP- and HexaPro S-elicited polyclonal Abs exhibited similar resilience to a range of mutations, we next investigated cross-reactivity with a panel of sarbecovirus RBDs. RBD-NP-elicited polyclonal Abs purified from pigtail macaque serum obtained 70 days post-prime strongly cross-reacted with the SARS-CoV-2-related Pangolin-GD and RaTG13 RBDs and bound more weakly to distantly related RmYN02, SARS-CoV, WIV16, and ZXC21 RBDs (**Figs. S9.2B and S9.2C**). Measurement of Ab binding titers using the SARS-CoV and SARS-CoV-2 S2P ectodomain trimers by ELISA showed that RBD-NPs and HexaPro S induced similar levels of cross-reactive Abs against each antigen, with responses roughly two orders of magnitude higher against SARS-CoV-2 S2P (detection antigen matching immunogen) compared to SARS-CoV S2P (**Fig. S9.2D; Data S9.1**). The two immunogens also elicited similar peak levels of SARS-CoV S ACE2-competing Abs (GMT 2×10^2 RBD-NP and 5×10^1 HexaPro) suggesting that cross-neutralizing Abs might have been induced to similar extents between the two vaccines (**Fig. S9.2E**).

Motivated by the cross-reactivity of RBD-NP-elicited polyclonal Abs with various sarbecovirus RBDs and the correlation between ACE2 competition and serum neutralization titers (Piccoli et al., 2020), we evaluated neutralization of a panel of HIV (D614G GMT 5.5×10^2) and VSV (D614G GMT 6×10^2) pseudoviruses harboring sarbecovirus S glycoproteins. RBD-NP-elicited sera efficiently neutralized pseudotyped

viruses harboring the S glycoprotein of the Pangolin-GD isolate (GMT 6.6 × 10^2 and 7.7 × 10^2 VSV) and RaTG13 (GMT VSV 6 × 10^2) (**Figs. 9.3E and 9.3F**) (Lam et al., 2020; Zhou et al., 2020b), in agreement with the close phylogenetic relationship of their RBDs with that of SARS-CoV-2 S (**Fig. S9.2B**). HexaPro S-elicited NHP sera also efficiently neutralized Pangolin-GD S (GMT HIV 4.7 × 10^3 and VSV 9 × 10^2) and RaTG13 S pseudotypes (GMT VSV 4.7 × 10^3) (**Figs. 9.3E and 9.3F**). Furthermore, we observed that both RBD-NP- and HexaPro S-induced polyclonal Abs that weakly neutralized HIV pseudovirus carrying SARS-CoV S (RBD-NP GMT 6 × 10^1 and HexaPro GMT 1 × 10^2) (**Figs. 9.3E and 9.3F**). Collectively, these data demonstrate that both immunogens elicited comparable neutralization breadth and potency against the pseudoviruses tested, reinforcing the notion that most sarbecovirus S-directed neutralizing Abs target the RBD (Greaney et al., 2021a, 2021b; Piccoli et al., 2020).

9.6 Design, assembly, and characterization of mosaic and cocktail sarbecovirus RBD-NPs

We and others have recently evaluated nanoparticle immunogens that display multiple antigenic variants of viral glycoproteins as a potential route toward broadly protective vaccines (Boyoglu-Barnum et al., 2021; Cohen et al., 2021a, 2021b; Kanekiyo et al., 2019). Given the large number of coronaviruses circulating in zoonotic reservoirs, such vaccines could be important for understanding vaccine-induced neutralization/protection breadth and preventing future pandemics (Menachery et al., 2015, 2016). We expressed and purified four RBDs from the S proteins of SARS-CoV-2, SARS-CoV, and the bat coronaviruses WIV1 and RaTG13 genetically fused to the I53-50A trimer. An equimolar mixture of these four proteins was

added to the I53-50B pentamer to assemble a mosaic RBD-NP (mRBD-NP) co-displaying the four RBDs on the same nanoparticle (**Fig. 9.4A and S9.3A**). We also assembled a trivalent mosaic RBD-NP designated “drop out” lacking the SARS-CoV RBD (mRBD-NP-DO), as well as cocktail immunogens with three (cRBD-NP-DO; lacking the SARS-CoV RBD) or four (cRBD-NP) independently assembled nanoparticles, each displaying a single type of the aforementioned RBDs, mixed after independent assembly. Finally, we made a bivalent mosaic RBD-NP co-displaying the SARS-CoV and SARS-CoV-2 RBDs (**Fig. S9.3A**) to directly confirm co-display on the nanoparticles using a sandwich binding assay. All of the nanoparticle immunogens formed the intended icosahedral architecture and retained native antigenicity, as shown by SDS-PAGE, dynamic light scattering, negative staining electron microscopy, and binding to hACE2-Fc (**Figs. S9.3B–S9.3E**). We found that the vaccine candidates were stable for at least 4 weeks at several temperatures except the highest temperature evaluated (37°C), at which we observed a decrease in hACE2 recognition over time beginning after 7 to 14 days, presumably due to aggregation (**Figs. 9.4B–9.4E**). Following immobilization using the SARS-CoV-2-specific mAb S2H14 (Piccoli et al., 2020), the quadrivalent, trivalent, and bivalent mRBD-NPs all bound the Fab of the SARS-CoV-specific Ab S230 (Piccoli et al., 2020; Rockx et al., 2008; Walls et al., 2019), confirming co-display, whereas the monovalent SARS-CoV-2 RBD-NP did not (**Fig. S9.3G**). We determined that the reactivity of the trivalent mRBD-NP derived from the inclusion of the WIV1 RBD in this vaccine, as the monovalent WIV1 RBD-NP also bound the S230 Fab after immobilization with hACE2 (as expected [Menachery et al., 2016]), whereas the monovalent SARS-CoV-2 and RaTG13 RBD-NPs did not

(**Fig. S9.3H**). Collectively, these data demonstrate that all of the nanoparticles were stable and (co-)displayed the various RBDs as intended.

9.7 Cocktail and mosaic RBD-NP vaccines elicit cross-reactive and broadly neutralizing sarbecovirus Abs

The mosaic and cocktail nanoparticle immunogens were compared to monovalent RBD-NP vaccines and a non-assembling control vaccine comprising all four RBD-I53-50A trimeric components and a non-assembling I53-50B pentamer (**Fig. S9.3F**) in an immunization study in BALB/c cByJ mice (**Figs. 9.4A and S9.3A**). All immunizations comprised 1 μg of total RBD antigen, such that ~0.25 μg of each RBD was given in each administration of the quadrivalent vaccines and ~0.33 μg of each RBD in the trivalent vaccines. After two immunizations, all four mosaic or cocktail RBD-NP vaccines elicited strong binding (GMT ~1–5 × 10^4) (**Fig. S9.4A; Data S9.1**) and potent serum neutralizing (GMT 2–8 × 10^3; **Fig. 5A; Data S1**) Ab titers against wild-type (D614G) SARS-CoV-2 S pseudovirus. Competition ELISAs using hACE2, CR3022, S309, and S2X259 demonstrated that all mosaic and cocktail vaccine candidates elicited robust Ab titers targeting all antigenic sites tested, showcasing the diversity of RBD-specific Abs elicited (**Figs. S9.4B–S9.4E; Data S9.1**). The neutralizing SARS-CoV-2 S (D614G) Ab responses were slightly higher for mRBD-NPs and slightly lower for cRBD-NPs than that of the monovalent RBD-NP (GMT 3 × 10^3) (**Fig. 9.5A**), similar to the aforementioned low-dose RBD-NP immunization study (**Fig. 9.1D**) and suggesting that the dose of strain-matched antigen in the multivalent vaccines is not a limiting factor of the magnitude of neutralizing Ab responses. The neutralizing activity against SARS-CoV-2 elicited by the other monovalent RBD-NPs and the non-

assembling control were ~2 orders of magnitude lower (GMT < 5×10^1) than for the monovalent RBD-NP, indicating poor elicitation of cross-neutralizing Abs by these heterologous monovalent RBD-NPs.

Although ELISA binding titers were comparable across mosaic and cocktail groups against SARS-CoV S2P (GMT 3–10×10^3), the corresponding pseudovirus neutralization titers showed more nuanced patterns (**Figs. 9.5B and S9.4F; Data S9.1**). Tetravalent mosaic and cocktail RBD-NPs elicited potent neutralizing activity (GMT 3–5.5×10^3) with magnitudes roughly comparable to that of the monovalent SARS-CoV-RBD-NP (GMT 5×10^3) (**Fig. 9.5B**). Strikingly, the trivalent nanoparticle immunogens (mRBD-NP-DO and cRBD-NP-DO)—which did not contain the SARS-CoV RBD—also elicited potent neutralization (GMT ~1–2×10^3). This cross-neutralization likely arose from the inclusion of the WIV1 RBD (**Fig. S9.4G; Data S9.1**) in the trivalent immunogens, as WIV1 cross-reacts with a SARS-CoV-specific mAb (**Fig. S9.3H**) (Menachery et al., 2016), and the monovalent WIV1-RBD-NP induced similar levels of pseudovirus neutralization (GMT 2×10^3) (**Fig. 9.5B; Data S9.1**). The non-assembling control immunogen, which contains all four RBD-I53-50A trimeric components, also elicited substantial neutralizing activity against SARS-CoV (GMT 1×10^3) but not against SARS-CoV-2. Furthermore, we show that sera elicited by the monovalent, mRBD-NPs, and cRBD-NPs neutralize SHC014 pseudotyped virus, a distantly related sarbecovirus sharing 76%, 76%, 82%, and 84% amino acid sequence identity with the SARS-CoV-2, RaTG13, SARS-CoV, and WIV1 RBDs, respectively (**Fig. 9.5C; Data S9.1**). These data show that both mRBD-NPs and cRBD-NPs are promising vaccine candidates for eliciting broad sarbecovirus immunity, in agreement with previous

findings using a different nanoparticle platform (Cohen et al., 2021a) or chimeric sarbecovirus S glycoproteins (Martinez et al., 2021b) and recent results obtained with influenza virus hemagglutinin nanoparticle vaccines (Boyoglu-Barnum et al., 2021).

9.8 Mosaic RBD-NPs protect mice against heterotypic SARS-CoV challenge

To gauge the ability of the multivalent RBD-NPs to confer protection against vaccine-matched and heterotypic sarbecoviruses, we challenged groups of eight mice with a high dose (10^5 PFUs) of the mouse-adapted SARS-CoV MA15 virus (Roberts et al., 2007). In agreement with the pseudovirus neutralization data, animals immunized with mRBD-NP, mRBD-NP-DO, cRBD-NP, and cRBD-NP-DO were protected from weight loss (ranging between 1% and 5%) and serious lung pathology throughout the four days of the experiment **(Figs. 9.6A and 9.6B**). Unvaccinated mice exhibited up to 15% average weight loss and signs of lung pathology (congestion score ~1.5) that were only minorly seen in all other vaccinated groups (**Figs. 9.6A and 9.6B**). mRBD-NP- and mRBD-NP-DO-vaccinated mice were completely protected from viral replication in the lungs (**Fig. 9.6C**). Overall, the trivalent mRBD-NP-DO provided protection that was indistinguishable from the tetravalent mRBD-NP, despite lacking the SARS-CoV RBD. Five out of eight and six out of eight cRBD-NP- and cRBD-NP-DO-vaccinated mice were completely protected from viral replication in the lungs, respectively, with viral burden ranging between 10^2 and 10^3 PFUs/lobe for the remaining animals (compared to unvaccinated mice with ~10^6 PFUs/lobe, **Fig. 9.6C**). To assess whether the RBD-NP platform causes vaccine-induced immune pathology, which was previously described with double-inactivated SARS-CoV vaccines (Bolles et al., 2011), we investigated eosinophil infiltration and inflammation in the lungs following SARS-CoV MA15

challenge. Although infection induced inflammation and infiltration of eosinophils in the lungs of unvaccinated mice (**Figs. 9.6D, 9.6I, and S9.5A**), no signs of eosinophilia were detected for any animals in the vaccinated groups using histological analysis of stained lung sections (**Figs. 9.6D–9.6I and S9.5A**). These data show that mRBD-NP, mRBD-NP-DO, cRBD-NP, and cRBD-NP-DO protect against weight loss, lung injury, and viral replication following challenge, with no detectable vaccine-induced immune pathology.

To identify the gain in protective breadth provided by multivalent display of sarbecovirus RBD-NP, relative to monovalent RBD-NPs, we performed a second SARS-CoV MA15 challenge experiment including monovalent RBD-NP and SARS-CoV-RBD-NP (**Fig. S9.5B**). In agreement with the pseudovirus neutralization data and first challenge (**Figs. 9.4B–9.4D and 9.6A–9.6C**), animals immunized with the SARS-CoV-RBD-NP, mosaic, and cocktail RBD-NP formulations, as well as the non-assembling control immunogen, were protected from weight loss and serious lung pathology throughout the four days of the experiment. The animals receiving the monovalent RBD-NP vaccine experienced up to 12% average weight loss, whereas the unvaccinated mice exhibited further (up to 15%) average weight loss and signs of lung pathology (congestion score ~2) that were only minorly seen in all other vaccinated groups (**Figs. S9.5B and S9.5C**). All mice vaccinated with SARS-CoV-RBD-NP, RBD-NP, and mRBD-NP were protected from viral replication in the lungs, whereas we detected $\sim 10^{2.5}$ and 10^{6} PFUs/lobe for half of the mice receiving the non-assembling control immunogen and unvaccinated mice, respectively (**Fig. S9.5D**). These results provide proof-of-principle that mosaic and cocktail RBD nanoparticle vaccines elicit

broad protection against heterotypic sarbecovirus challenge and could represent the next generation of vaccines developed in anticipation of future spillovers.

9.9 Chapter Discussion

The data presented here show that two ultra-low dose immunizations or a single immunization with RBD-NP produces potent neutralizing Ab responses in mice. The latter strategy confers protection against SARS-CoV-2 MA10 challenge, suggesting that the nanoparticle platform could enable dose-sparing regimens to achieve global vaccination, especially given the high shelf stability of this vaccine (Walls et al., 2020b). NHP vaccination with RBD-NP was concurrently shown to also elicit CD4 T cell responses and to protect from viral replication (Arunachalam et al., 2021). Moreover, RBD-NP vaccination, which is currently under phase I/II evaluation in the clinic, elicits diverse Ab responses neutralizing a broad spectrum of SARS-CoV-2 variants detected in clinical isolates with similar potency to those elicited by HexaPro S vaccination. Several RBM mutations, including at position E484, however, lead to reductions in neutralizing activity elicited by either RBD-NP or HexaPro S in NHPs. Although both RBD-NP- and HexaPro S-elicited sera robustly neutralize the B.1.1.7 S variant, which does not contain the E484 substitution, neutralization of the B.1.351, P.1, and B.1.1.7/E484K variants was dampened similarly to sera from individuals vaccinated twice with the Pfizer-BioNTech BNT162b2 mRNA. These findings are in agreement with reports showing that the serum neutralizing activity against the B.1.351 variant from mRNA-1273-vaccinated individuals was comparably reduced (Wu et al., 2021), as was also the case for neutralization of authentic SARS-CoV-2 B.1.351 by HCP (Wibmer et al., 2021). Collectively, these data indicate that a significant fraction of vaccine-

elicited neutralizing activity is directed to the RBM, which is the target of potent neutralizing Abs, irrespective of the antigen design strategy (RBD- or S-based), the vaccine modality (protein subunit or mRNA), or species (NHPs or humans). We show that receiving a third immunization, even of an unaltered vaccine, improves the neutralizing Ab responses to VOCs, which is reminiscent of what has been shown in vaccinated humans previously infected with SARS-CoV-2 (Abu Jabal et al., 2021; Krammer et al., 2021; Stamatatos et al., 2021), suggesting that further booster immunizations may limit the impact of VOCs without vaccine update.

The ongoing global spread of SARS-CoV-2 and the circulation of a large number of sarbecoviruses in bats (Menachery et al., 2015, 2016) and other animal reservoirs strongly motivate the development of vaccines that protect against a broad spectrum of coronaviruses. We observed that vaccination of NHPs with RBD-NPs or HexaPro S induced comparable but moderate neutralization breadth against genetically distinct sarbecoviruses. We demonstrate here that co-display (mRBD-NP) and co-immunization (cRBD-NP) of multivalently arrayed sarbecovirus RBDs elicit robust neutralizing Ab responses against SARS-CoV-2, SARS-CoV, and SHC014 pseudotyped viruses and outperform their monovalent vaccine counterparts. The observation that mosaic and cocktail RBD-NPs elicit greater titers of Abs competing with the broadly neutralizing sarbecovirus mAb S2X259 than monovalent vaccines suggest that this vaccine design strategy could enhance neutralization breadth. Accordingly, mice vaccinated with multivalent vaccines were protected from disease upon SARS-CoV MA15 challenge, including with formulations that did not include the SARS-CoV RBD, whereas RBD-NP-vaccinated mice experienced weight loss. Previous studies showed that double-

inactivated SARS-CoV vaccines could lead to eosinophil infiltration in the lungs and subsequent immune pathology (Bolles et al., 2011). Our data show that none of the multivalent vaccine candidates evaluated here induced eosinophilia whereas SARS-CoV MA15 challenge did, pointing to the safety and efficacy of these immunogens. Our data highlight the potential of this approach to achieve broad sarbecovirus immunity, overcoming both the emergence of SARS-CoV-2 variants and putative future zoonosis of genetically distinct sarbecoviruses not included in the vaccine. Based on the RBD-centric nature of neutralizing Ab responses resulting from infection and vaccination (Dejnirattisai et al., 2021; Greaney et al., 2021a, 2021b; Piccoli et al., 2020), irrespective of immunogen format or vaccine modality, and the enhanced elicitation of Abs targeting all RBD antigenic sites for multivalent RBD-NPs compared to S-based vaccines, this study paves the way for advancing RBD pan-sarbecovirus subunit vaccines to the clinic.

9.10 Methods

Limitations of study

Due to the cost of challenge and BSL-3 vivarium limitations, the mice vaccinated with a low dose of RBD-NP were not challenged, and therefore we can only infer the likelihood of protection from their sera's ability to neutralize SARS-CoV-2 pseudovirus better than other mice (i.e., those given a single vaccine dose) that were challenged. Neutralization is a correlate of protection in NHPs (Arunachalam et al., 2021), but we recognize that neutralization assays using HEK-ACE2 cell lines may underestimate the effects of non-RBM targeting Abs (Lempp et al., 2021; McCallum et al., 2021a; Suryadevara et al., 2021). Due to the limitations of the number of animals and

the cost of challenge, we cannot challenge with all desired point mutant or variant strains and must use neutralization as the best correlate for interpreting the expected challenge outcome.

Cell lines

Expi293F (derived from 293 cells which are female) cells are derived from the HEK293F cell line (Life Technologies). Expi293F cells were grown in Expi293 Expression Medium (Life Technologies), cultured at 36.5°C with 8% CO_2 and shaking at 150 rpm. HEK293T/17 is a female human embryonic kidney cell line (ATCC). The HEK-ACE2 (derived from HEK293T cells which are female) adherent cell line was obtained through BEI Resources, NIAID, NIH: Human Embryonic Kidney Cells (HEK293T) Expressing Human Angiotensin-Converting Enzyme 2, HEK293T-hACE2 Cell Line, NR-52511. All adherent cells were cultured at 37°C with 8% CO_2 in flasks with DMEM + 10% FBS (Hyclone) + 1% penicillin-streptomycin. Cell lines other than Expi293F were not tested for mycoplasma contamination nor authenticated**.**

Mice

Female BALB/c mice (Stock # 000651, BALB/c cByJ mice) four weeks old were obtained from Jackson Laboratory, Bar Harbor, Maine, and maintained at the Comparative Medicine Facility at the University of Washington, Seattle, WA, accredited by the American Association for the Accreditation of Laboratory Animal Care International (AAALAC). Animal procedures were performed under the approvals of the Institutional Animal Care and Use Committee (IACUC) of University of Washington, Seattle, WA, and University of North Carolina, Chapel Hill, NC.

Pigtail macaques

Two adult male Pigtail macaques (*Macaca nemestrina*) were immunized in this study. All animals were housed at the Washington National Primate Research Center (WaNPRC), an AAALAC International accredited institution. All experiments were approved by The University of Washington's IACUC. Animals were singly housed in comfortable, clean, adequately-sized cages with ambient temperatures between 72–82°F. Animals received environmental enrichment for the duration of the study including grooming contact, perches, toys, foraging experiences and access to additional environment enrichment devices. Water was available through automatic watering devices and animals were fed a commercial monkey chow, supplemented daily with fruits and vegetables. Throughout the study, animals were checked twice daily by husbandry staff.

Rhesus macaques

Male Rhesus macaques (Macaca mulatta) of Indian origin, aged 3–7 years were assigned to the study (Arunachalam et al., 2021). Animals were distributed between the groups such that the age and weight distribution were comparable across the groups. Animals were housed and maintained at the New Iberia Research Center (NIRC) of the University of Louisiana at Lafayette, an AAALAC International accredited institution, in accordance with the rules and regulations of the Guide for the Care and Use of Laboratory Animal Resources. The entire study (protocol 2020-8808-15) was reviewed and approved by the University of Louisiana at Lafayette IACUC. All animals were negative for SIV, simian T cell leukemia virus, and simian retrovirus.

Convalescent human sera

Samples collected between 1–60 days post infection from individuals who tested positive for SARS-CoV-2 by PCR were profiled for anti-SARS-CoV-2 S antibody responses and those with anti-S Ab responses were maintained in the cohort (Walls et al., 2020b). Individuals were enrolled as part of the HAARVI study at the University of Washington in Seattle, WA. Baseline sociodemographic and clinical data for these individuals are summarized in Table S1. This study was approved by the University of Washington Human Subjects Division Institutional Review Board (STUDY00000959 and STUDY00003376). All experiments were performed in at least two replicates. One sample is the 20/130 COVID-19 plasma from NIBSC (https://www.nibsc.org/documents/ifu/20-130.pdf).

Pfizer vaccinated human sera

Blood samples were collected from participants who had received both doses of the Pfizer mRNA vaccine and were 7–30 days post second vaccine dose. Clinical data for these individuals are summarized in Table S1. Individuals were enrolled in the UWARN: COVID-19 in WA study at the University of Washington in Seattle, WA. This study was approved by the University of Washington Human Subjects Division Institutional Review Board (STUDY00010350).

Plasmid construction

The SARS-CoV-2-RBD-Avi construct was synthesized by GenScript into pcDNA3.1- with an N-terminal mu-phosphatase signal peptide and a C-terminal octa-histidine tag, flexible linker, and avi tag (GHHHHHHHHGGSSGLNDIFEAQKIEWHE).

The boundaries of the construct are N-$_{328}$RFPN$_{331}$ and $_{528}$KKST$_{531}$-C (Walls et al., 2020a). The GD-Pangolin (326-527), WIV1 (316-518), RaTG13 (359-562), RmYN02 (307-492), and ZXC21 (323-507) were synthesized by GenScript into vector pcDNA3.1- or CMVR with a preceding mu-phosphatase signal peptide and a C-terminal octahistidine tag. SARS-CoV-1 (306-575) was subcloned from a GenArt synthesized SARS-CoV-1 Spike ectodomain. The SARS-CoV S2P ((Pallesen et al., 2017)) was synthesized by GeneArt and placed into a modified pOPING vector with its original N-terminal mu-phosphatase signal peptide, and an engineered C-terminal extension: SG-RENLYFQG (TEV protease site), GGGSG-YIPEAPRDGQAYVRKDGEWVLLSTFL (foldon trimerization motif), G-HHHHHH (hexa-histidine tag), just upstream of the predicted transmembrane region (YIK). The SARS-CoV-2 S2P ectodomain trimer (GenBank: YP_009724390.1, BEI NR-52420) was synthesized by GenScript into pCMV with an N-terminal mu-phosphatase signal peptide and a C-terminal TEV cleavage site (GSGRENLYPQG), T4 fibritin foldon (GGGSGYIPEAPRDGQAYVRKDGEWVLLSTPL), and octa-histidine tag (GHHHHHHHH) (Walls et al., 2020a). The construct contains the 2P mutations (proline substitutions at residues 986 and 987; and an $_{682}$SGAG$_{685}$ substitution at the furin cleavage site. The SARS-CoV-2 RBD was genetically fused to the N terminus of the trimeric I53-50A nanoparticle component using 12 or 16 glycine and serine residues. RBD-12GS-I53-50A fusions were synthesized and cloned by Genscript into pCMV. The RBD-16GS-I53-50A fusion was cloned into pCMV/R using the Xba1 and AvrII restriction sites and Gibson assembly (Gibson et al., 2009). All RBD-bearing components contained an N-terminal mu-phosphatase signal peptide and a C-terminal octa-histidine tag. Human ACE2

ectodomain was genetically fused to a sequence encoding a thrombin cleavage site and a human Fc fragment at the C-terminal end. hACE2-Fc was synthesized and cloned by GenScript with a BM40 signal peptide. Genes encoding CR3022 heavy and light chains were ordered from GenScript and cloned into pCMV/R. S309 construct as previously described (Pinto et al., 2020). SARS-CoV-2 HexaPro construct is as previously described (Hsieh et al., 2020) and placed into CMVR with an octa-his tag.

Transient transfection

Proteins were produced using endotoxin free DNA in Expi293F cells grown in suspension using Expi293F expression medium (Life Technologies) at 33°C, 70% humidity, 8% CO_2 rotating at 150 rpm. The cultures were transfected using PEI-MAX (Polyscience) with cells grown to a density of 3.0 million cells per mL and cultivated for 3 days. Supernatants were clarified by centrifugation (5 min at 4000 r*cf.*), addition of PDADMAC solution to a final concentration of 0.0375% (Sigma Aldrich, #409014), and a second spin (5 min at 4000 r*cf.*).

Microbial protein expression and purification

The I53-50A and I53-50B.4.PT1 proteins (Bale et al., 2016) were expressed in Lemo21(DE3) (NEB) in LB (10 g Tryptone, 5 g Yeast Extract, 10 g NaCl) grown in 2 L baffled shake flasks or a 10 L BioFlo 320 Fermenter (Eppendorf). Cells were grown at 37°C to an OD600 ~0.8, and then induced with 1 mM IPTG. Expression temperature was reduced to 18°C and the cells shaken for ~16 h. The cells were harvested and lysed by microfluidization using a Microfluidics M110P at 18,000 psi in 50 mM Tris, 500 mM NaCl, 30 mM imidazole, 1 mM PMSF, 0.75% CHAPS. Lysates were clarified by

centrifugation at 24,000 g for 30 min and applied to a 2.6 × 10 cm Ni Sepharose 6 FF column (Cytiva) for purification by IMAC on an AKTA Avant150 FPLC system (Cytiva). Protein of interest was eluted over a linear gradient of 30 mM to 500 mM imidazole in a background of 50 mM Tris pH 8, 500 mM NaCl, 0.75% CHAPS buffer. Peak fractions were pooled, concentrated in 10K MWCO centrifugal filters (Millipore), sterile filtered (0.22 μm) and applied to either a Superdex 200 Increase 10/300, or HiLoad S200 pg GL SEC column (Cytiva) using 50 mM Tris pH 8, 500 mM NaCl, 0.75% CHAPS buffer. I53-50A elutes at ~0.6 column volume (CV). I53-50B.4PT1 elutes at ~0.45 CV. After sizing, bacterial-derived components were tested to confirm low levels of endotoxin before using for nanoparticle assembly.

Protein purification

Proteins containing His tags were purified from clarified supernatants via a batch bind method where each clarified supernatant was supplemented with 1 M Tris-HCl pH 8.0 to a final concentration of 45 mM and 5 M NaCl to a final concentration of ~310 mM. Talon cobalt affinity resin (Takara) was added to the treated supernatants and allowed to incubate for 15 min with gentle shaking. Resin was collected using vacuum filtration with a 0.2 μm filter and transferred to a gravity column. The resin was washed with 20 mM Tris pH 8.0, 300 mM NaCl, and the protein was eluted with 3 column volumes of 20 mM Tris pH 8.0, 300 mM NaCl, 300 mM imidazole. The batch bind process was then repeated and the first and second elutions combined. SDS-PAGE was used to assess purity. RBD-I53-50A fusion protein IMAC elutions were concentrated to > 1 mg/mL and subjected to three rounds of dialysis into 50 mM Tris pH 7.4, 185 mM NaCl, 100 mM Arginine, 4.5% glycerol, and 0.75% w/v 3-[(3-cholamidopropyl)dimethylammonio]-1-

propanesulfonate (CHAPS) in a hydrated 10K molecular weight cutoff dialysis cassette (Thermo Scientific). S2P and HexaPro IMAC elution fractions were concentrated to ~1 mg/mL and dialyzed three times into 50 mM Tris pH 8, 150 mM NaCl, 0.25% L-Histidine in a hydrated 10K molecular weight cutoff dialysis cassette (Thermo Scientific). Clarified supernatants of cells expressing monoclonal antibodies and human ACE2-Fc were purified using a MabSelect PrismA 2.6 × 5 cm column (Cytiva) on an AKTA Avant150 FPLC (Cytiva). Bound antibodies were washed with five column volumes of 20 mM $NaPO_4$, 150 mM NaCl pH 7.2, then five column volumes of 20 mM $NaPO_4$, 1 M NaCl pH 7.4 and eluted with three column volumes of 100 mM glycine at pH 3.0. The eluate was neutralized with 2 M Trizma base to 50 mM final concentration. SDS-PAGE was used to assess purity.

Recombinant S309 was expressed as a Fab in expiCHO cells transiently co-transfected with plasmids expressing the heavy and light chain, as described above (see Transient transfection) (Pinto et al., 2020). The protein was affinity-purified using a HiTrap Protein A Mab select Xtra column (Cytiva) followed by desalting against 20 mM $NaPO_4$, 150 mM NaCl pH 7.2 using a HiTrap Fast desalting column (Cytiva). The protein was sterilized with a 0.22 μm filter and stored at 4°C until use.

In vitro nanoparticle assembly and purification

Total protein concentration of purified individual nanoparticle components was determined by measuring absorbance at 280 nm using a UV/vis spectrophotometer (Agilent Cary 8454) and calculated extinction coefficients. The assembly steps were performed at room temperature with addition in the following order: RBD-I53-50A trimeric fusion protein, followed by additional buffer (50 mM Tris pH 7.4, 185 mM NaCl,

100 mM Arginine, 4.5% glycerol, and 0.75% w/v CHAPS) as needed to achieve desired final concentration, and finally I53-50B.4PT1 pentameric component (in 50 mM Tris pH 8, 500 mM NaCl, 0.75% w/v CHAPS), with a molar ratio of RBD-I53-50A:I53-50B.4PT1 of 1.1:1. All RBD-I53-50 *in vitro* assemblies were incubated at 2-8°C with gentle rocking for at least 30 min before subsequent purification by SEC in order to remove residual unassembled component. Different columns were utilized depending on purpose: Superose 6 Increase 10/300 GL column was used analytically for nanoparticle size estimation, a Superdex 200 Increase 10/300 GL column used for small-scale pilot assemblies, and a HiLoad 26/600 Superdex 200 pg column used for nanoparticle production. Assembled particles were purified in 50 mM Tris pH 7.4, 185 mM NaCl, 100 mM Arginine, 4.5% glycerol, and 0.75% w/v CHAPS, and elute at ~11 mL on the Superose 6 column and in the void volume of Superdex 200 columns. Assembled nanoparticles were sterile filtered (0.22 μm) immediately prior to column application and following pooling of fractions.

Endotoxin measurements

Endotoxin levels in protein samples were measured using the EndoSafe Nexgen-MCS System (Charles River). Samples were diluted 1:50 or 1:100 in Endotoxin-free LAL reagent water, and applied into wells of an EndoSafe LAL reagent cartridge. Charles River EndoScan-V software was used to analyze endotoxin content, automatically back-calculating for the dilution factor. Endotoxin values were reported as EU/mL which were then converted to EU/mg based on UV/vis measurements. Our threshold for samples suitable for immunization was < 50 EU/mg.

Pigtail macaque immunization

Two Pigtail macaques were immunized with 250 µg of RBD-12GS-I53-50 nanoparticle (88 µg RBD antigen) with AddaVax at day 0, day 28, and O/W at day 168. Blood was collected every 14 days post-prime. Blood was collected in serum collection tubes and allowed to clot at room temperature. Serum was isolated after a 15 min spin at 1455 x g for 15 min and stored at −80C until use. Prior to vaccination or blood collection, animals were sedated with an intramuscular injection (10 mg/kg) of ketamine (Ketaset®; Henry Schein). Prior to inoculation, immunogen suspensions were gently mixed 1:1 vol/vol with AddaVax adjuvant for immunizations 1 and 2 and O/W for immunization 3 (VFI) to reach a final concentration of 0.250 mg/mL antigen. The vaccine was delivered intramuscularly into both quadriceps muscles with 1 mL per injection site on days 0, 28, and 168. All injection sites were shaved prior to injection. Animals were observed daily for general health (activity and appetite, urine/feces output) and for evidence of reactogenicity at the vaccine inoculation site (swelling, erythema, and pruritus) for up to 1 week following vaccination. They also received physical exams including temperature and weight measurements at each study time point. None of the animals became severely ill during the course of the study nor required euthanasia.

Rhesus macaque immunization

Adapted from Arunachalam et al. (2021). AS03 was kindly provided by GSK Vaccines. AS03 is an oil-in-water emulsion that contains 11.86 mg a-tocopherol, 10.69 mg squalene, and 4.86 mg polysorbate 80 (Tween-80) in PBS. For each dose, RBD-NP was diluted to 50 µg/ml (RBD component) in 250 µL of Tris-buffered saline

(TBS) and mixed with an equal volume of AS03.The dose of AS03 was 50% v/v (equivalent of one human dose). Soluble HexaPro was diluted to 50 µg/ml in 250 µL of Tris-buffered saline (TBS) and mixed with an equal volume of AS03. All immunizations were administered via the intramuscular route in right forelimbs. The volume of each dose was 0.5 ml.

Deep mutational scanning

All mutations that escape serum antibody binding were mapped via a deep mutational scanning approach (Greaney et al., 2020, 2021b). We used previously described yeast-display RBD mutant libraries (Greaney et al., 2020; Starr et al., 2020a). Briefly, duplicate mutant libraries were constructed in the spike receptor binding domain (RBD) from SARS-CoV-2 (isolate Wuhan-Hu-1, GenBank: MN908947, residues N331-T531) and contain 3,804 of the 3,819 possible amino-acid mutations, with > 95% present as single mutants. Each RBD variant was linked to a unique 16-nucleotide barcode sequence to facilitate downstream sequencing. As previously described, libraries were sorted for RBD expression and ACE2 binding to eliminate RBD variants that are completely misfolded or non-functional (i.e., lacking modest ACE2 binding affinity) (Greaney et al., 2020).

Antibody escape mapping experiments were performed in biological duplicate using two independent mutant RBD libraries, with minor modifications from Greaney et al. (2020), and exactly as described in Greaney et al. (2021b). The antibody escape mapping for the vaccinated NHP serum was performed in this study; the antibody escape mapping from convalescent human plasma was performed in Greaney et al. (2021b). Briefly, mutant yeast libraries induced to express RBD were washed and

incubated with plasma or serum at a range of dilutions for 1 h at room temperature with gentle agitation. For each sample, we chose a sub-saturating dilution such that the amount of fluorescent signal due to plasma antibody binding to RBD was approximately equal across samples. A 1:1000 dilution was used for the vaccinated NHP serum, and the exact dilutions of human convalescent plasma are reported in Greaney et al. (2021b). After the antibody incubations, the libraries were secondarily labeled with 1:100 FITC-conjugated anti-MYC antibody (Immunology Consultants Lab, CYMC-45F) to label for RBD expression and and 1:200 Alexa-647-conjugated goat anti-human-IgA+IgG+IgM (Jackson ImmunoResearch 109-605-064) to label for bound serum or plasma antibodies. A flow cytometric selection gate was drawn to capture approximately 5% of the RBD mutant libraries with the lowest amount of plasma binding for their degree of RBD expression. For each sample, approximately 10 million RBD+ cells were processed on the cytometer. Antibody-escaped cells were grown overnight in SD-CAA (6.7g/L Yeast Nitrogen Base, 5.0g/L Casamino acids, 1.065 g/L MES acid, and 2% w/v dextrose) to expand cells prior to plasmid extraction.

Plasmid samples were prepared from pre-selection and overnight cultures of antibody-escaped cells (Zymoprep Yeast Plasmid Miniprep II) as previously described (Greaney et al., 2020). The 16-nucleotide barcode sequences identifying each RBD variant were amplified by PCR and sequenced on an Illumina HiSeq 2500 with 50 bp single-end reads as described in Greaney et al. (2020and Starr et al. (2020a).

Escape fractions were computed as described in Greaney et al. (2020), and exactly as described in Greaney et al. (2021b). We used the dms_variants package (https://jbloomlab.github.io/dms_variants/, version 0.8.2) to process Illumina sequences

into counts of each barcoded RBD variant in each pre-sort and antibody-escape population using the barcode/RBD look-up table from Starr et al. (2021).

For each serum selection, we computed the "escape fraction" for each barcoded variant using the deep sequencing counts for each variant in the original and serum-escape populations and the total fraction of the library that escaped antibody binding via the formula provided in Greaney et al. (2020). These escape fractions represent the estimated fraction of cells expressing that specific variant that fall in the antibody escape bin, such that a value of 0 means the variant is always bound by serum and a value of 1 means that it always escapes serum binding. We then applied a computational filter to remove variants with low sequencing counts or highly deleterious mutations that might cause antibody escape simply by leading to poor expression of properly folded RBD on the yeast cell surface (Greaney et al., 2020; Starr et al., 2020a). Specifically, we removed variants that had (or contained mutations with) ACE2 binding scores < −2.35 or expression scores < −1, using the variant- and mutation-level deep mutational scanning scores from Starr et al. (2020a). Note that these filtering criteria are slightly more stringent than those used in Greaney et al. (2020)but are identical to those used in Greaney et al. (2021b) and Starr et al. (2020b).

We next deconvolved variant-level escape scores into escape fraction estimates for single mutations using global epistasis models (Otwinowski et al., 2018) implemented in the dms_variants package, as detailed at (https://jbloomlab.github.io/dms_variants/dms_variants.globalepistasis.html) and described in Greaney et al. (2020). The reported escape fractions throughout the paper are the average across the libraries (correlations shown in **Figure S8.2**); these scores

are also in Table S1 and at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_RBD-nano-vax-NHP1/blob/main/results/supp_data/NHP_HCS_raw_data.csv. Sites of strong escape from each antibody were determined heuristically as sites whose summed mutational escape scores were at least 10 times the median sitewise sum of selection, and within 10-fold of the sitewise sum of the most strongly selected site. Sites shown in **Figs. 2 and S2A** are the sites of strong escape for any of the three human convalescent plasma, plus sites 417, 452, and 501 due to their prevalence in circulating SARS-CoV-2 variants. For each plasma, the y axis is scaled to be the greatest of (a) the maximum site-wise escape metric observed for that plasma, (b) 20x the median site-wise escape fraction observed across all sites for that plasma, or (c) an absolute value of 1.0 (to appropriately scale plasma that are not "noisy" but for which no mutation has a strong effect on plasma binding). Full documentation of the computational analysis is at https://github.com/jbloomlab/SARS-CoV-2-RBD_MAP_RBD-nano-vax-NHP1. These results are also available in a zoomable, interactive form at https://jbloomlab.github.io/SARS-CoV-2-RBD_MAP_RBD-nano-vax-NHP1/.

ELISA

For anti-S2P ELISA, 25 µL of 2 µg/mL S2P was plated onto 384-well Nunc Maxisorp (ThermoFisher) plates in PBS and sealed overnight at RT. The next day plates were washed 4 × in Tris Buffered Saline Tween (TBST) using a plate washer (BioTek) and blocked with SuperBlock (ThermoFisher) for 1 h at 37°C. Plates were washed 4 × in TBST and 1:5 serial dilutions of mouse, NHP, or human sera were made in 25 µL TBST and incubated at 37°C for 1 h. Plates were washed 4 × in TBST, then

anti-mouse (Invitrogen), anti-NHP (AlphaDiagnostics), or anti-human (Invitrogen) horseradish peroxidase-conjugated antibodies were diluted 1:5,000 and 25 µL added to each well and incubated at 37°C for 1 h. Plates were washed 4 × in TBST and 25 µL of TMB (SeraCare) was added to every well for ~5 min at room temperature. The reaction was quenched with the addition of 25 µL of 1 N HCl. Plates were immediately read at 450 nm on a BioTek plate reader and data plotted and fit in Prism (GraphPad) using nonlinear regression sigmoidal, 4PL, X is log(concentration) to determine EC_{50} values from curve fits.

Competition ELISA of NHP sera with hACE2, CR3022 IgG, and S309 IgG for immobilized SARS-CoV-2 S2P

384-well Maxisorp plates (Thermo Fisher) were coated overnight at room temperature with 3 µg/mL of SARS-CoV-2 S2P (Pallesen et al., 2017) in 20mM Tris pH 8 and 150mM NaCl. Plates were slapped dry and blocked with Blocker Casein in TBS (Thermo Fisher) for one hour at 37°C. Plates were slapped dry and NHP sera was serially diluted 1:4 in TBST with an initial dilution of 1:4 for hACE2 competition or 1:2 for antibody competition. Random primary amine biotinylated (Thermo Fisher) hACE2-Fc, CR3022 (Yuan et al., 2020), or S309 (Pinto et al., 2020) were added, bringing the concentration of each well to the EC_{50} values of 0.2nM, 2nm, and 0.01nM, respectively. Plates were left for one hour at 37°C, then washed 4x with TBST using a 405 TS Microplate Washer (BioTek) followed by addition of 1:500 streptavidin-HRP (Thermo Fisher) for one hour at 37°C. Plates were washed 4x and TMB Microwell Peroxidase (Seracare) was added. The reaction was quenched after 1-2 minutes with 1 N HCl and the A450 of each well was read using a BioTek plate reader (BioTek). Data plotted and

fit in Prism (GraphPad) using nonlinear regression sigmoidal, 4PL, X is log(concentration) to determine EC_{50} values from curve fits with upper and lower constraints determined by uncompeted ELISA per antigen.

Competition ELISA of mouse sera and immobilized hACE2 or mAbs with SARS-CoV-2 or SARS-CoV S2P

384-well Maxisorp plates (Thermo Fisher) were coated overnight at room temperature with 3 µg/mL of hACE2-Fc, CR3022 (Yuan et al., 2020), or S309 (Pinto et al., 2020) in 20mM Tris pH 8 and 150mM NaCl. Plates were slapped dry and blocked with Blocker Casein in TBS (Thermo Fisher) for one hour at 37°C. Plates were slapped dry and a 30-minute pre-incubated 1:5 serial dilution of mouse sera in TBST, with in initial dilution of 1:50 for hACE2-Fc competition or 1:10 for antibody competition, and a constant concentration of biotinylated (Avidity) SARS-CoV-2 S2P or SARS-CoV 2P at their EC_{50} values were added. Spike concentrations were 0.63nM, 5.98nM, and 0.22nM of SARS-CoV-2 S2P or 4.11nM, 2.89nM, and 0.19nM of SARS-CoV S2P for immobilized hACE2, CR3022, and S309, respectively. Plates were left for one hour at 37°C, then washed 4x with TBST using a 405 TS Microplate Washer (BioTek) followed by addition of 1:500 streptavidin-HRP (Thermo Fisher) for one hour at 37°C. Plates were washed 4x and TMB Microwell Peroxidase (Seracare) was added. The reaction was quenched after 1-2 minutes with 1 N HCl and the A450 of each well was read using a BioTek plate reader (BioTek). Data plotted and fit in Prism (GraphPad) using nonlinear regression sigmoidal, 4PL, X is log(concentration) to determine EC_{50} values from curve fits.

Pseudovirus production

MLV-based D614G SARS-CoV-2 S and SARS-CoV S pseudotypes were prepared as previously described (Millet and Whittaker, 2016; Walls et al., 2020a, 2020b). Briefly, HEK293T cells were co-transfected using Lipofectamine 2000 (Life Technologies) with an S-encoding plasmid, an MLV Gag-Pol packaging construct, and the MLV transfer vector encoding a luciferase reporter according to the manufacturer's instructions. Cells were washed 3 × with Opti-MEM and incubated for 5 h at 37°C with transfection medium. DMEM containing 10% FBS was added for 60 h. The supernatants were harvested by spinning at 2,500 g, filtered through a 0.45 μm filter, concentrated with a 100 kDa membrane for 10 min at 2,500 g and then aliquoted and stored at −80°C.

HIV-based pseudotypes were prepared as previously described (Crawford et al., 2020). Briefly, HEK293T cells were cotransfected using Lipofectamine 2000 (Life Technologies) with an S-encoding plasmid (D614G SARS-CoV-2 S (YP 009724390.1), Pangolin-Guangdong S (QLR06867.1), SARS-CoV S (YP 009825051.1), P1, B.1.351 S, B.1.1.7, and B.1.1.7+E484K S) an HIV Gag-Pol, Tat, Rev1B packaging construct, and the HIV transfer vector encoding a luciferase reporter according to the manufacturer's instructions. Cells were washed 3 × with Opti-MEM and incubated for 5 h at 37°C with transfection medium. DMEM containing 10% FBS was added for 60 h. The supernatants were harvested by spinning at 2,500 g, filtered through a 0.45 μm filter, concentrated with a 100 kDa membrane for 10 min at 2,500 g and then aliquoted and stored at −80°C.

D614G SARS-CoV-2 S (YP 009724390.1), D614G SARS-CoV-2 point mutants, SHC014 (AGZ48806.1), RaTG13 S (QHR63300.2), Pangolin-Guangdong S (QLR06867.1), SARS-CoV S (YP 009825051.1), WIV1 S (AGZ48831.1), B.1.351 S, P1, and B.1.1.7 S pseudotyped VSV viruses were prepared as described previously (McCallum et al., 2021a; Sauer et al., 2021). Briefly, 293T cells in DMEM supplemented with 10% FBS, 1% PenStrep seeded in 10-cm dishes were transfected with the plasmid encoding for the corresponding S glycoprotein using lipofectamine 2000 (Life Technologies) following manufacturer's indications. One day post-transfection, cells were infected with VSV(G*ΔG-luciferase) and after 2 h were washed five times with DMEM before adding medium supplemented with anti-VSV-G antibody (I1- mouse hybridoma supernatant, CRL- 2700, ATCC). Virus pseudotypes were harvested 18-24 h post-inoculation, clarified by centrifugation at 2,500 x g for 5 min, filtered through a 0.45 μm cut off membrane, concentrated 10 times with a 30 kDa cut off membrane, aliquoted and stored at −80°C.

Pseudovirus neutralization

HEK293-hACE2 cells (Crawford et al., 2020) were cultured in DMEM with 10% FBS (Hyclone) and 1% PenStrep with 8% CO_2 in a 37°C incubator (ThermoFisher). One day prior to infection, 40 μL of poly-lysine (Sigma) was placed into 96-well plates and incubated with rotation for 5 min. Poly-lysine was removed, plates were dried for 5 min then washed 1 × with water prior to plating with 40,000 cells. The following day, cells were checked to be at 80% confluence. In an empty half-area 96-well plate a 1:3 serial dilution of sera was made in DMEM and diluted pseudovirus was then added to the serial dilution and incubated at room temperature for 30-60 min. After incubation, the

sera-virus mixture was added to the cells at 37°C and 2 hours post-infection, 40 µL 20% FBS-2% PenStrep DMEM was added. After 17-20 hours (VSV) or 48 hours (HIV, MLV) 40 µL/well of One-Glo-EX substrate (Promega) was added to the cells and incubated in the dark for 5-10 min prior reading on a BioTek plate reader. Measurements were done in at least duplicate. Relative luciferase units were plotted and normalized in Prism (GraphPad). Nonlinear regression of log(inhibitor) versus normalized response was used to determine IC_{50} values from curve fits. Kruskal Wallis tests were used to compare two groups to determine whether they were statistically different.

Sarbecovirus biolayer interferometry binding analysis

Purification of Fabs from NHP serum was adapted from Boyoglu-Barnum et al. (2021). Briefly, 1 mL of day 70 serum was diluted to 10 mL with PBS and incubated with 1 mL of 3 × PBS-washed protein A beads (GenScript) with agitation overnight at 37°C. The next day beads were thoroughly washed with PBS using a gravity flow column and bound Abs were eluted with 0.1 M glycine pH 3.5 into 1M Tris-HCl (pH 8.0) to a final concentration of 100 mM. Serum and early washes that flowed through were re-bound to beads overnight again for a second, repeat elution. IgGs were concentrated (Amicon 30 kDa) and buffer exchanged into PBS. 2 × digestion buffer (40 mM sodium phosphate pH 6.5, 20 mM EDTA, 40 mM cysteine) was added to concentrated and pooled IgGs. 500 µL of resuspended immobilized papain resin (ThermoFisher Scientific) freshly washed in 1 × digestion buffer (20 mM sodium phosphate, 10 mM EDTA, 20 mM cysteine, pH 6.5) was added to purified IgGs in 2 × digestion buffer and samples were agitated for 5 h at 37°C. The supernatant was separated from resin and resin washes were collected and pooled with the resin flow through. Pooled supernatants were sterile-

filtered at 0.22 µm and applied 6 × to PBS-washed protein A beads in a gravity flow column. The column was eluted as described above and the papain procedure repeated overnight with undigested IgGs to increase yield. The protein A flow throughs were pooled, concentrated (using an Amicon 10 kDa), and buffer exchanged into PBS. Purity was checked by SDS-PAGE.

Assays were performed and analyzed using biolayer interferometry on an Octet Red 96 System (Pall Forte Bio/Sartorius) at ambient temperature with shaking at 1000 rpm. Different Sarbeco RBDs were purified like in Walls et al., 2020 and were diluted with different acetate buffers and applied to a black 96-well Greiner Bio-one microplate at 200ul per well. GD-Pangolin RBD was diluted in pH 6 buffer to 5 µg/mL, RmNY02 were diluted in pH 5 to 25 µg/mL, WIV16 was diluted in pH 5 to 10 µg/mL, SARS-CoV-2 was diluted in pH 6 to 5 µg/mL, RaTG13 was diluted in pH 6 to 10 µg/mL, RaTG13 was diluted in pH 6 to 10 µg/mL, SARS-CoV was diluted in pH 6 to 50 µg/mL, and finally ZXC21 was diluted in pH 6 to 10µg/mL. AR2G biosensors (ForteBio/Sartorius) following 600 s hydration were normalized in water for 180 s. Then tips were NHS-EDC activated for 300 s and the different sarbecovirus RBDs were loaded up to a 1.50 nm threshold for up to 600 s. Immobilized RBDs on the tips were quenched for 300 s in ethanolamine and dipped into kinetics buffer for a 60 s baseline. The association step was performed by dipping the mobilized RBDs into diluted purified polyclonal pigtail macaque IgGs for 600 s. Dissociation was measured by inserting the biosensors in kinetics buffer for 600 s. The data were baseline subtracted and the plots fitted using the Pall ForteBio/Sartorius analysis software (version 12.0).

Cocktail and mosaic bio-layer interferometry (antigenicity)

Binding of hACE2-Fc to monovalent RBD-I53-50 nanoparticles, mosaic-RBD-I53-50 nanoparticles, and cocktail of RBD-nanoparticles was analyzed for antigenicity experiments and real-time stability studies using an Octet Red 96 System (Pall FortéBio/Sartorius) at ambient temperature with shaking at 1000 rpm. Protein samples were diluted to 100 nM in Kinetics buffer (1 × HEPES-EP+ (Pall Forté Bio), 0.05% nonfat milk, and 0.02% sodium azide). Buffer, receptor, and analyte were then applied to a black 96-well Greiner Bio-one microplate at 200 µL per well. Protein A biosensors (FortéBio/Sartorius) were first hydrated for 10 minutes in Kinetics buffer, then dipped into hACE2-Fc diluted to 10 µg/mL in Kinetics buffer in the immobilization step. After 150 s, the tips were transferred to kinetics buffer for 60 s to reach a baseline. The association step was performed by dipping the loaded biosensors into the immunogens for 300 s, and subsequent dissociation was performed by dipping the biosensors back into Kinetics buffer for an additional 300 s. The data were baseline subtracted prior for plotting using the FortéBio analysis software (version 12.0). Plots in **Fig. S8.3** show the 600 s of association and dissociation.

Sandwich bio-layer interferometry (mosaic display antigenicity)

Binding of hACE2-Fc or S2H14 mAb and S230 Fab to WIV1-RBD-I53-50, RaTG13-RBD-I53-50, SARS-CoV-SARS-CoV2-RBD-I53-50, SARS-CoV2-I53-50, and mosaic-RBD-I53-50 nanoparticles were analyzed for co-display of RBDs using an Octet Red 96 System (Pall FortéBio/Sartorius) at ambient temperature with shaking at 1000 rpm. Nanoparticles were diluted to 100 nM in Kinetics buffer. Kinetics buffer, mAb, nanoparticles and Fab were then applied to a black 96-well Greiner Bio-one microplate

at 200 µL per well. Protein A biosensors (FortéBio/Sartorius) were first hydrated for 10 minutes in Kinetics buffer, then dipped into hACE2-Fc or S2H14 mAb diluted to 10 µg/mL in Kinetics buffer in the immobilization step. After 150 s, the tips were transferred to Kinetics buffer for 60 s to reach a baseline. The receptor or mAb was then loaded with nanoparticle by dipping the loaded biosensors into the immunogens for 300 s, and subsequent baseline was performed by dipping the biosensors back into the Kinetics buffer for an additional 60 s. Association of S230 Fab diluted to 100 nM in Kinetics buffer was then measured for 300 s and subsequent dissociation in Kinetics buffer of S230 Fab for 300 s. The data were baseline subtracted prior for plotting using the FortéBio analysis software (version 12.0). Plots in **Fig. S8.3** exclude the initial mAb loading and the first baseline.

Cocktail and mosaic negative stain electron microscopy

Monovalent RBD-I53-50 nanoparticles, mosaic-RBD-I53-50 nanoparticles, and cocktail of RBD-nanoparticles were first diluted to 75 µg/mL in 50 mM Tris pH 7.4, 185 mM NaCl, 100 mM Arginine, 4.5% v/v Glycerol, 0.75% w/v CHAPS prior to application of 3 µL of sample onto freshly glow-discharged 300-mesh copper grids. Sample was incubated on the grid for 1 minute before the grid was dipped in a 50 µL droplet of water and excess liquid blotted away with filter paper (Whatman). The grids were then dipped into 6 µL of 0.75% w/v uranyl formate stain. Stain was blotted off with filter paper, then the grids were dipped into another 6 µL of stain and incubated for ~70 s. Finally, the stain was blotted away and the grids were allowed to dry for 1 minute. Prepared grids were imaged in a Talos model L120C electron microscope at 57,000 × (nanoparticles).

Cocktail and mosaic dynamic light scattering

Dynamic Light Scattering (DLS) was used to measure hydrodynamic diameter (Dh) and % Polydispersity (%Pd) of monovalent RBD-I53-50 nanoparticles, mosaic-RBD-I53-50 nanoparticles, and cocktail of RBD-nanoparticles on an UNcle Nano-DSF (UNchained Laboratories). Sample was applied to a 8.8 µL quartz capillary cassette (UNi, UNchained Laboratories) and measured with 10 acquisitions of 5 s each, using auto-attenuation of the laser. Increased viscosity due to 4.5% v/v glycerol in the RBD nanoparticle buffer was accounted for by the UNcle Client software in Dh measurements.

Mouse immunizations and challenge

At six weeks of age, 8 female BALB/c mice per dosing group were vaccinated with a prime immunization, and three weeks later mice were boosted with a second vaccination (IACUC protocol 4470.01). Prior to inoculation, immunogen suspensions were gently mixed 1:1 vol/vol with AddaVax adjuvant (Invivogen, San Diego, CA) to reach a final concentration of 0.01 mg/mL antigen. Mice were injected intramuscularly into the gastrocnemius muscle of each hind leg using a 27-gauge needle (BD, San Diego, CA) with 50 µL per injection site (100 µL total) of immunogen under isoflurane anesthesia. To obtain sera all mice were bled two weeks after prime and boost immunizations. Blood was collected via submental venous puncture and rested in 1.5 mL plastic Eppendorf tubes at room temperature for 30 min to allow for coagulation. Serum was separated from red blood cells via centrifugation at 2,000 g for 10 min. Complement factors and pathogens in isolated serum were heat-inactivated via incubation at 56°C for 60 min. Serum was stored at 4°C or −80°C until use. The study

was repeated twice. Five weeks post-boost, mice (aged 14 weeks) were exported from Comparative Medicine Facility at the University of Washington, Seattle, WA to an AAALAC accredited Animal Biosafety Level 3 (ABSL3) Laboratory at the University of North Carolina, Chapel Hill. After a 7-day acclimation time, mice were anesthetized with a mixture of ketamine/xylazine and challenged intranasally with 10^5 plaque-forming units (pfu) of mouse-adapted SARS-CoV-2 MA10 or SARS-CoV MA15 strain for the evaluation of vaccine efficacy (IACUC protocol 20-114.0). After infection, body weight and congestion score were monitored daily until the termination of the study two days post-infection, when lung and nasal turbinate tissues were harvested to evaluate viral load by plaque assay.

Histopathology of post challenge mouse lungs

For each mouse the left lung was incubated in formalin at 4°C for at least 7 days to fix tissue and inactivate virus. The fixed tissue was processed and embedded in paraffin. 5 µm sections were cut and stained with either Congo red or hematoxylin and eosin (H&E). Eosinophils were enumerated by counting the Congo red positive leukocytes in 10 high power fields (400X final magnification) per mouse lung. Representative images were minimally and similarly adjusted with Adobe Photoshop 2020 to enhance contrast. Airway pathology was assessed in H&E stained sections to assess bronchial epithelial cell death (score: 0 = no dead cells, 1 = 1-5 dead cells, 2 = 6-10 dead cells, 3 = 11-20 dead cells and 4 = > 20 dead cells; scored for 10 400X fields per mouse lung), peribronchial inflammation (score 0 = none, 1 = 0%–25% circumference with > 1 leukocyte cell layer, 2 = 26%–50% circumference with > 1 leukocyte cell layer, and 3 = 50%–100% circumference with > 1 leukocyte cell layer;

scored for 10 400X fields per mouse lung), and interstitial pneumonitis (score = percentage of pulmonary alveolar parenchyma with septae expanded by leukocyte; scored for 10 100X fields).

Quantification and statistical analysis

Statistical details of experiments can be found in the figure legends. For NHP experiments, 2-6 sera samples were used and experiments were done in at least duplicate unless mentioned. For mouse ELISAs, neutralization, and challenge experiments, sera from 6, 8, or 10 BALB/c animals were used and experiments were completed in at least duplicate unless mentioned. Geometric mean titers were calculated. Kruskal Wallis tests were performed to compare two groups to determine whether they were statistically different for ELISA and neutralization experiments. Significance is indicated with stars: *, $p < 0.05$; **** $p < 0.0001$ and non significant groups are not shown.

9.11 Figures and Tables

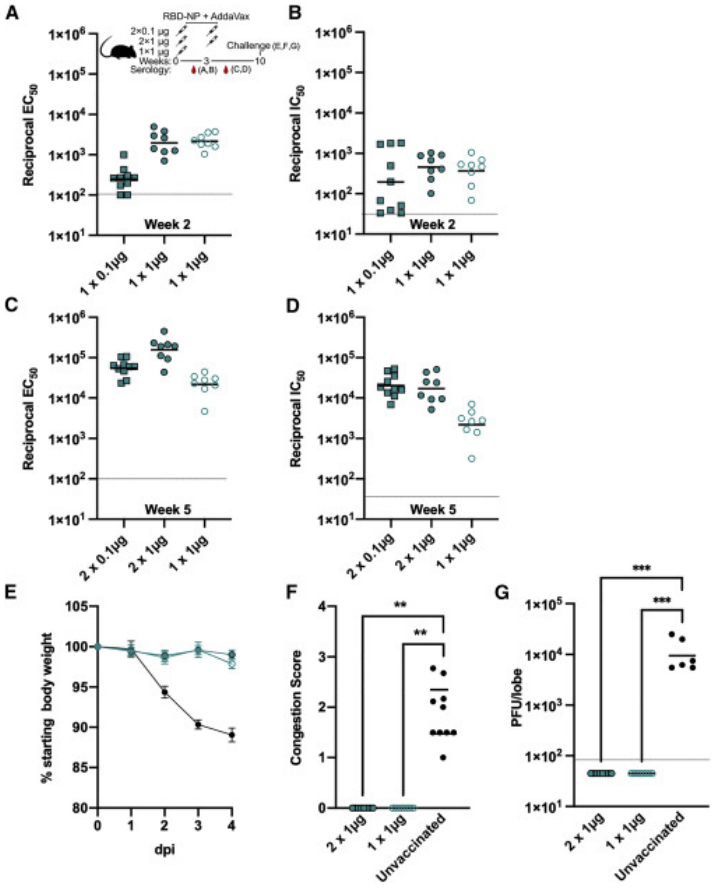


Figure 9.1: A single immunization with RBD-NP protects BALB/c cByJ mice from SARS-CoV-2 MA10 challenge.

(A and C) S2P-binding Abs were measured 2 (A) or 5 (C) weeks post-prime; teal squares: 2 doses of 0.1 µg (n = 10); filled teal circles: 2 doses of 1 µg (n = 8); open teal circles: one dose of 1 µg (n = 8) with a limit of detection (LOD) of $1 \times 10^{2.}$ (B and D) Serum neutralizing Ab titers at 2 (B) or 5 (D) weeks post-prime determined using an MLV pseudotyping system with an LOD of 3.3×10^{1}. (E) Weight loss following SARS-CoV-2 MA10 challenge up to 4 days post infection (n = 8 vaccinated; n = 6 naive mice shown as black filled circles). (F) Congestion score following SARS-CoV-2 MA10 challenge with a score of 0 indicating unchanged lung color and 4 indicating a darkened and diseased lung. (G) Viral titers in the mice lungs (expressed in plaque forming units [PFUs] per lobe) following challenge with an LOD of 9×10^{1}. Statistical significance was determined by Kruskal-Wallis test and shown only when significant. $^{**}p < 0.01$. LODs are shown as gray horizontal dotted lines.

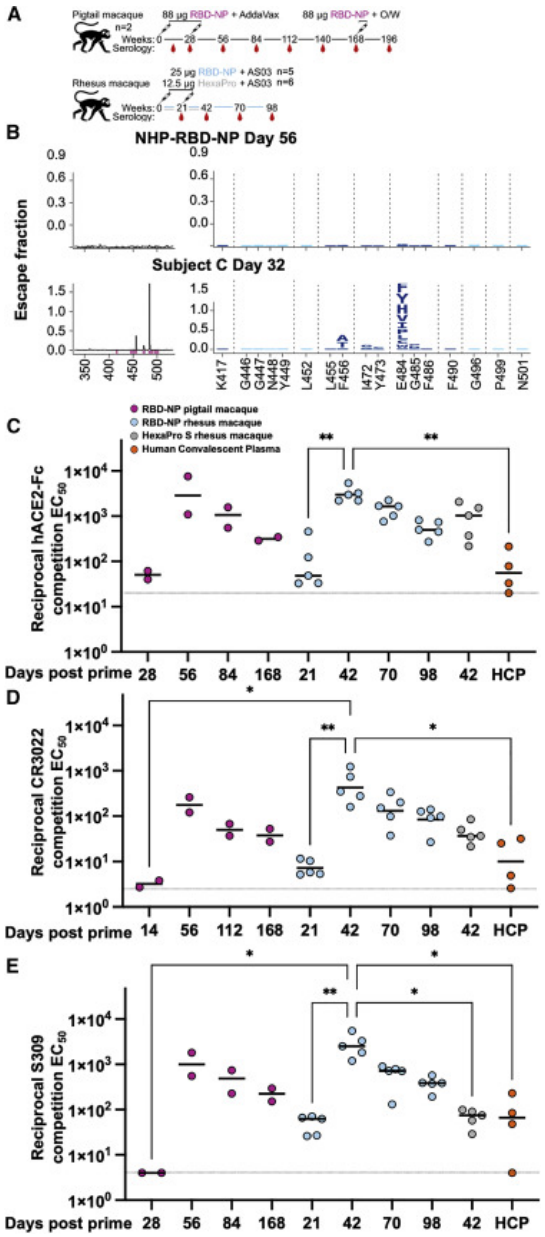
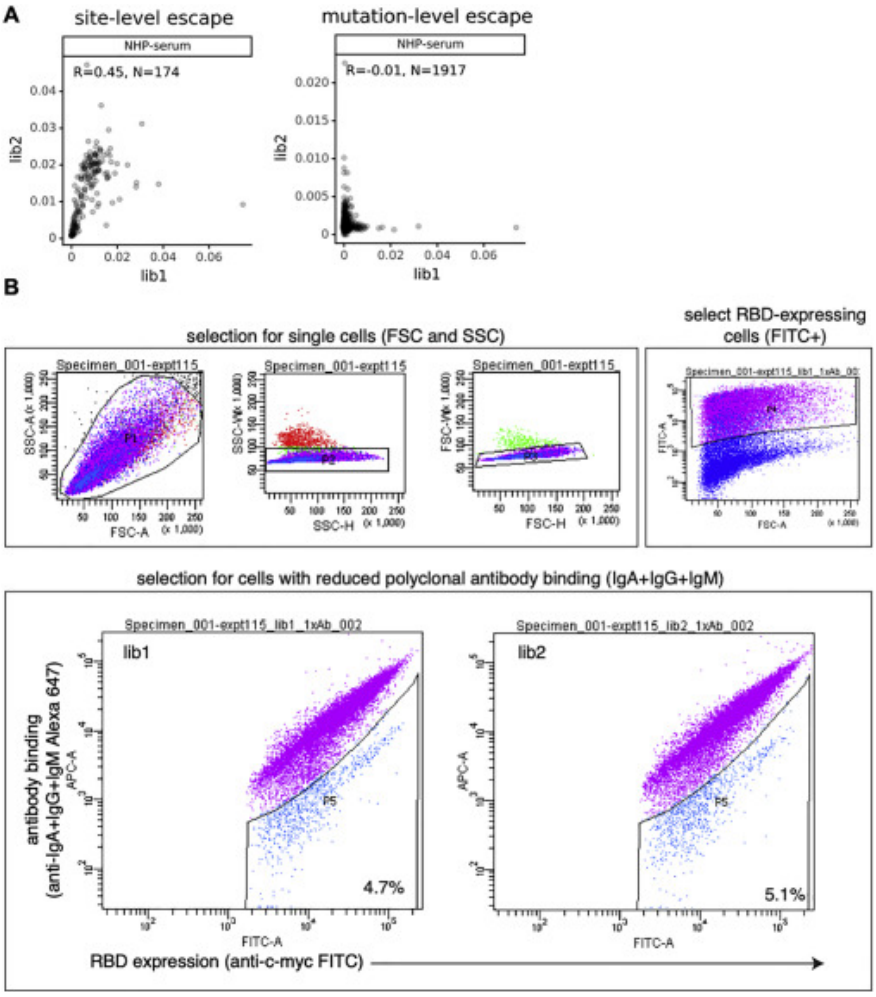


Figure 9.2: RBD-NP vaccination elicits high titers of Abs targeting diverse RBD antigenic sites in NHPs.

(A) NHP study designs.
(B) Effect of RBD mutations on polyclonal Ab binding measured by DMS analysis of serum obtained 8 weeks post-prime from an immunized pigtail macaque (n = 1) compared to a previously reported DMS measurement from a representative COVID-19 HCP sample (reproduced here for comparison; Greaney et al., 2021b). The line plots on the left show the summed effect of all mutations at a site in the RBD on serum or plasma binding, with RBD residues on the x axis and Ab escape on the y axis (Table S1). Due to the use of sample-specific FACS gates (**Fig. S9.1**), the y axes are scaled independently. Sites in the logo plots are colored dark blue if located in the receptor-

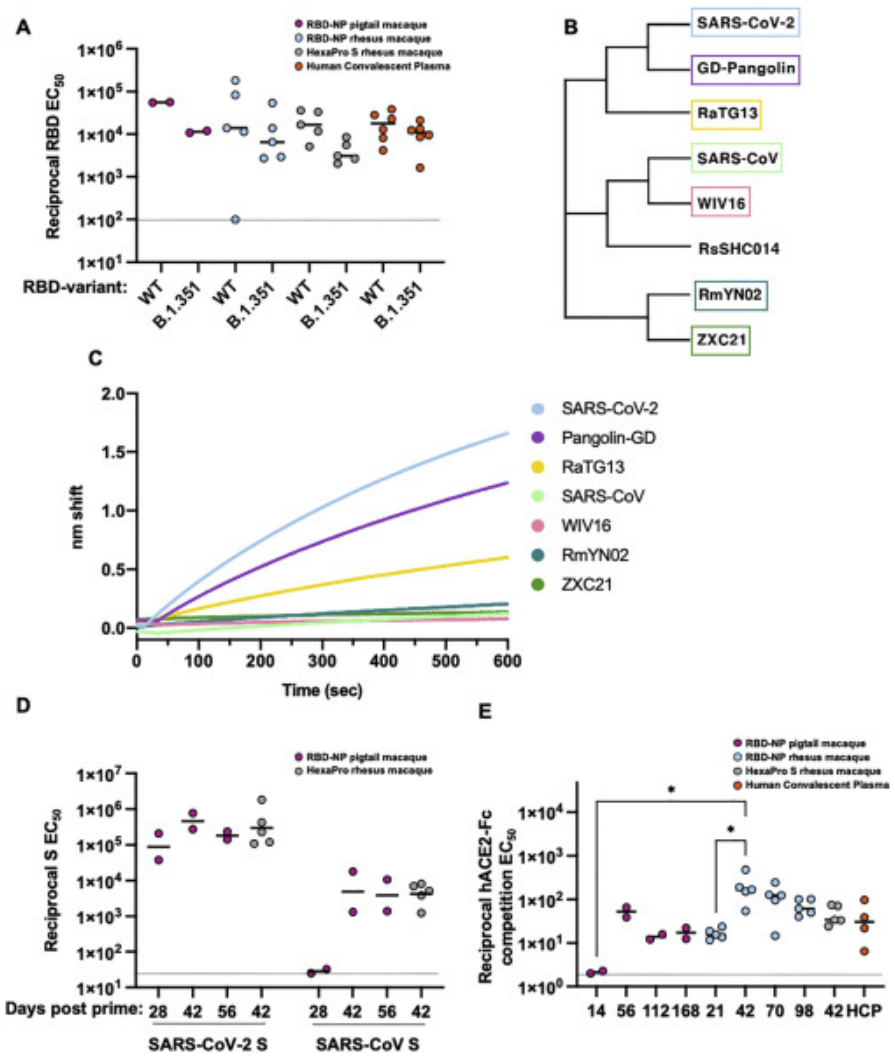
binding ridge or cyan if located in the RBD 443–450 loop. Larger values indicate more Ab escape.
(C–E) Competition ELISA between 0.2-nM hACE2-Fc (LOD of 2 × 10^{1}) (C), 2-nM CR3022 mAb (LOD of 2 × 10°) (D), or 0.01-nM S309 mAb (LOD of 4 × 10°) (E), and RBD-NP-elicited sera in pigtail macaques (n = 2), RBD-NP-elicited sera in rhesus macaques (n = 5), or HexaPro S-elicited sera in rhesus macaques (n = 5) at various time points following vaccination, benchmarked against HCP (n = 4) (**Table S9.2**). Each plot shows the magnitude of inhibition of hACE2/mAb binding to immobilized SARS-CoV-2 S2P, expressed as reciprocal serum dilution blocking 50% of the maximum binding response. Statistical significance was determined by Kruskal-Wallis test and shown only when significant. **p < 0.01. LODs are shown as gray horizontal dotted lines. Raw data curves shown in **Data S9.1**.



Supplemental Figure 9.1: Effects of mutations on binding of HCP Abs to RBD and FACS gating strategy, related to Fig. 2.

To measure the magnitude of vaccine-elicited Abs against distinct RBD antigenic sites, we used quantitative competition ELISAs with human ACE2-Fc (hACE2-Fc), which binds to the receptor-binding motif (RBM) corresponding to antigenic sites Ia and Ib, as

well as with structurally characterized mAbs recognizing antigenic sites II (CR3022) and IV (S309) (Huo et al., 2020; Piccoli et al., 2020; Pinto et al., 2020; Yuan et al., 2020) (**Figs. 2C–2E; Data S1**). These experiments used sera from the aforementioned pigtail macaques (RBD GMT EC_{50} 5×10^4) followed for 168 days and sera from 5 rhesus macaques immunized twice on days 0 and 21 with AS03-adjuvanted RBD-NPs containing 25 μg of the SARS-CoV-2 RBD (with a 16-GS linker between the RBD and the NP; RBD GMT EC_{50} 1×10^4) and followed for 98 days (Arunachalam et al., 2021; Walls et al., 2020b; **Figs. 2A and S2A**). These two sets of RBD-NP sera were compared to sera from 5 rhesus macaques immunized twice on days 0 and 21 with 12.5 μg of AS03-adjuvanted prefusion-stabilized HexaPro S trimer (RBD EC_{50} 1×10^4) and followed for 42 days (Arunachalam et al., 2021; Hsieh et al., 2020), as well as a panel of COVID-19 HCP (Table S2) (RBD EC_{50} 2×10^4) (**Figs. 2A and S2A**). All NHP sera post boost had high titers of Abs targeting sites Ia and Ib in the immunodominant RBM (day 42 hACE2 competition RBD-NP GMT 3×10^3 and HexaPro GMT 7×10^2) (**Fig. 2C; Data S1**), a correlate of neutralization potency (Piccoli et al., 2020), in agreement with the potent immunogenicity and protective efficacy of RBD-NP and HexaPro S in NHPs (Arunachalam et al., 2021). We also observed strong Ab responses against antigenic sites II (RBD-NP day 42 GMT 4×10^2 and HexaPro GMT 4×10^1) with CR3022 competition and IV (RBD-NP day 42 GMT 2.5×10^3 and HexaPro GMT 6.5×10^1) (**Figs. 2D and 2E; Data S1**), which comprise conserved sarbecovirus epitopes recognized by broadly neutralizing sarbecovirus mAbs such as S2X259 (site II) and S309 (site IV) (Piccoli et al., 2020; Pinto et al., 2020; Tortorici et al., 2021). Ab responses were durable against all 3 antigenic sites regardless of dose or adjuvant, with a decrease of competition titers by roughly half an order to one order of magnitude 168 days (pigtail macaque) or 98 days (rhesus macaque) post-prime. RBD-NP elicited much higher peak binding titers toward all RBD antigenic sites evaluated compared to HexaPro S and HCP, showcasing the potency, durability, and multi-specificity of Ab responses induced by multivalent display of the RBD, possibly as a result of enhanced accessibility of the antigen.



Supplemental Figure 9.2: Evaluation of vaccine-elicited binding and neutralizing Ab titers against SARS-CoV-2 variants and sarbecoviruses, related to Fig. 3.

(A) Wild-type (Wuhan-Hu-1) and B.1.351 SARS-CoV-2 RBD-specific Ab binding titers of RBD-NP-elicited sera in pigtail macaques (magenta, n = 2) and rhesus macaques (blue, n = 5), HexaPro S-elicited sera in rhesus macaques (gray, n = 5), or HCP (orange, n = 6, Table S1) analyzed by ELISA with an LOD of $1x10^2$. (B) Cladogram based on sarbecovirus RBD amino acid sequences. (C) Biolayer interferometry analysis of binding of 1 μM purified polyclonal pigtail macaque IgGs (obtained 70 days post prime) to sarbecovirus RBDs immobilized at the surface of biosensors. (D) SARS-CoV-2 S2P (left) or SARS-CoV S2P (right) Ab binding titers of RBD-NP-elicited sera in pigtail macaques (magenta) or HexaPro S-elicited sera in rhesus macaques (gray) analyzed by ELISA with an LOD of 2.5×10^1. (E) Competition ELISA between 0.13 nM human ACE2-Fc and RBD-NP-elicited sera in pigtail macaques (magenta) and rhesus macaques (blue), or HexaPro S-elicited sera in rhesus macaques (gray) at various time points following vaccination, benchmarked against COVID-19 HCP with an LOD of 4 × 10°, showing the magnitude of inhibition of ACE2 binding to immobilized SARS-CoV S2P expressed as reciprocal serum dilution blocking 50% of the maximum binding

response. Statistical significance was determined by Kruskal-Wallis test and shown when significant. **, $p < 0.01$. All data repeated twice. LODs are shown as gray horizontal dotted lines. Raw data curves shown in Data S1.

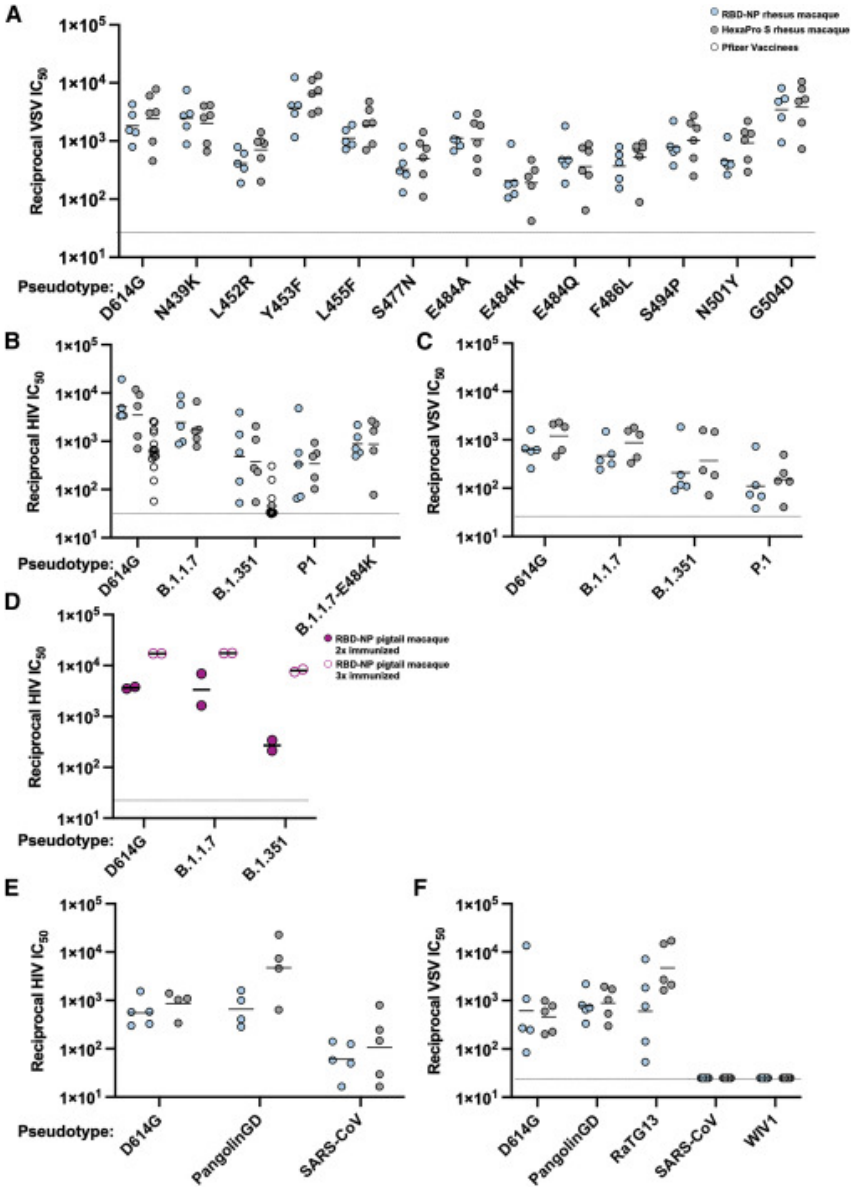


Figure 9.3: RBD-NP and HexaPro S elicit Abs with similar neutralization breadth toward SARS-CoV-2 variants.

(A) Neutralizing Ab titers against wild-type (D614G) SARS-CoV-2 S and RBD point mutants determined using RBD-NP-elicited sera in rhesus macaques (blue, n = 5) and HexaPro S-elicited sera in rhesus macaques (gray, n = 6) with an VSV pseudotyping system with an LOD of 3.3×10^1. Neutralization performed once.
(B) Neutralizing Ab titers against HIV pseudotyped viruses harboring wild-type (D614G) SARS-CoV-2 S, B.1.1.7 S, B.1.1.7-E484K S, B.1.351 S, or P.1 S, determined using RBD-NP-elicited sera in rhesus macaques (blue), HexaPro S-elicited sera in rhesus macaques (gray), or plasma from individuals who received two doses of Pfizer mRNA

vaccine (open circles) with an LOD of 3.3×10^1. Neutralization performed twice and a representative shown.
(C) Neutralizing Ab titers against VSV pseudotyped viruses harboring wild-type (D614G) SARS-CoV-2 S, B.1.1.7 S, B.1.351 S, or P.1 S, determined using RBD-NP-elicited sera in rhesus macaques (blue) or HexaPro S-elicited sera in rhesus macaques (gray) with an LOD of 2.5×10^1. Neutralization performed twice and a representative shown.
(D) Neutralizing Ab titers against D614G SARS-CoV-2 S, B.1.1.7, and B.1.351 S HIV pseudoviruses in pigtail macaque sera collected 28 days after a second (filled symbols, n = 2) or third (open symbols, n = 2) immunization with 88-μg RBD-NP (RBD antigen dose) with an LOD of 1×10^2. Neutralization performed twice and a representative shown.
(E) Neutralizing Ab titers against HIV pseudotyped viruses harboring wild-type (D614G) SARS-CoV-2 S, Pangolin-GD S, or SARS-CoV S, determined using RBD-NP-elicited sera in in rhesus macaques (blue) or HexaPro S-elicited sera in rhesus macaques (gray) with an LOD of 1×10^1. Neutralization performed twice and a representative shown.
(F) Neutralizing Ab titers against VSV pseudotyped viruses harboring wild-type (D614G) SARS-CoV-2 S, Pangolin-GD S, RaTG13 S, SARS-CoV S, or WIV1 S determined using SARS-CoV-2 RBD-NP-elicited sera in rhesus macaques (blue) or HexaPro S-elicited sera in rhesus macaques (gray) with an LOD of 2.5×10^1. Neutralization performed twice and a representative shown. Statistical significance was determined by Kruskal-Wallis test and shown in Table S3. LODs are shown as gray horizontal dotted lines. Raw data curves shown in Data S1 and GMTs in Table S4. The various pseudovirus backbones were benchmarked against NIBSC standard and are shown in Table S5.

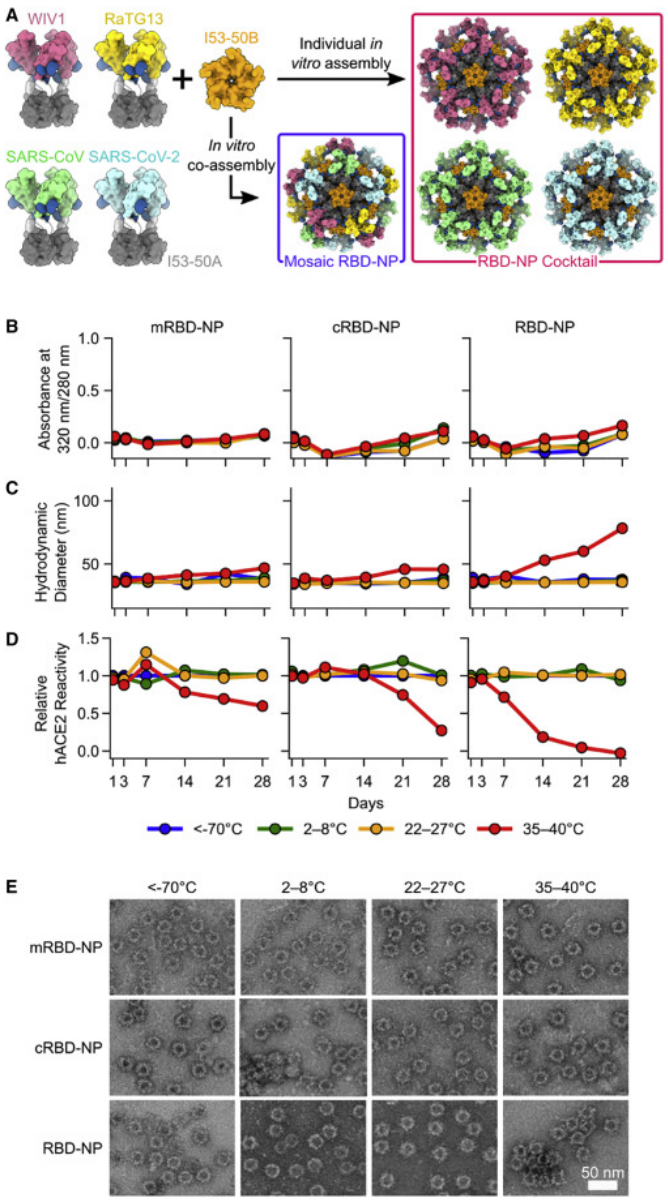
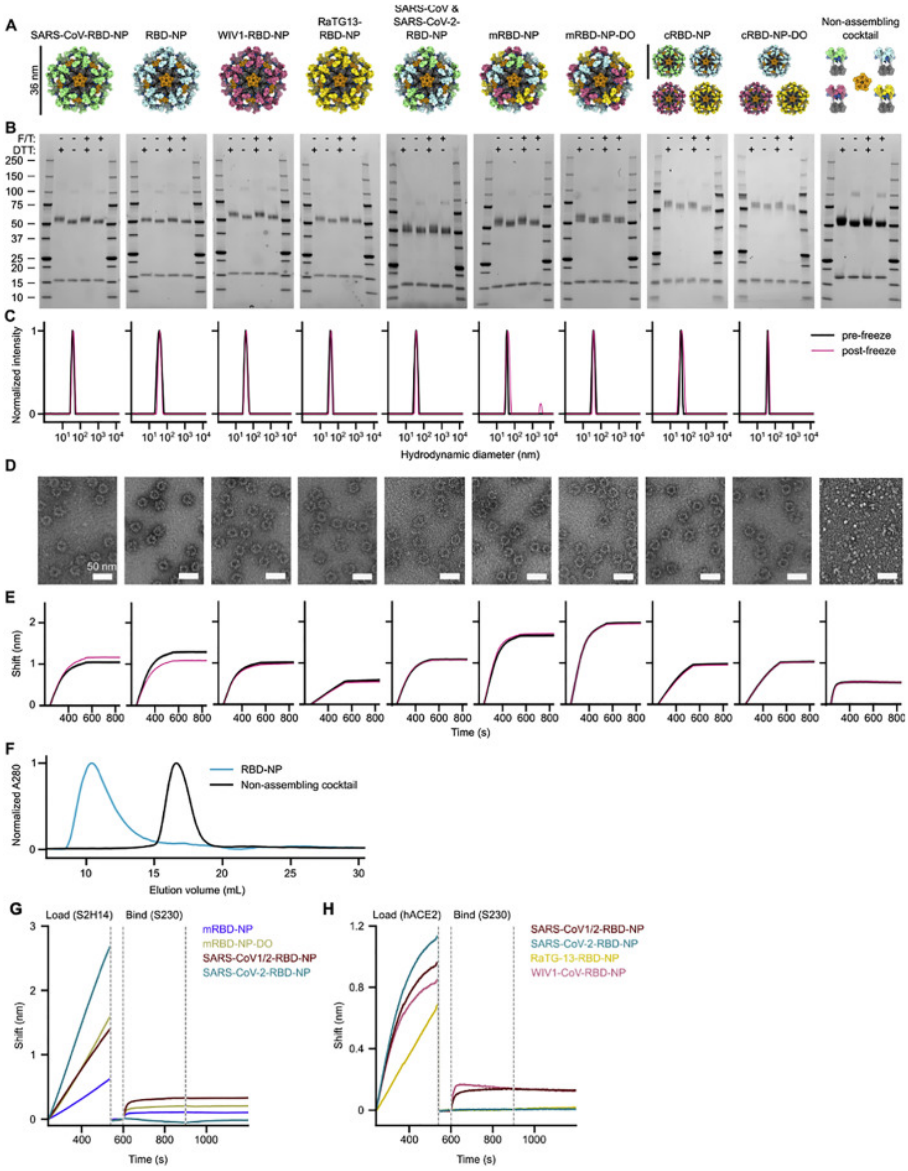


Figure 9.4: In vitro assembly and accelerated stability studies of mosaic and cocktail nanoparticle immunogens.

(A) Schematic of *in vitro* assembly of mRBD-NPs and cRBD-NPs.
(B–E) The physical and antigenic stability of mRBD-NP, cRBD-NP, and (SARS-CoV-2) RBD-NP samples incubated at four different temperatures was followed for four weeks.
(B) The ratio of UV/vis absorbance at 320 nm/280 nm is a measure of turbidity (proxy for aggregation).
(C) Hydrodynamic diameter of the nanoparticles measured using dynamic light scattering.
(D) hACE2-Fc binding measured by comparing the peak amplitude of hACE2-Fc binding for each sample to a reference sample stored at < −70°C using biolayer interferometry.

(E) Electron micrographs of negatively stained samples after incubation for 28 days at the indicated temperatures. Scale bar, 50 nm.



Supplemental Figure 9.3: In vitro characterization and confirmation of co-display of sarbecovirus RBD-NP immunogens, related to Fig. 4.

(A) Design models of the various vaccine candidates evaluated. Scale bars, 36 nm. (B) SDS-PAGE analysis of purified nanoparticles. DTT, dithiothreitol; F/T, freeze/thaw. (C) Dynamic light scattering. (D) Electron micrographs of negatively stained samples. Scale bars, 50 nm. (E) Binding of 100 nM SEC-purified nanoparticle immunogens and the non-assembling cocktail immunogen (which was not purified with SEC) to immobilized hACE2-Fc. (F) SEC chromatogram overlay of purified RBD-NP and non-assembling cocktail. (G-H) Sandwich biolayer interferometry. The SARS-CoV-2 S-specific mAb S2H14 immobilized on protein A biosensors was used to capture various nanoparticle immunogens from 300-480 s. The captured nanoparticles were subsequently exposed

to a Fab derived from the SARS-CoV S-specific mAb S230 from 600-900 s (G). hACE2-Fc immobilized on protein A biosensors was used to capture various nanoparticle immunogens from 300-480 s. The captured nanoparticles were subsequently exposed to a Fab derived from the SARS-CoV S-specific mAb S230 from 600-900 s (H).

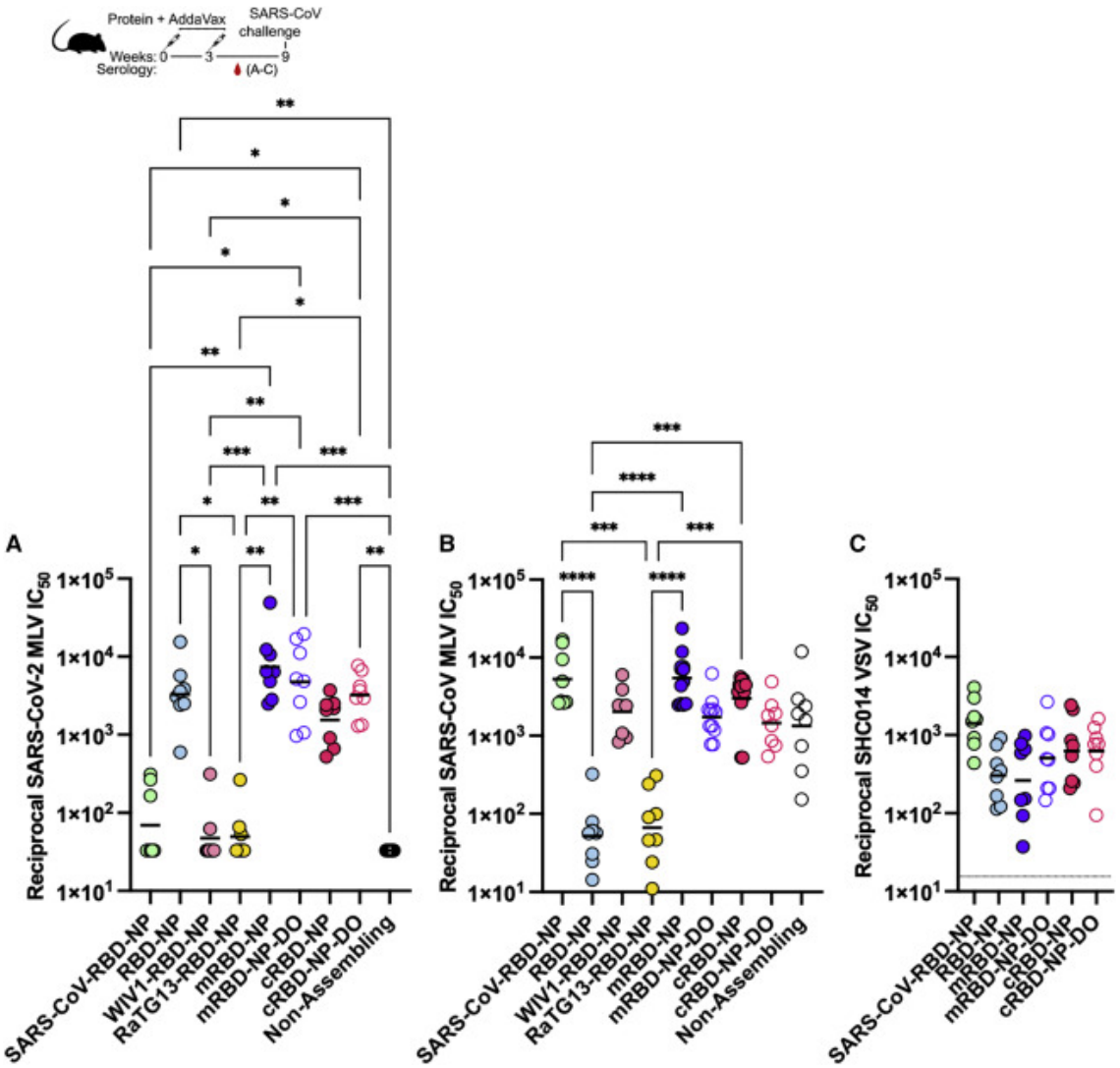


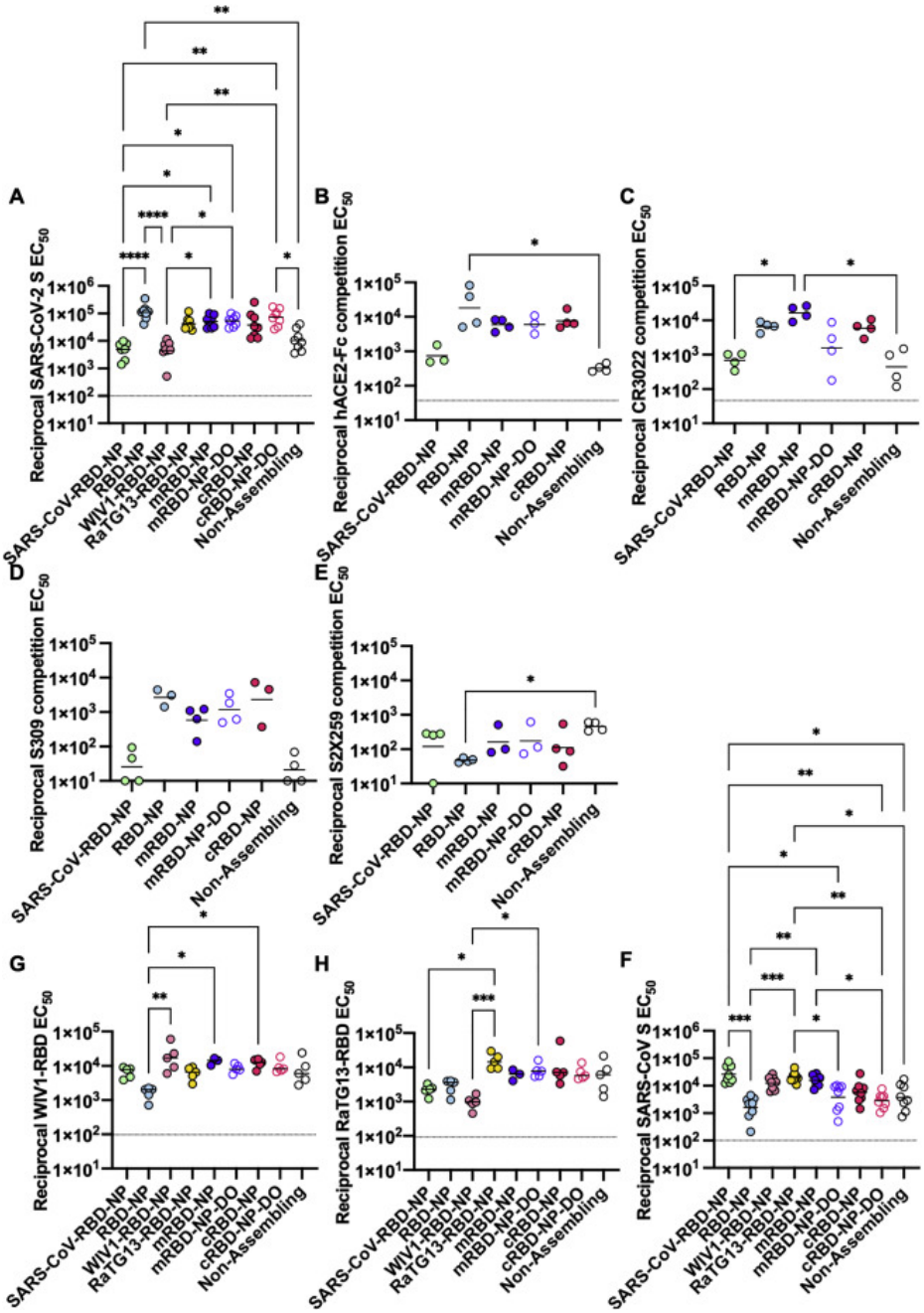
Figure 9.5: Mosaic and cocktail RBD-NP vaccines elicit neutralizing Abs against multiple sarbecoviruses.

(A) Neutralizing Ab titers in mice (n = 8) against wild-type (D614G) SARS-CoV-2 S MLV pseudovirus five weeks post-prime elicited by monovalent, mosaic, and cocktail RBD-NPs with an LOD of 1 × 10^1.

(B) Neutralizing Ab titers in mice against SARS-CoV S MLV pseudovirus five weeks post-prime elicited by monovalent, mosaic, and cocktail RBD-NPs with an LOD of 1 × 10^1.

(C) Neutralizing Ab titers in mice against SHC014 VSV pseudovirus five weeks post-prime elicited by monovalent, mosaic, and cocktail RBD-NPs with an LOD of 1.7 × 10^1.

Raw data curves shown in Data S1. Statistical significance was determined by Kruskal-Wallis test and shown only when significant. ∗∗p<0.01. LOD is shown as a gray horizontal dotted line in (C).



Supplemental Figure 9.4: Serum Ab binding titers elicited by mosaic and cocktail RBD-NPs, related to Fig. 5.

(A) Ab binding titers to SARS-CoV-2 S2P at five weeks post prime analyzed by ELISA with an LOD of $1x10^2$. (B–E) Titers of SARS-CoV-2 S-specific Abs competing with ACE2-Fc with an LOD of $5x10^1$ (B), CR3022 with an LOD of $5x10^1$ (C), S309 with an LOD of $1x10^1$ (D), and S2X259 with an LOD of $1x10^1$(E) in immunized mouse sera

analyzed by competition ELISA. (F) Ab binding titers to SARS-CoV S2P at week 5 analyzed by ELISA. (G–H) Ab binding titers to the WIV1 (G), and RaTG13 (H) RBDs analyzed by ELISA with an LOD of $1x10^2$. Statistical significance was determined by Kruskal Wallis test and shown when significant. $^{**}p < 0.01$. LODs are shown as gray horizontal dotted lines. Raw data curves shown in Data S1.

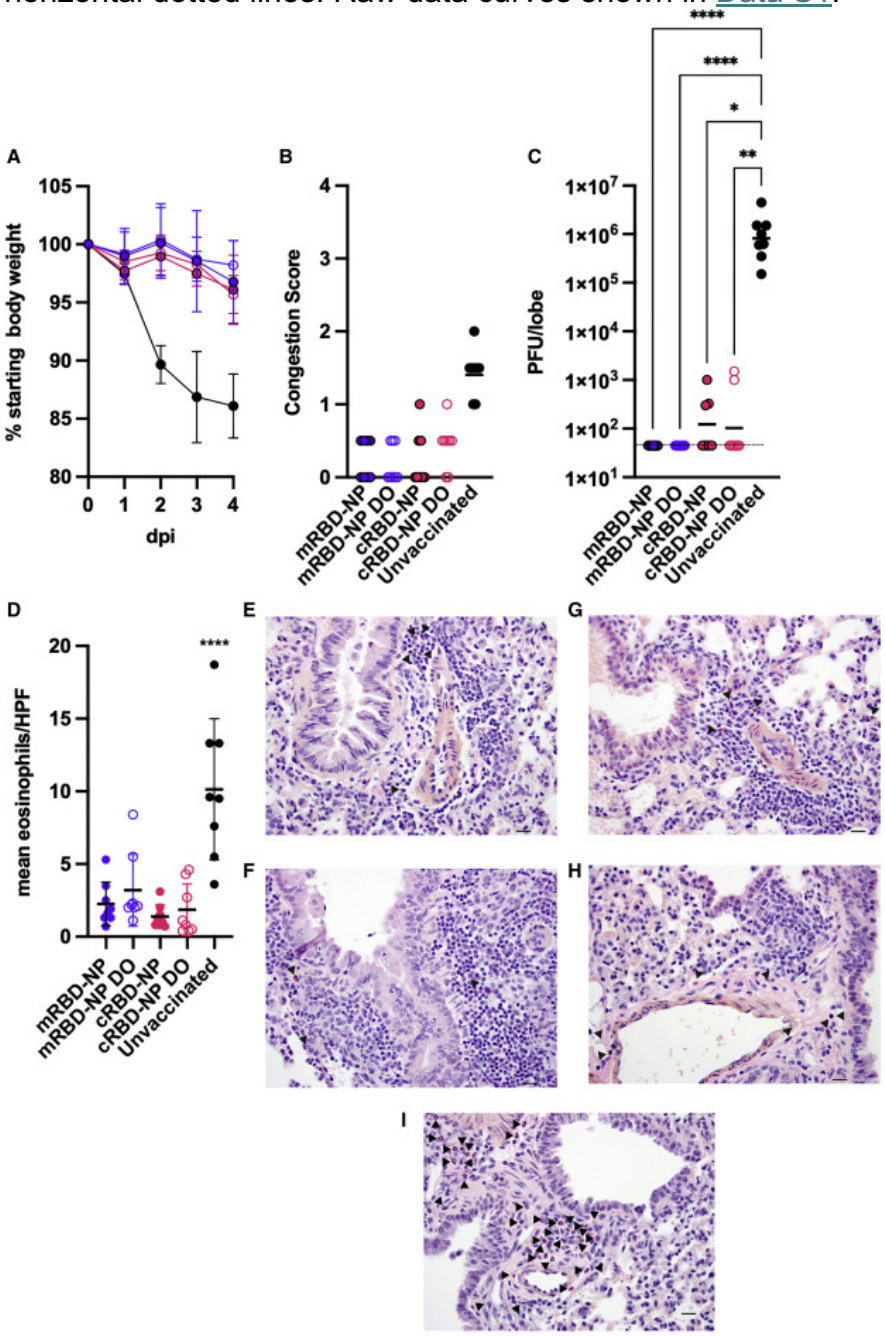
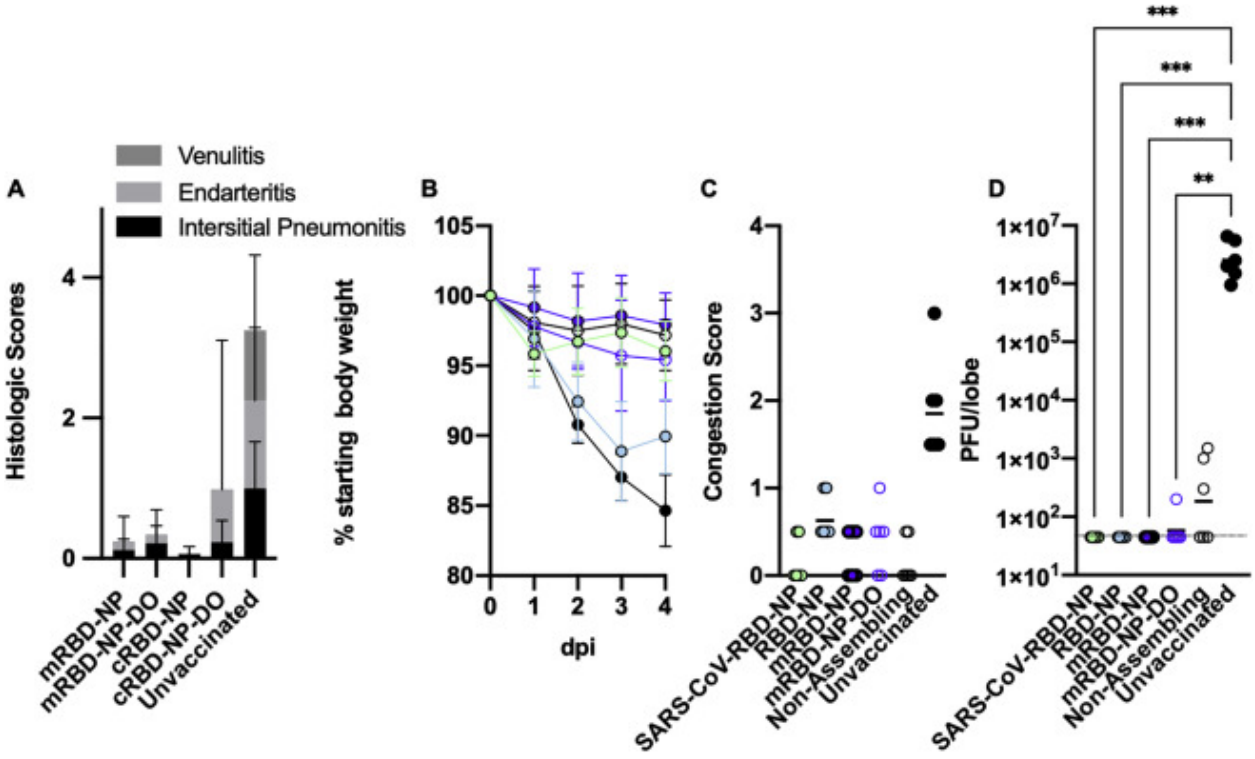


Figure 9.6: Mosaic and cocktail RBD-NP vaccines protect against heterotypic SARS-CoV-MA15 challenge in 15-week-old BALB/c cByJ mice.

(A) Weight loss following SARS-CoV MA15 challenge up to 4 days post infection (n = 8). Unvaccinated animals are shown as black circles.

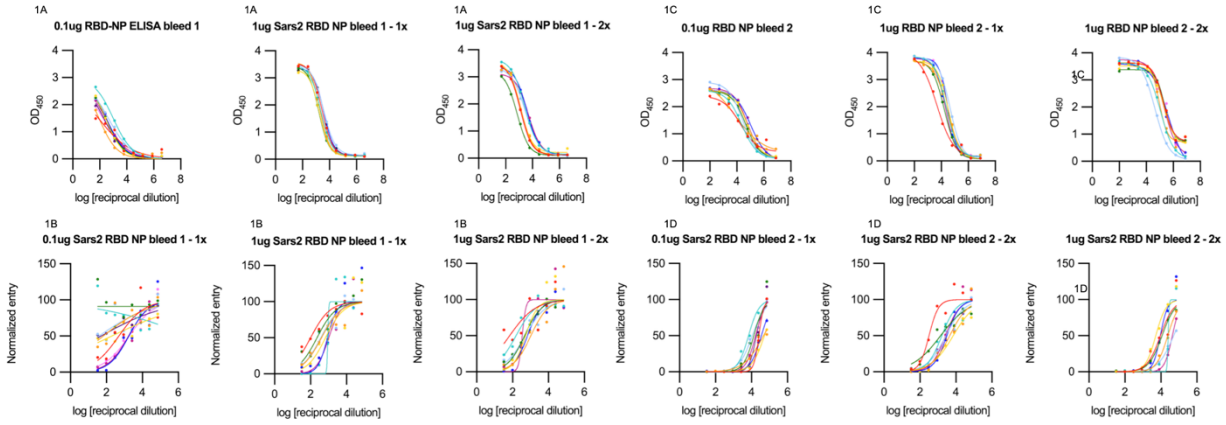
(B) Congestion score following SARS-CoV MA15 infection with a score of 0 indicating unchanged lung color and 4 indicating a darkened and diseased lung (n = 8).
(C) Viral titers in mice lungs (expressed in PFUs per lobe) following challenge (n = 8) with an LOD of 9×10^1. LOD is shown as a gray horizontal dotted line. Statistical significance was determined by Kruskal-Wallis and shown when significant. $^{**}p < 0.01$.
(D) Mean eosinophils per high power field (HPF) per sample run over 10 HPF per lung stained with congo red. Significance was determined using one-way ANOVA and shown where significant.
(E–I) Histological analysis of stained lung sections for mRBD-NP (E), mRBD-NP-DO (F), cRBD-NP (G), cRBD-NP-DO (H), and unvaccinated mice (I). Arrowheads indicate eosinophils. Scale bar, 20 μm.



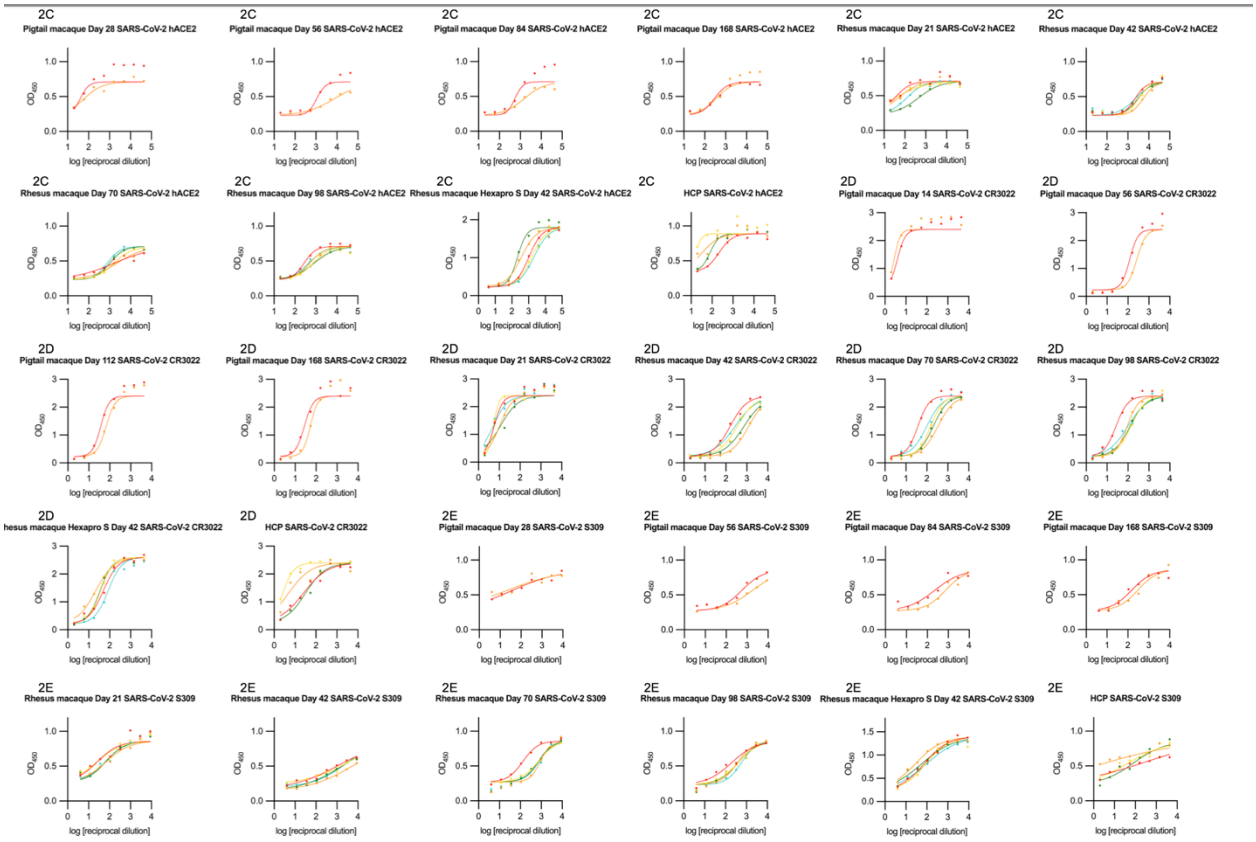
Supplemental Figure 9.5: Monovalent, mosaic, and cocktail RBD-NPs protect against heterotypic SARS-CoV-MA15 challenge in 15-week-old BALB/c cByJ mice, related to Fig. 6

(A) Normalized active inflammation following SARS-CoV MA 15 challenge shown in **Fig. 6** with venulitis, endarteritis, and interstitial pneumonitis shown as stacked bar graphs in dark gray, light gray, and black respectively. (B) Weight loss following SARS-CoV MA15 challenge (N = 6). Unvaccinated animals are shown as black circles. (C) Congestion score following SARS-CoV MA15 challenge with a score of 0 indicating unchanged lung color and 4 indicating a darkened and diseased lung (N = 6). (D) Viral titers in mice lungs (expressed in plaque forming units per lobe) following challenge (N = 8) with an LOD of $9x10^1$. Statistical significance was determined by Kruskal Wallis test and shown when significant and $^{**}p < 0.01$. LOD is shown as a gray horizontal dotted line.

Data 9.1: Raw data fits and curves of ELISA and Neutralization plots (related to Fig 1, Fig. 2, Fig. S2, Fig. 3, Fig. S4, Fig. S5)

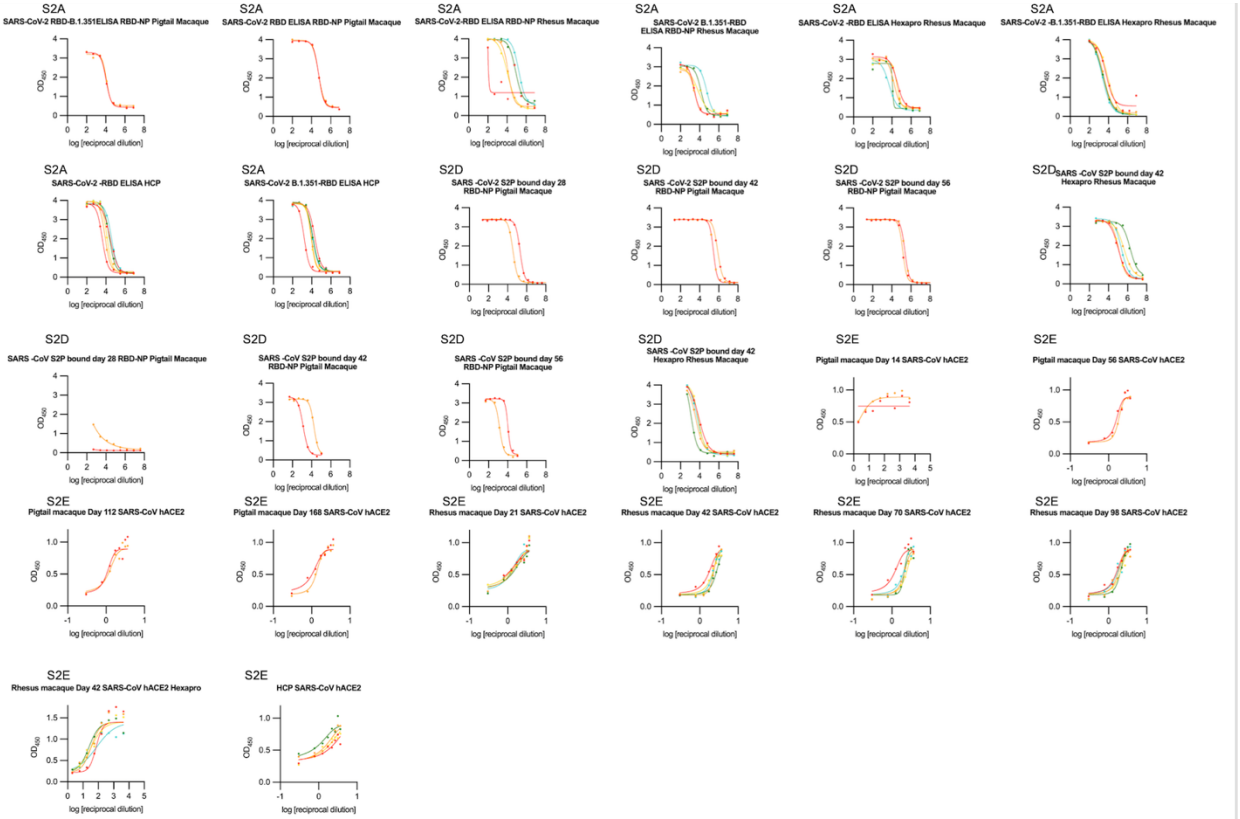


Associated curves with Fig. 1
Each colored curve represents a mouse sera sample

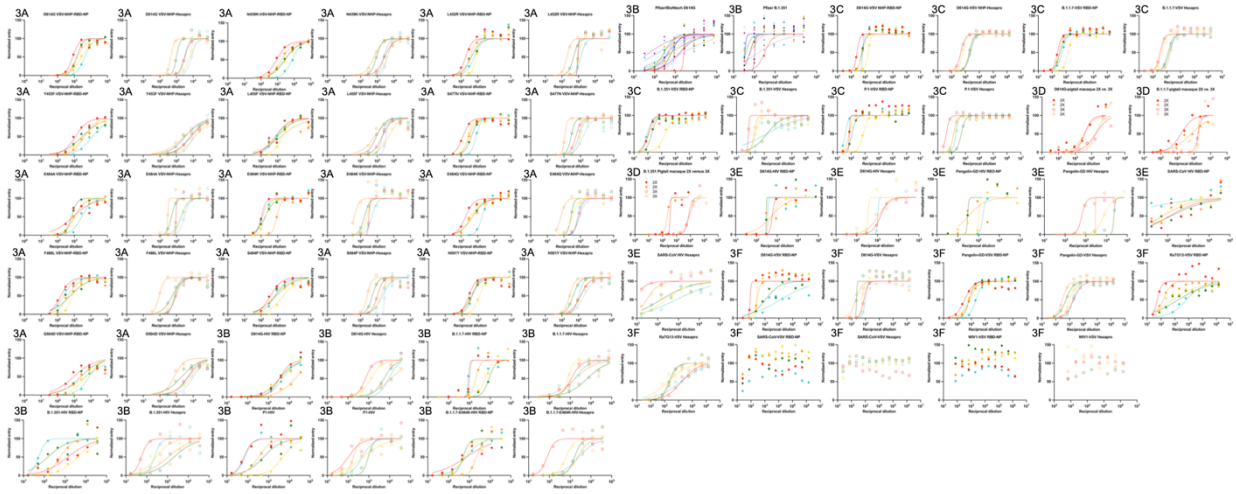


Associated curves with Fig. 2

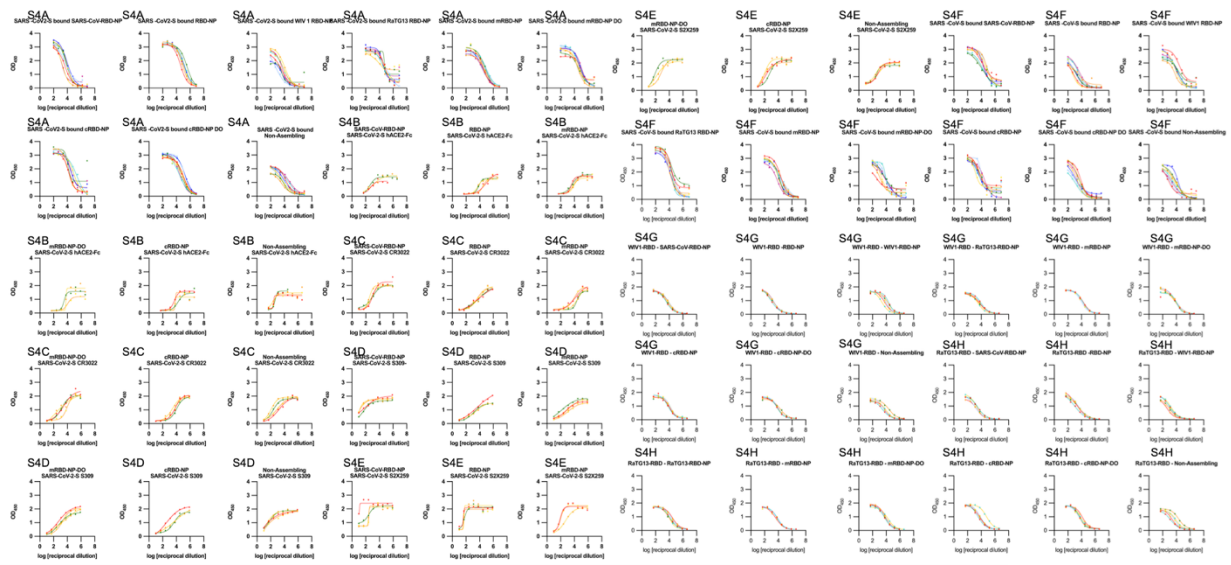
Each colored curve represents an NHP sera or human convalescent plasma sample



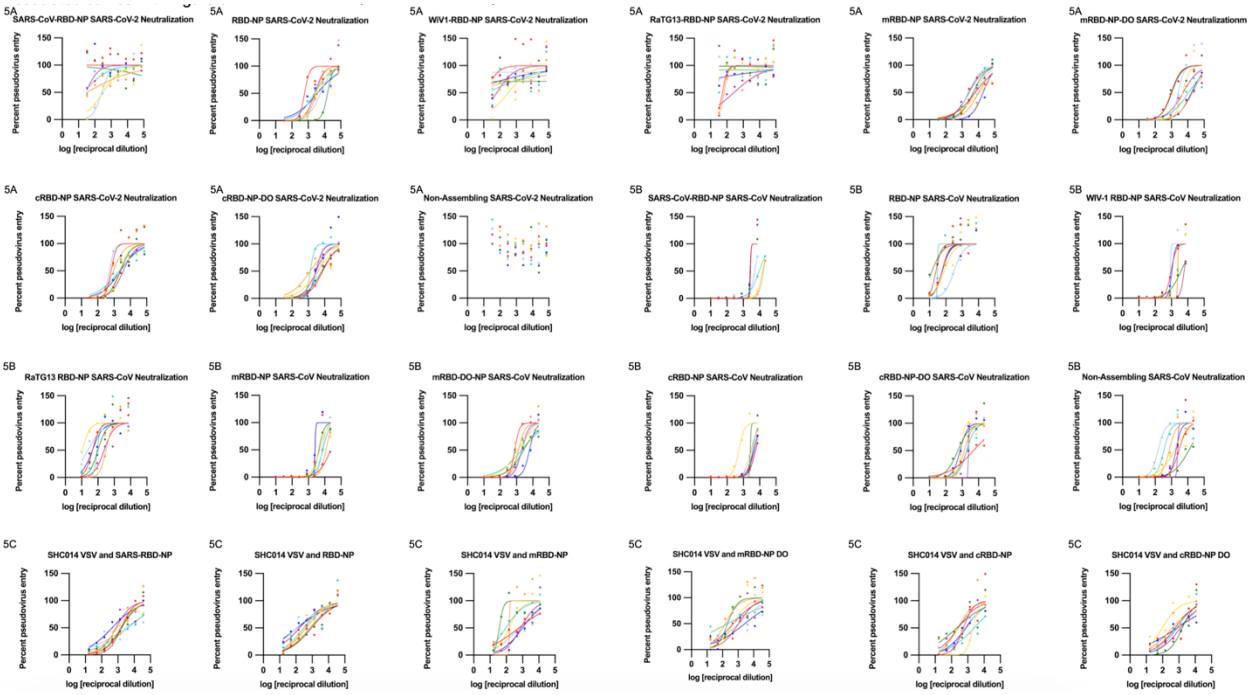
Associated curves with Fig. 3
Each colored curve represents an NHP (closed: RBD-NP, open: HexaPro) sera or human vaccinated plasma sample



Associated curves with Fig. S4
Each colored curve represents a mouse sera sample



Associated curves with Fig. 5
Each colored curve represents a mouse sera sample



CHAPTER 10. A POTENT PAN-SARBECOVIRUS NEUTRALIZING ANTIBODY RESILIENT TO EPITOPE DIVERSIFICATION.

SARS-CoV-2 evolution has resulted in viral escape from clinically authorized monoclonal antibodies (mAbs), creating a need for mAbs that are resilient to epitope diversification. In this story, we identified a human mAb, designated VIR-7229, which targets the viral receptor-binding motif (RBM) with unprecedented cross-reactivity to all sarbecovirus clades, including non-ACE2-utilizing bat sarbecoviruses, while potently neutralizing SARS-CoV-2 variants since 2019, including the recent EG.5, BA.2.86, and JN.1. In the subsequent sections, I detail VIR-7229's extraordinary toleration of epitope variability, partly attributed to its high binding affinity, receptor molecular mimicry, and interactions with RBM backbone atoms. Consequently, VIR-7229 features a high barrier for selection of escape mutants, which are rare and associated with reduced viral fitness, underscoring its potential to be resilient to future viral evolution. VIR-7229 is currently in development to become a next-generation COVID-19 medicine.

Adapted From: Rosen, L., Tororici, M.A., De Marco, A., Pinto, D., Foreman, W.B., Taylor, A.L., Bohan, D., Rietz, T., Errico, J.M., Hauser, K., Dang, H.V., Chartron, J.W., Giurdanella,M., Cusumano, G., Saliba, C., Zatta, F., Sprouse, K., Park, Y.P., Addetia, A., **Zepeda, S.K**., Brown, J, Dellota, E., Rajesh, A., Noack, J., Tao, Q., DaCosta, Y., Dias de Melo, G., Kergoat, L., Zhang, I., Liu, Z., Guarino, B., Schmid, M.A., Miller, J.L., Lempp, F.A., Czudnochowski, N., Cameroni, E., Whelan, S.P.J., Bourhy, H., Purcell, L.A., Benigni, F., di Iulio, J, Pizzuto, M.S., Lanzavecchia, A., Telenti, A., Snell, G., Corti, D., Veesler,D., Starr, T.N. (2024) A potent pan-sarbecovirus neutralizing

antibody resilient to epitope diversification. Cell.. doi: 10.1016/j.cell.2024.09.026. Online ahead of print. PMID: 39383863.

10.1 Introduction

Four and a half years after the emergence of SARS-CoV-2, the disease burden from COVID-19 remains high particularly for immunocompromised individuals and those at risk of severe disease (Antinori, 2023; Evans, 2023). Monoclonal antibodies (mAbs) have proven to be an important tool in the prevention and treatment of COVID-19 for these patient populations(Taylor, 2021; Corti, 2021; Suribhatla, 2023; Yang, 2022; Gentile, 2023), with millions of doses administered during the pandemic(US Department of Health and Human Services). However, mutations arising from continued SARS-CoV-2 evolution have abolished the activity of most approved therapeutic mAbs(Cameroni, 2022; Yang, 2023; Wang, 2023; Arora, 2023). Therefore, there remains a need for potent COVID-19 mAbs with durable activity in the face of continuous antigenic change.

One strategy for identifying mAbs resilient to viral evolution is to evaluate epitope conservation in distantly related viruses. The only therapeutic or prophylactic mAbs with activity against the majority of SARS-CoV-2 variants to-date are sotrovimab (Addetia, 2023; Stanford HVDB team, 2023), derived from the S309 mAb (Pinto, 2020), and pemivibart (VYD222)(Walker, 2024; FDA, 2024), which was affinity-matured from the ADI-55688 mAb(Corti, 2021; Rappazzo, 2021; Brandyn West). S309 and ADI-55688 were selected from the memory B cells of 2003 SARS-CoV-1 survivors based on their ability to recognize epitopes conserved between SARS-CoV-2 and SARS-CoV-1(Pinto, 2020; Rappazzo, 2021). Following the identification of S309 and ADI-55688, additional conserved SARS-CoV-2 spike (S) epitopes have been identified as targets of mAbs that

are broadly reactive with all sarbecoviruses(Starr, 2021; Tortorici, 2021), beta-coronaviruses(Pinto, 2021; Dacon, 2023), or even with multiple coronavirus genera(Low, 2022; Dacon, 2022). However, these highly-conserved epitopes are distinct from the ones targeted by the most potent mAbs, which usually compete directly with the ACE2 host receptor for binding to the SARS-CoV-2 S receptor-binding motif (RBM)(Tortorici, 2020; Nutalai, 2022). The RBM is one of the most variable S regions (**Fig. S1A**) due to the strong immune pressure at this site(Piccoli, 2020; Cao, 2022) and the plasticity of the binding interface between the receptor-binding domain (RBD) and ACE2(Cameroni, 2022; Starr, 2020; Taft, 2022; Thomson, 2021). As a result, RBM-targeting mAbs are generally SARS-CoV-2-specific(Park, 2022) and their neutralizing activity is frequently abolished due to viral evolution(Mykytyn, 2023). This apparent trade-off between neutralization breadth and potency has posed a major challenge for the clinical development of mAbs with broad neutralizing activity as their low neutralization potency would require using high doses. High doses leads to higher manufacturing costs, more frequent administration, and/or lengthy intravenous infusions, thereby reducing access and convenience for patients.

Here, we describe a highly potent RBM-targeting human mAb, designated VIR-7229, which has unprecedented activity across the entire family of sarbecoviruses, including neutralization of all SARS-CoV-2 variants that have arisen during the four and a half years since the start of the COVID-19 pandemic. VIR-7229 tolerates a remarkable sequence diversity in its epitope, and the rare escape mutations discovered in vitro are associated with reduced viral fitness, resulting in a high barrier for the emergence of

resistance. Our findings indicate VIR-7229 has a high probability to be resilient to future SARS-CoV-2 evolution, positioning it as a promising investigational COVID-19 medicine.

10.2 Isolation of a potent pan-sarbecovirus neutralizing mAb

To identify broadly reactive sarbecovirus mAbs, we interrogated with a high-throughput method the memory B cells of individuals who had received two to three doses of a Wuhan-Hu-1 S SARS-CoV-2 vaccine and who were subsequently infected by Omicron variants in 2022. A candidate mAb (S2V29) was selected based on its potent neutralization of pre- and post-Omicron SARS-CoV-2 variants, as well as for its cross-reactivity with a panel of sarbecovirus RBDs, including SARS-CoV-1. To further improve its cross-reactivity and neutralization potency, S2V29 was affinity-matured using both SARS-CoV-1 and SARS-CoV-2 BQ.1.1 RBDs as target antigens. Due to the urgency of the pandemic, including the deauthorization of most prophylactic and therapeutic mAbs, we desired to reach a lead sequence in the most efficient manner, and therefore utilized a machine learning (ML)-guided approach (Mason, 2021; Makowski, 2022a, Makowski, 2022a) combined with a yeast display system. This approach probes a larger number of antibody variants than possible with traditional affinity maturation(Khang) by sequencing every mAb variant in a training library after it has been sorted for binding affinity to the target antigens, and then using ML models to make predictions about mAb variants not present in the library. Additionally, evaluating multiple target antigens simultaneously rather than sequentially is anticipated to improve the efficiency and avoid bias towards a specific target. After two rounds of library screening and ML model training, a set of 56 mAb sequences were selected for recombinant production. These mAb variants were screened for binding and

neutralization of a wide panel of SARS-CoV-2 variants and other sarbecoviruses to select the mAbs with the highest breadth and potency.

Although the affinity maturation approach yielded candidate mAb sequences containing up to 11 mutations relative to the S2V29 mAb, one of the top neutralizers contained only two mutations (heavy chain V50Y and N57D) and was selected as the lead candidate (designated VIR-7229). Remarkably, VIR-7229 neutralized SARS-CoV-1 two orders of magnitude more potently than S2V29 (**Figs. 1A, S1B-C**). VIR-7229 competed with ACE2 for binding to the SARS-CoV-2 RBD (**Fig. S2A**) and neutralized a large panel of pre- and post-Omicron vesicular stomatitis virus (VSV)-based SARS-CoV-2 S pseudotyped viruses (pseudoviruses) with high potency (IC_{50} ranging between 1.8 and 435 ng/ml, median 7.3 ng/ml, for a 28-strain panel; **Fig. 1A** and **Data S1**), as well as representative VSV-based pseudoviruses selected from all ACE2-utilizing sarbecovirus clades (**Fig. 1A** and **Data S1**). Neutralization of authentic SARS-CoV-2 viral isolates was consistent with the high potency observed with pseudoviruses (IC_{50} 1.3-10.3 ng/ml, median 2.0 ng/ml, for a 10-strain panel; **Fig. 1B**).

To further investigate the breadth of this unique RBM-specific mAb, we evaluated its binding to a yeast-displayed library of RBDs encompassing 71 strains spanning the known sarbecovirus phylogenetic diversity (**Fig. 1C** and **Data S1**). Strikingly, VIR-7229 recognized every sarbecovirus RBD tested known to bind or enter cells using the human ACE2 receptor, as well as divergent bat-ACE2-utilizing sarbecoviruses (e.g. RsYN04, PRD-0038, and BtKY72), some of which can evolve to bind human ACE2 via single amino acid changes(Starr, 2022; Lee, 2023). Furthermore, VIR-7229 bound to all the divergent, non-ACE2-utilizing clade 2 sarbecovirus RBDs tested, which is an unprecedented result

due to the presence of two large deletions in clade 2 RBMs which typically disrupt recognition by SARS-CoV-2 RBM-targeting mAbs (Starr, 2021) (**Data S1**). The only three RBDs not recognized by VIR-7229 are viruses from Japan related to Rc-o319, which have a narrow host specificity for the geographically isolated *Rhinolophus cornutus* bat species(Murakami, 2022). Concurring with the yeast-display data, surface plasmon resonance (SPR) experiments showed that the VIR-7229 Fab fragment binds with sub-nanomolar affinity to most clade 1b (K_D 0.05-19 nM, median 0.29 nM, for a 14-strain panel of SARS-CoV-2 variants) and clade 3 RBDs and with nanomolar affinity to the other sarbecovirus RBDs (**Fig. 1D** and **Data S1**). To ensure that the extensive cross-reactivity of VIR-7229 was not due to binding promiscuity, we performed assessments of polyreactivity (Hep2 cells) and tissue cross-reactivity (immunohistochemical screening of 39 human tissues) with both S2V29 and VIR-7229, and observed no off-target binding for either mAb (**Data S2**).

To benchmark the breadth and potency of VIR-7229, we compared pseudovirus neutralization and RBD binding side-by-side with a panel of previously-described broadly reactive sarbecovirus mAbs(Pinto, 2020; Tortorici, 2021; Nutalai 2022; Park, 2022; Cao, 2022) (**Figs. 1C, 1E** and **S1D-J**). Similar to VIR-7229, S2X259 (antigenic site IIa) cross-reacted with members of all four sarbecovirus RBD clades although it recognized fewer clade 2 RBDs. Moreover, S2X259 had lower neutralization potency against SARS-CoV-2 Wuhan-Hu-1, relative to VIR-7229, and lost activity against recent Omicron variants. S309 (sotrovimab parent, antigenic site IV) cross-reacted with clade 1b and 1a RBDs, including the divergent bat ACE2-utilizing Rc-o319-related viruses, but not with clade 2 and 3 RBDs. The moderately-potent VYD222 mAb(Walker, 2024; FD, 2024) (antigenic

site Ia; renamed pemivibart and which received an emergency use authorization in the United States during the revision of this manuscript) cross-reacted with clade 1a RBDs, as well as two clade 3 RBDs tested, but did not cross-react with all clade 3 or with any clade 2 RBDs tested (**Figs. S1I-J**). The potent RBM-targeting (antigenic site Ia) mAbs S2K146, Omi-42, and SA55 revealed distinct patterns of cross-reactivity against sarbecoviruses. S2K146 is a unique example of an RBM-specific mAb isolated in the pre-Omicron era that retains activity, albeit reduced, against the majority of Omicron variants to date(Cao, 2022b), possibly due to its receptor molecular mimicry(Park, 2022). S2K146 cross-reacted with SARS-CoV-1 and some related clade 1a RBDs, as well as select clade 3 RBDs. Omi-42, which is in clinical development in the United States(Zhou, 2023), bound to human ACE2-utilizing clade 1b RBDs, neutralized the recently circulating JN.1 variant with moderate potency, and did not neutralize F456L-harboring SARS-CoV-2 XBB-lineage(Wang, 2023) or JN.1-lineage variants. Finally, SA55, in clinical development in China (SA55, 2023), broadly reacted with most ACE2-utilizing RBDs but did not cross-react with any of the clade 2 RBDs in our panel. Overall, the broad sarbecovirus reactivity of VIR-7229 is unique among potently neutralizing RBM-directed antibodies, establishing it as a best-in-class neutralizing mAb.

We next characterized the ability of VIR-7229 to promote S_1 shedding and Fc-mediated effector functions. As observed for other RBM-targeting mAbs(Piccoli, 202; Walls, 2019), VIR-7229 efficiently triggered S_1 shedding (**Fig. S2B**), a mechanism that can contribute to viral neutralization(Walls, 2019). Possibly due to S_1 shedding kinetics, VIR-7229 weakly activated FcγRIIa and FcγRIIIa in vitro using a Jurkat cell line-based reporter assay (**Figs. S2C** and **S2D**). Using primary human effector cells and highly

sensitive HiBiT target cells, we observed only moderate antibody-dependent cell cytotoxicity (ADCC) (**Fig. S2E**). These findings suggest that the main mechanisms of VIR-7229 antiviral activity are ACE2 competition and possibly S_1 shedding, with only a minor potential contribution of effector function in vivo.

To investigate whether the potent VIR-7229-mediated in vitro neutralization translates into effective in vivo protection, we evaluated mAb prophylactic activity using a Syrian hamster model of infection. Animals were administered intraperitoneally with VIR-7229-hm-Fc (VIR-7229 with a species-matched Fc) at various doses one day before challenge with SARS-CoV-2 XBB.1.5 or JN.1. VIR-7229 reduced viral RNA and infectious viral titers in the lungs of challenged animals in a dose-dependent manner, with the highest dose (1.5 mg/kg and 5 mg/kg, for XBB.1.5 and JN.1 challenge, respectively) resulting in infectious virus titers below the limit of detection (**Fig. 2** and **Data S3**). Furthermore, VIR-7229 administration protected the XBB.1.5-challenged animals from weight loss (**Fig. 2** and **Data S3**; JN.1 infection induced mild clinical symptoms in hamsters and none of the JN.1-challenged animals experienced weight loss at day 4). Overall, VIR-7229 is endowed with broad sarbecovirus cross-reactivity, potent neutralizing activity and protective prophylactic efficacy that are collectively unparalleled among previously characterized COVID-19 mAbs.

10.3 Structural basis of VIR-7229 breadth and potency

To understand the molecular basis of the remarkable VIR-7229 breadth, we determined a cryoEM structure of the VIR-7229 Fab fragment bound to the BA.2.86 S ectodomain trimer at an overall resolution of 2.9 Å. Subsequent local refinement of the VIR-7229 Fab variable domains and the BA.2.86 RBD yielded a reconstruction at 3.3 Å

resolution (**Figs. 3A** and **S3** and **Table S1**). Moreover, we determined crystal structures of the VIR-7229 Fab fragment bound to the XBB.1.5 and to the EG.5 (XBB.1.5 + F456L) RBDs at 2.4 and 1.95 Å resolution, respectively, as well as of the parent S2V29 Fab bound to the BQ.1.1 RBD at 1.67 Å resolution (**Fig. 3B-G** and **Table S2**). VIR-7229 recognizes an epitope in the RBD antigenic site Ia (Piccolim 2020), which overlaps with the ACE2-binding site (i.e. the RBM), burying an average of 950 $Å^2$ at the interface between the epitope and the paratope (**Fig. 3A-B)**. VIR-7229 interacts with the RBD via polar interactions and shape complementarity mediated by all six CDR loops, with the heavy chain CDR3 (H3) dominating the paratope. The epitope comprises amino acid residues 403, 405, 409, 415-417, 420-421, 453-460, 473-477, 487, 489, 493 and 505 (**Fig. 3C**). Thirteen out of these 25 residues participate in binding to human ACE2, explaining the competition observed for receptor engagement (**Figs. S2A** and **S4A**)**.** Strikingly, VIR-7229 binding induces a rearrangement of RBD residues 473-489 (**Fig. S4B**), which are shifted approximately 5.5 Å relative to structures of apo S or of the RBD in complex with ACE2 (Addetia, 2023; Nutalai, 2022, Wrobel, 2020, Lan, 2020).

VIR-7229 CDR H3 forms extensive contacts with the RBD, burying ~470 $Å^2$ of its surface at the interface with SARS-CoV-2 RBD residues 415-417, 420-421, 454-460, 473, 489 and 493 (**Fig. 3D-E**). Five out of 16 hydrogen bonds between VIR-7229 and the RBD involve the RBD backbone rather than amino acid side chains (RBD residues N417, L455, R457, K458), which may contribute to the resilience of VIR-7229 to epitope diversification (**Fig. 3E**). The hydrogen bonds formed by the RBD-L455 backbone carbonyl oxygen with H3-Y108 side chain hydroxyl and by the RBD-R457 backbone amide and carbonyl oxygen with H3-L104 backbone carbonyl oxygen and amide, respectively, are examples

of these contacts. Furthermore, comparison of the VIR-7229-bound XBB.1.5 and EG.5 structures shows that the VIR-7229 binding mode enables H3 to accommodate equally effectively F456 or L456 in these variant backgrounds (**Fig. 3F**), the latter mutation mediating immune evasion of many mAbs targeting antigenic site Ia(Jian, 2023). Molecular dynamics (MD) simulations performed on both the VIR-7229:XBB.1.5 RBD and the VIR-7229:EG.5 RBD structures (total simulation time 4.0 μs for each) indicated that residue 456 is one of the RBD positions (along with 415) with which VIR-7229 makes the largest number of persistent contacts for both F456 (XBB.1.5) and L456 (EG.5) (**Fig. S4C** and **Data S4**). The affinity maturation of S2V29 to VIR-7229 selected for two amino acid changes in CDR H2 (V50Y and N57D) and resulted in marked improvement of neutralization potency against SARS-CoV-1 (**Fig. S1B**), which harbors a leucine at the position equivalent to SARS-CoV-2 residue 456 (**Fig. 4A**). Comparison of the S2V29-bound and VIR-7229-bound RBD structures suggests that the N57D substitution allows formation of a salt bridge with RBD K460 (clearly resolved in the EG.5 RBD structure), most likely strengthening binding (**Fig. 3G**). Although CDR H2 residue 50 does not directly contact the RBD, the V50Y substitution leads to reorientation of the neighboring H3-Y106 side chain and formation of T-shaped pi stacking interactions between the two aromatic rings, possibly preconfiguring CDR H3 for binding (**Fig. 3G**).

The above structural findings likely explain the overall resilience of VIR-7229 to mutations observed in circulating strains at RBD residues 455 and 456 (**Figs. 1A, 1B** and **1D**), which individually or jointly appeared in descendant lineages of XBB.1.5 (i.e. HK.3) and BA.2.86 (i.e. JN.1) and have dampened the neutralizing activity of polyclonal plasma antibodies in individuals exposed to XBB.1.5(Jian, 2023; Yang, 2023; Tortorici, 2024).

Furthermore, our structural data explain the potent VIR-7229-mediated neutralization of BA.2.86 and JN.1; the latter variant is the parental lineage for the dominant currently circulating strains. Out of the 11 mutated residues in the BA.2.86 RBD relative to XBB.1.5, only R403K is found in the VIR-7229 epitope, a substitution which preserves electrostatic interactions with the VIR-7229 light chain N33 and D52 amide and carboxylate side chains, respectively, as observed in our cryoEM structure (**Fig. 3H**). The JN.1 variant harbors the immune-evasive L455S mutation relative to BA.2.86 that is also compatible with the VIR-7229 paratope interface due to the small size of the introduced serine side chain, concurring with the preserved binding and neutralization of BA.2.86 and JN.1 variants by VIR-7229 (**Figs. 1A, 1B** and **1D**).

The footprint of Omi-42 largely overlaps with that of VIR-7229 (24 residues are shared between VIR-7229 and Omi-42 out of 25 and 26 epitope residues, respectively, **Fig. S4D**)(Nutalai, 2022) and both mAbs bury a comparable surface area at the interface with the RBD. The more extensive hydrogen-bonding network of VIR-7229 with the RBD, relative to Omi-42, might explain its increased cross-reactivity and neutralization breadth (**Fig. 1C** and **S1H**). VIR-7229 forms four hydrogen bonds with the backbone of residues 455, 457, and 458, whereas Omi-42 forms a single one with backbone atoms in this RBD region. These results possibly explain the reduced neutralizing activity of Omi-42 for the JN.1 strain harboring the L455S mutation, and the markedly reduced neutralizing activity for XBB-descendant and JN.1-descendant variants harboring the F456L mutation(Wang, 2023) (**Fig. S1H**). Conversely, VIR-7229 neutralizes XBB-descendant F456L variants and JN.1 with high potency, and the F456L-harboring JN.1.16 variant with moderate potency (**Fig. 1A**). Furthermore, RBD residue K458 is hydrogen-bonded via its backbone carbonyl

to the VIR-7229 heavy chain Y53 side chain whereas it is the side chain of K458 that interacts with the Omi-42 heavy chain D31 and W53 side chains. Therefore, K458 mutations (observed in clades 1a and 3 sarbecoviruses) could impair interactions with Omi-42 but not with VIR-7229 and may explain the limited sarbecovirus cross-reactivity of Omi-42.

10.4 Structural basis of VIR-7229 breadth across animal sarbecoviruses

Our structural data explain the broad VIR-7229-cross-reactivity with phylogenetically distinct sarbecovirus RBDs spanning all four defined clades (**Fig. 4**). One part of the VIR-7229 epitope is highly conserved across sarbecoviruses and maps mainly outside of the RBM (**Fig. 4B,** dark orange). However, the VIR-7229 epitope also comprises residues with considerable variation among sarbecovirus RBDs: some substitutions introduce residues of similar size and properties compared to those found in SARS-CoV-2 whereas other mutations introduce distinct residues that are nevertheless accommodated by VIR-7229 (**Figs. 1A, 1C-D** and **4**). RsSHC014, the most weakly neutralized bat sarbecovirus in our panel, harbors W455 which is expected to disrupt the interface with VIR-7229, as none of the energetically favored side chain rotamers at this position can be accommodated without steric hindrance with surrounding residues (**Fig. S4E**). Furthermore, the deletion of residues 473-477 in the RBD of the non-ACE2-utilizing bat clade 2 sarbecoviruses would reduce interactions with the heavy chain CDR1 and CDR3, leading to a reduction of epitope buried surface area of ~180 $Å^2$, in line with the dampened VIR-7229 binding observed by both flow cytometry analysis of yeast-displayed RBDs and by SPR (**Figs. 1C-D**).

Collectively, our data show that although some VIR-7229 epitope residues are mutational hotspots for SARS-CoV-2 or are positions substituted in distinct sarbecoviruses, these substitutions are accommodated by VIR-7229, illustrating a high mutational tolerance of this RBM-targeting mAb.

10.5 VIR-7229 has an unusually high barrier to viral escape

To investigate the potential for viral resistance to VIR-7229, we exhaustively mapped its escape profile using deep mutational scanning (DMS) of the Wuhan-Hu-1, BA.2, BQ.1.1, XBB.1.5, EG.5, and BA.2.86 yeast-displayed RBDs. VIR-7229 featured a remarkably narrow escape profile (**Figs. 5A** and **S5A-C**) as compared to its parent mAb S2V29 (**Fig. S5C**) and even more so relative to Omi-42, SA55, and S2K146 which we evaluated side-by-side (**Fig. 5B-C** and **S5D**) or compared to published SARS-CoV-2 mAb DMS profiles(Starr, 2021; Tortorici, 2021; Cao, 2022; Starr, 2023). These results concur with the sub-nanomolar binding affinity of VIR-7229 to most SARS-CoV-2 variant RBDs (**Fig. 1D**) and our previous observation that binding affinity correlates with escape profile width(Starr, 2021). Concurring with our structural analysis, VIR-7229 is largely unaffected by mutations at position K458, whereas Omi-42 binding is abrogated by several substitutions at this position in multiple backgrounds (**Fig. 5B**). Many of these mutations correspond to residue changes found in sarbecoviruses, such as K458H found in clade 1a and clade 3 RBDs or K458S/A which is present in some clade 2 RBDs, consistent with the limited sarbecovirus breadth of Omi-42 relative to VIR-7229 despite overlapping epitopes.

Most VIR-7229 DMS escape mutations mapped to the RBD position 456, which was the only position of escape for the BQ.1.1, XBB.1.5, EG.5, and BA.2.86 RBDs, with

the exception of Y421W observed in the EG.5 background. All VIR-7229 DMS escape mutants reduced ACE2-binding affinity (yellow/orange letters in **Fig. 5A**; see also **Fig. S6A**) and those observed in the most recent variant backgrounds (BQ.1.1, XBB.1.5, EG.5, BA.2.86) require 2-3 nucleotide mutations from the wildtype codon, with the exception of P456 which is only 1 nucleotide mutation away from the recently-circulating L456 (**Fig. S6A**). None of these mutations are found in SARS-CoV-2 sequenced genomes (each mutation has at most one occurrence in the GISAID database, which is below the anticipated rate of error for deposited sequences), likely due to the high barrier to sampling as well as to the reduced fitness (**Fig. S6A**). The importance of residue 456 for viral fitness was confirmed by a >4 μs MD simulation of the XBB.1.5 RBD:ACE2 complex, revealing that residue 456 is one of the RBM positions with which ACE2 makes multiple persistent contacts (**Figs. 5D** and **S4F** and **Data S4**). This observation explains the large reduction of RBD:ACE2 binding affinity resulting from non-conservative substitutions at position 456 (**Figs. 5A** and **S6A**). Residue 456 is also a key Omi-42 DMS escape position (**Fig. 5B**) with many more amino acid substitutions at that position impacting binding relative to VIR-7229, including several mutations, such as F456L, which do not have a significant impact on RBD:ACE2 binding affinity (dark red letters in **Fig. 5B**).

To directly evaluate viral escape from VIR-7229, we used replicating VSV (rVSV) chimeras harboring SARS-CoV-2 variant S glycoproteins instead of the endogenous VSV G (**Fig. 6A** and **Data S5**). These experiments were performed with VIR-7229 alongside SA55 and Omi-42 mAbs as benchmarks. A single round of passaging was sufficient to select viral escapes for the SA55 mAb (G504D in the XBB.1.5, EG.5, and XBB.1.5.70 S

backgrounds) and Omi-42 (F456L in the XBB.1.5 S background) (**Fig. 6A**); a previous study showed similar escape from S2K146 via the Y489H mutation in the Wuhan-Hu-1 S background(Park, 2022). In all cases, these escape mutations concur with our DMS data (**Figs. 5B-C** and **S5D**). In contrast, we could not isolate any VIR-7229 escape mutants for the Wuhan-Hu-1 and XBB.1.5 S backgrounds after ten and seven rounds of serial passaging, respectively. These results were consistent with orthogonal plaque-based selection assays with BQ.1.1 S and XBB.1 S rVSV, with which we selected several escape mutants for the SA55 mAb but none for VIR-7229 (**Fig. 6B** and **Data S5**). We observed escape from VIR-7229 only with the EG.5 S or XBB.1.5.70 S rVSV after two or three rounds of serial passaging, leading to the emergence of the L455W mutation combined with R357I or T415I (EG.5 S) or of the D420N mutation (XBB.1.5.70) (**Fig. 6A**). The results obtained with EG.5 S concur with the reduced binding and neutralization of the RsSHC014 S pseudovirus (**Figs. 1A** and **1C**), which harbors W455 at the equivalent RBD residue position.

To validate the DMS and serial passaging results, we evaluated VIR-7229-mediated neutralization of a large panel of SARS-CoV-2 S pseudovirus mutants (**Figs. 5E-F** and **S6A-C** and **Data S1**). All epitope substitutions which appear in the GISAID database with >0.005% prevalence (as of May 8, 2024) were potently neutralized by VIR-7229 when introduced as single mutants in the XBB.1.5 and JN.1 backgrounds or when tested in the context of a circulating variant harboring that mutation (**Figs. 1A, 5F** and **Data S1**), underscoring the resilience of this mAb to epitope mutations found in circulating variants (**Figs. 5E** and **5G**). We observed a complete or near-complete loss of neutralization with the G416L, F456E, F456P, F456K and F456R mutations in all S

backgrounds tested (**Fig. S6A**; each mutation has a maximum of 2 occurrences in GISAID), all of which severely reduce ACE2 binding affinity and have a notable defect in pseudovirus infectivity (**Fig. S6A**). The effect of several other mutations on neutralizing activity was dependent on the S background in which they were evaluated. For instance, A475N promoted full neutralization escape in the Wuhan-Hu-1 background where it creates a new glycosylation site (due to the presence of S477), but not in an Omicron background (given the S477N mutation abrogating the glycosylation sequon). S459P promoted full or partial escape from VIR-7229-mediated neutralization in early-Omicron (BA.2 and BA.5) but not in later-Omicron (BQ.1.1 and XBB.1.5) S backgrounds (**Fig. S6A** and **Data S1**). These results might be explained by remodeling of the putative BA.2/BA.5 RBD-N460/VL-Y97 hydrogen bond to an RBD-K460/VH-D57 salt bridge (the latter interaction being stronger and observed in the VIR-7229-bound EG.5 structure). L455W did not promote neutralization escape in the BQ.1.1, XBB.1.5, or BA.2.86/JN.1 S backgrounds, but led to reduced neutralization in the BA.5 and EG.5 S backgrounds, likely due to the presence of the N460 or F456L mutations, respectively (**Fig. S6B** and **Data S1**). Consistent with the serial passaging results, either the R357I or T415I mutations in combination with L455W were required to promote complete escape from VIR-7229-mediated neutralization or binding in the EG.5 S background (neutralization IC_{50}: EG.5-L455W 236 ng/ml, EG.5-L455W/R357I >1250 ng/ml, EG.5-L455W/T415I >1250 ng/ml; binding affinity: EG.5-L455W K_D 160 nM; **Figs. 6C** and **S6B**). Although D420N modestly attenuated VIR-7229 potency in the BQ.1.1, XBB.1.5, EG.5, and JN.1 S backgrounds (~2-7-fold reduction), it had a larger impact in the XBB.1.5.70 S background (neutralization IC_{50}: 708 ng/ml, ~25-fold reduction; binding affinity: 50 nM;

Fig. 6D). This result concurs with the resistance selection experiments and is potentially explained by the additional contributions of the L455F/F456L mutations present in XBB.1.5.70 S.

The above findings point to D420N and L455W as key mutations promoting viral escape from VIR-7229 in a subset of SARS-CoV-2 S variant backgrounds. RBD residue D420 is >99.99% conserved among circulating SARS-CoV-2 isolates (based on GISAID as of May 8, 2024) and 100% conserved among sarbecoviruses, including divergent bat ACE2- and non-ACE2-utilizing sarbecoviruses (**Fig. 4** and **Data S1**), suggesting a likely constraint for viral fitness. D420 is hydrogen-bonded to the Y369 side chain from a neighboring RBD in the closed S trimer and this interaction is conserved in SARS-CoV-1 S (clade 1a) and PRD-0038 S (clade 3) (**Fig. S4G**). Though the D420N substitution would be compatible with this interaction, it would form a weaker hydrogen bond, possibly altering RBD opening propensity within the S trimer and modulating both ACE2 binding and exposure of the VIR-7229 epitope.

In contrast to position 420, residue 455 has mutated in recently circulating SARS-CoV-2 variants and is key for ACE2 binding (**Fig. 5B**). We therefore assessed the impact of the L455W substitution on markers of viral fitness. Whereas the L455W substitution reduced ACE2 binding affinity (5.9-fold) of the XBB.1.5 RBD, it enhanced ACE2 binding of the EG.5 RBD (XBB.1.5 + F456L) (**Fig. 6E**), concurring with S:ACE2-mediated cell-cell fusion data (**Fig. S6D**). Additionally, L455W is anticipated to be equivalently or more immune evasive than L455F or L455S(Dadonaite, 2023); the latter two mutations have recently been associated with epidemic spread, likely driven by convergent immune pressure at the RBD positions 455 and 456 (**Fig. 6F**). Given that these observations are

at odds with the very low L455W frequency (<0.004% in all backgrounds; <0.15% in F456L background, **Fig. S6B**), we performed a bioinformatic analysis of intra-individual SARS-CoV-2 genomic variability to determine if this could be explained by low sampling (i.e., observation) frequency of the required T to G nucleotide mutation in the context of the adjacent nucleotides throughout the SARS-CoV-2 genome. Our analysis on the whole SARS-CoV-2 genome revealed an average sampling frequency of 0.0023% for any TTG to TGG mutation, required for L455W (**Table S3** and **Data S5**), which is 17-fold lower than the most frequently sampled nucleotide change that gives rise to L455F (**Table S3**). Nevertheless, it is sufficiently high to anticipate recurrent sampling of L455W and subsequent growth and transmission given the favorable ACE2 binding affinity conferred by this mutation in the F456L background and immune evasion at a site under high selective pressure. Therefore, the observation that L455W has remained very rare points to reduced viral fitness not only in the F456 S background (in which it decreased ACE2 affinity) but also in the F456L S background. A possible mechanism for the fitness defect of W455 may be that this large residue would not be accommodated in a fully closed SARS-CoV-2 S trimer without some degree of structural remodeling of adjacent residues (**Fig. S4E**).

Although VIR-7229 potently neutralized recently circulating XBB-descendant strains harboring F456L (e.g. EG.5, FL.1.5.1, and HK.3; **Figs. 1A** and **1B**), this mutation reduced neutralization potency by two orders of magnitude in an N460-harboring (BA.5) S background (**Fig. S6A** and **Data S1**). N460K (present in circulating strains since BQ.1.1) is the most relevant background for assessing future potential mAb escape, as it has become fixed in SARS-CoV-2 variants, consistent with evidence of improved

fitness(Qu, 2022) and with its high conservation across sarbecoviruses outside of clade 1b (**Figs. 4A** and **Data S1**). During revision of this manuscript, we observed that F456L also dampens VIR-7229 pseudovirus neutralization potency in the JN.1 background (JN.1.16 sub-lineage; **Fig. 1** and **Fig. 6G**), which was unexpected given that the L455S/F456L combination is neutralized with high potency in the EG.5 background (EG.5-L455S IC50 33.7 ng/ml compared to JN.1-F456L IC50 435 ng/ml; both of them harboring the L455S/F456L combination; **Fig. 6G**) despite only one conservative substitution in the VIR-7229 epitope between these variants (R403K in JN.1 sub-lineages). Moreover, the clade 1a sarbecovirus WIV1, which also contains the L455S/F456L combination, is potently neutralized by VIR-7229 (IC50 7.8 ng/ml in VeroE6 cells; **Fig. 6G**). Given that the VIR-7229 Fab 1:1 binding affinity is similar for these three variant RBDs (**Fig. 6G**), these findings suggest that other properties may influence neutralization potency, such as S protein dynamics, ACE2 binding affinity and/or epistasis, underlying the importance of evaluating mutations in the context of each strain background. It is notable that, in general, mutations L455S +/- F456L are associated with reduced ACE2 binding affinity (**Fig. 6H**), which increases the probability of reversion in future SARS-CoV-2 circulating strains. It is further notable that the growth of F456L in JN.1-decendant lineages is in the context of a marked decline in overall levels of circulating SARS-CoV-2 virus, as measured by viral activity in U.S. wastewater (**Fig. 6F**).

Overall, VIR-7229 exhibits a very high barrier to escape, as illustrated by extraordinarily narrow DMS profiles, the difficulty to select for escape mutations, and the fitness defects associated with mutations that lead to complete escape, which are rarely (if at all) observed among circulating SARS-CoV-2 isolates. This high barrier to escape,

combined with unparalleled breadth and neutralization potency, establish VIR-7229 as a promising next-generation investigational COVID-19 medicine anticipated to remain active in the context of continuing SARS-CoV-2 evolution.

10.6 Chapter Discussion

Identifying epitopes resilient to viral evolution remains fundamental to the development of durable anti-viral mAbs. One strategy has been the identification of epitopes with high phylogenetic conservation as this may be predictive of a future low likelihood of mutation if it is due to a functional constraint. However, sequence conservation may frequently result from low immune pressure rather than functional constraint[58] and these epitopes may be more vulnerable than they first appear. For example, the SD1 region (residues 323-331 and 532-591, domain C) has an average conservation of 99.6% among SARS-CoV-2 isolates in the GISAID database, but the E554K mutation present in the BA.2.86 and JN.1 variants results in full escape from SD1-targeting neutralizing mAbs(Yang, 2023, Wang, 2023). There was at least one SD1-targeting mAb in clinical development which may have been impacted by this mutation(Regeneron). Likewise, whereas the stem helix (residues 1139-1160) is highly conserved among SARS-CoV-2 variants (>99.5% in GISAID) and beta-coronaviruses, there appears to be little functional pressure to maintain the epitope residues targeted by anti-stem helix mAbs, as escape mutants are easily selected(Pinto, 2021). Therefore, prioritizing mAbs with demonstrated ability to accommodate epitope diversity may be a better strategy for long-term resilience than relying solely on evolutionary epitope conservation.

Another approach for identifying epitopes resilient to viral evolution is to have an overlap with an area of functional importance for the virus, such as the RBM, with the expectation that antigenic changes will be restrained by fitness constraints. This approach was employed for most of the SARS-CoV-2 mAbs developed at the beginning of the COVID-19 pandemic, but fell short because only a small subset of RBM residues are actually constrained by receptor binding, while the rest of them are highly variable under immune pressure(Cameroni, 2022, Starr, 2020; Taft, 2022; Thomson, 2021). Consequently, all RBM-targeting mAbs developed early in the pandemic lost their ability to neutralize circulating variants and in many cases even lost activity prior to the Omicron antigenic shift.

In this study we describe the identification and characterization of a next-generation mAb, VIR-7229, which neutralizes all SARS-CoV-2 variants that have arisen thus far, and which has potential durability to future viral evolution. The parent mAb of VIR-7229, designated S2V29, was isolated from an individual vaccinated with the ancestral Wuhan-Hu-1 S, and subsequently infected with an Omicron variant, resulting in the recall of cross-reactive memory B cells(Addetia, 2023; Tortorici, 2024, Park, 2022). S2V29 is endowed with high potency and cross-reactivity to all SARS-CoV-2 variants and sarbecovirus clades, properties that were further improved by machine learning-guided affinity maturation utilizing the SARS-CoV-1 RBD as a target antigen (the SARS-CoV-1 RBD differs from the SARS-CoV-2 RBD at positions 455 and 456, among other positions; **Fig. 4**), yielding VIR-7229. VIR-7229 is one of very few mAbs described to date capable of neutralizing all SARS-CoV-2 variants which have emerged after four and a half years of antigenic evolution, and is the only RBM-directed mAb with pan-sarbecovirus cross-

reactivity. VIR-7229 has a very high barrier to viral resistance driven by the fact that the key epitope contacts are also important for ACE2 binding and therefore functionally and evolutionary constrained, which is a form of receptor molecular mimicry. In addition, the high tolerance for epitope diversification is supported by its high-affinity binding (**Fig. 1D**), as well as by the extensive contacts with the RBD backbone which are unchanged upon residue mutations (**Fig. 3E**). Receptor molecular mimicry has also been attributed to the P4J15 and S2K146 mAbs, though their breadth and resistance to viral escape is more limited(Park, 2022; Fenwick, 2023).

The unique binding mode of VIR-7229, extraordinary pan-sarbecovirus breadth, and high tolerance for epitope variation, suggest that VIR-7229 may become an enduring SARS-CoV-2 medicine, as a stand-alone product or in combination with other mAbs. We anticipate that the remarkable characteristics of this mAb may provide a roadmap for the identification of future anti-viral mAbs with resilience to viral evolution. Finally, VIR-7229 could be considered as a component of a pandemic preparedness strategy due to its neutralization of divergent bat-infecting sarbecoviruses, including strains known to be able to evolve human ACE2 binding via single amino acid changes(Starr, 2022; Lee, 2023), in the event of a future spillover from a zoonotic reservoir.

10.7 Methods

Limitations of the Study

The diversity of SARS-CoV-2 variants which have emerged over the last four and a half years, as well as the speed with which new variants continue to emerge, has made it challenging to characterize new neutralizing SARS-CoV-2 mAbs. Ideally each experiment would be performed in every strain background, as results are impacted by the characteristics of each strain. This is especially the case for characterization of mAb

escape profiles, whether by DMS, resistance selection, or evaluation of single amino acid substitutions in binding and neutralization assays. We have performed our escape profiling experiments in multiple backgrounds, creating a comprehensive picture of the remarkably narrow VIR-7229 escape profile. However, due to experiment lead times, several experiments were not performed with the current circulating strains, e.g. we do not have DMS profile or resistance selection in the JN.1 background. Additionally, at the time of our final preparation of this manuscript (June 2024) it is a period of very low SARS-CoV-2 viral circulation. It is unknown what variant will drive the next wave of SARS-CoV-2; however, based on the unprecedented breadth of VIR-7229, as well as its molecular receptor mimicry, we anticipate that VIR-7229 will exhibit high potency to future variants.

Human participants

Serum samples utilized for mAb discovery were obtained from SARS-CoV-2 infected individuals under study protocols approved by the local Institutional Review Boards (Canton Ticino Ethics Committee, Switzerland, the Ethical committee of Luigi Sacco Hospital, Milan, Italy, and WCG North America, Princeton, NJ, US). All donors provided written informed consent for the use of blood and blood components (such as PBMCs, sera or plasma).

Cell lines

Cell lines were obtained from ATCC (HEK293T, Vero-E6), Thermo-Fisher Scientific (HEK293F), Invitrogen (Expi-CHO cells) and Takara (Lenti-X 293T cells). Vero-TMPRSS2 (Vero-T) cells were generated in-house.[62] Expi-293 (Invitrogen) and Expi-CHO cells (Invitrogen) were maintained in Expi-293 Expression Medium (Invitrogen) and ExpiCHO- Expression Medium (Gibco), respectively. Vero-E6 and Lenti-X cells were

cultured in DMEM high glucose with GlutaMAXTM (Gibco) supplemented with 10% fetal bovine serum (FBS) (Integro) and 1% Penicillin-Streptomycin (Gibco). Vero-T cells were cultured in DMEM high glucose with GlutaMAXTM (Gibco) supplemented with 10% fetal bovine serum (FBS) (Integro), 1% Penicillin-Streptomycin (Gibco) and 8 ug/mL puromycin (Gibco). All cell lines used in this study, except HEK293F and Expi293 used for protein expression and HEK293T used for sarbecovirus neutralizations, were routinely tested for mycoplasma and found to be mycoplasma-free.

Animals

All animal experiments were performed according to the French legislation and in compliance with the European Communities Council Directives (2010/63/UE, French Law 2013–118, February 6, 2013) and according to the regulations of Institut Pasteur Animal Care Committees. The Animal Experimentation Ethics Committee (CETEA 89) of the Institut Pasteur approved this study (200023; APAFIS#25326-2020050617114340 v2) before experiments were initiated. Hamsters were housed by groups of 3-4 animals in isolators with ad libitum access to water and food. The animals were manipulated in class III safety cabinets in the Institut Pasteur animal facilities accredited by the French Ministry of Agriculture for performing experiments on live rodents. All animals were handled in strict accordance with good animal practice. Before any manipulation, animals underwent an acclimation period of one week. Male golden Syrian hamsters (Mesocricetus auratus; RjHan:AURA) of 5-6 weeks of age (average weight 60-80 grams) were purchased from Janvier Laboratories (Le Genest-Saint-Isle, France) and handled under specific pathogen-free conditions.

Antibody isolation and recombinant production

S2V29 mAb was isolated from peripheral blood mononuclear cells (PBMC) of a SARS-CoV-2 convalescent and vaccinated individual under study protocols approved by a local institutional review board (Canton Ticino Ethics Committee, Switzerland). The donor provided written informed consent for the use of blood and blood derivatives for research.

PBMC were isolated by Ficoll density gradient centrifugation and B cells were enriched by staining with CD19 PE-Cy7 (BD Bioscience, cat. 341113) and incubation with anti-PE microbeads (Miltenyi Biotec, cat. 130-048-801), followed by positive selection using LS columns (Miltenyi Biotec, cat. 130-042-401). Enriched B cells were stained with anti-IgM, anti-IgD, anti-CD14 and anti-IgA, all PE labelled, and prefusion SARS-CoV-2 S with a biotinylated AviTag conjugated to Streptavidin Alexa-Fluor 647 (Fisher scientific, cat. 10308062). SARS-CoV-2 S-specific IgG+ memory B cells were sorted by flow cytometry via gating for PE-4 negative and Alexa-Fluor 647 positive cells. Antigen-specific memory B cells were co-cultured with mesenchymal stromal cells (MSC) in the presence of a cocktail of stimuli that induces the proliferation and differentiation of B cells into antibody secreting cells. After 7 days of culture, B cell supernatants were screened for the presence of mAbs of interest.

S2V29 VH and VL sequences were obtained by RT-PCR and subcloned in IgG1 expression vectors; the amino-acid sequence of the original S2V29-VL isolated from B cells carried a germline-encoded cysteine residue which was mutated to serine to reduce the risk of forming improper disulfide bonds (the mAb carrying the Cys to Ser mutation is

mAb variant S2V29-v1.2, referred to as S2V29 in this manuscript). The VH and VL amino-acid sequences of the comparator IgGs were available from previous work(Pinto, 2020, Starr, 2021, Park, 2022) or retrieved from publications(Nutalai, 2022; Murakami, 2022) and the DNA sequence was produced with codon optimization for expression in hamster cells, then subcloned into IgG1 expression plasmids. The antibodies were expressed as recombinant human IgG1 (G1m17 allotype for all, except G1m3 allotype for S2X259) carrying the half-life extending M428L/N434S (LS) mutation in the Fc region (except Omi-42 was produced with the M252Y/S254T/T256E [YTE] mutation in the Fc region). ExpiCHO cells were transiently transfected with heavy and light chain expression vectors as described previously(Pinto, 2020). For in vivo experiments in Syrian hamsters, VIR-7229 and a control mAb (specific to Plasmodium falciparum sporozoite) were produced with a Syrian hamster IgG2 Fc.

For binding and ACE2 competition measurements, VIR-7229 Fab and SA55 Fab were obtained by fragmentation of the corresponding IgG using the FabLACTICA Fab kit (Genovis, Cat #: A2-AFK-025) according to manufacturer's protocol. The Fab-containing fraction was concentrated and buffer-exchanged into filtered HBS buffer (10mM HEPES pH 7.5, 150mM NaCl) using an Amicon 10kDa cutoff concentrator (Millipore Sigma, Cat #: ACS501024). The IgG digestion reactions were analyzed by SDS-PAGE. Recombinant S309 Fab and S2X259 Fab used for the ACE2 competition experiment were expressed in HEK293 suspension cells, purified using CaptureSelect IgG-CH1 resin and buffer exchanged into PBS (ATUM Bio; Newark, CA). Recombinant

S2K146 Fab used for the ACE2 competition experiment was expressed in ExpiCHO and purified using CaptureSelect CH1-XL MiniChrom columns (Thermo Fisher Scientific).

Recombinant Fabs for crystallography were produced by ATUM Bio (Newark, CA). Engineered Fabs have been previously reported to improve crystallization by rigidifying the Fab elbow hinge(Bailey, 2018) and by replacing the human kappa constant domain FG loop (HQGLSSP) with a shorter rabbit kappa loop (QGTTS)(Lieu, 2020). These designs were incorporated into VIR-7229 Fab and S309 Fab, resulting in VIR-7229^{E} Fab and S309^{RK} Fab, respectively.

Machine-learning-assisted affinity maturation

S2V29 was affinity matured using the following approach: (1) training libraries comprising 10^6 to 10^7 S2V29 variants were designed; (2) cell surface display and FACS were used to separate the training libraries by their relative affinities to SARS-CoV-2 BQ.1.1 and SARS-CoV-1 RBDs, and next-generation sequencing (NGS) was used to determine the sequences of the different populations; (3) these first-round data were used to train a variety of AI/ML models to predict binding affinity of mAb variants towards the RBDs of SARS-CoV-1 and SARS-CoV-2 BQ.1.1; (4) the model predictions guided the design of an optimized 8 x 10^6 library which was enriched through serial rounds of FACS for improved binders; (5) the second-round data provided further training for the AI/ML models, which then helped to select a set of ~50 candidate mAbs which were produced as purified protein; and (6) the ~50 candidate mAbs were evaluated in vitro for neutralization of a panel of SARS-CoV-2 pseudoviruses as well as

SARS-CoV, and also for binding to a panel of sarbecovirus RBDs. Many of these steps are similar to previously-described experimentally-driven ML approaches.[36–38] Further details on each step are provided below.

First-round library design

No high-resolution structural information on the interaction between S2V29 and its epitope was available during the affinity maturation campaign. Therefore, two parallel approaches were taken for the design of the first-round library: (a) libraries with 1 x 10^7 mAb sequence variants mutated all possible CDR positions to determine positions that when mutated yielded a range of effects on binding, and (b) additional training libraries were designed focusing on mutations to optimize thermodynamic stability of a homology-model and structure-based design using TRIAD(Lee, 2023), a physics-based computational design suite that incorporates the Rosetta(Rohl, 2004), Dreiding(Mayo, 1990), and Phoenix force-fields(Lee, 2023). mAb variants in the libraries harbored up to eight mutations each.

Evaluation of mAb libraries by cell surface display and FACS

The coding sequences of the S2V29 VH and VL domains were subcloned into a yeast vector for cell surface Fab display, with DNA encoding a V5 epitope tag fused in-frame to the CL domain. The libraries were constructed by PCR-amplifying the CDRs with DNA oligonucleotides containing degenerate codons and reassembling the vector *in vitro* with either Golden Gate Assembly(Engler, 2008), Gibson Assembly(Gibson, 2009), or overlap extension PCR, as appropriate for each library design. Each library was transformed into *S. cerevisiae(Benatuil, 2010)*. After expressing Fabs, cells were washed with PBS

containing 0.1% BSA and incubated in the same buffer containing mouse anti-V5 antibody (SV5-Pk1, Bio-Rad Laboratories, Inc.) and either 1 μM of biotinylated SARS-CoV-1 RBD or 100 nM of biotinylated SARS-CoV-2 BQ.1.1 RBD. Cells were washed and stained with PE-streptavidin (Jackson ImmunoResearch, Inc.) and goat anti-mouse StarBright Blue 700 (Bio-Rad Laboratories, Inc.), and then subject to FACS (WOLF Cell Sorter, NanoCellect Biomedical, Inc.). Cells were binned according to levels of Fab display and antigen binding, and antibody coding sequences were sequenced using a MiniSeq (Illumina, Inc.) or a MinION (Oxford Nanopore Technologies, plc).

Training of AI/ML models and second-round screening

Data from the first-round library screening was used to train a variety of AI/ML models to predict binding affinity of mAb variants towards the RBDs of SARS-CoV-1 and SARS-CoV-2 BQ.1.1. Several types of algorithms were used, including logistic regression, neural networks, support-vector machines, and decision trees(Pedregosa, 2012). When generating models, 20% of the data was withheld as a final test set, and hyperparameters were tuned using 5-fold cross-validation with the training set. The predictions of these models were manually examined and used to guide the design of an optimized 8×10^6 combinatorial library, which contained antibody variants harboring up to 15 mutations each that were expected to bind RBDs from both SARS-CoV-2 and SARS-CoV with greater affinity than the parent mAb S2V29. Using serial rounds of FACS, as described above, the optimized library was enriched for variants that bind to the SARS-CoV-1 RBD, and the enriched populations were subsequently screened to ensure retention of binding to the SARS-CoV-2 BQ.1.1 RBD. Enriched populations from each round were deep-

sequenced, and used to develop updated ML models distinguishing variants that bound in the most stringent conditions. ML prediction scores and antibody variant abundance following FACS were used to select candidate antibodies for expression and purification from mammalian cells and downstream evaluation. Fifty-six clones were tested in a panel of in vitro neutralization and binding assays to determine the top mAb variants.

Recombinant RBDs, BA.2.86 S ectodomain, and ACE2 production

SARS-CoV-2 RBD proteins (residues 328-531 of S protein from GenBank NC_045512.2, modified as needed with mutations from other SARS-CoV-2 strains, with N-terminal signal peptide from mouse Ig heavy chain and C-terminal 8xHis-AviTag or Thrombin-8xHis-AviTag) and other sarbecovirus RBD proteins for SPR binding assays (except for Khosta-2 RBD, see below) were expressed in Expi293 cells at 37°C and 8% CO_2. See **Data S1** for full sequences. Transfections were performed using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatants were collected four to five days after transfection and supplemented with 10x PBS to a final concentration of 2.5x PBS (342.5 mM NaCl, 6.75 mM KCl and 29.75 mM phosphates). RBD proteins were purified by IMAC using Cobalt resin and buffer exchanged into PBS by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva). For BLI experiments, recombinant SARS-CoV-2 Wuhan-Hu-1 RBD was purified by cobalt affinity chromatography and biotinylated using a BirA biotin-protein ligase bulk reaction kit (Avidity) followed by size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva) into PBS.

The Khosta-2 RBD contains an N-terminal mu-phosphatase signal peptide and includes residues N_{316}RFPN$_{319}$ and C_{513}KQST$_{516}$ and a C-terminal 8xHis tag followed by an AviTag (HHHHHHHHGGSSGLNDIFEAQKIEWHE). The Khosta-2 RBD was expressed in Expi293 cells (Thermo) at 37°C and 8% CO_2. Cells were transfected with the corresponding plasmid using Expifectamine (Thermo) following the manufacturer's protocol. Four to five days post-transfection, supernatant was clarified by centrifugation at 4,121g for 30 minutes, supplemented with 25 mM phosphate pH 8.0, and 300 mM NaCl. Supernatant was then bound to a 1 mL His trap HP or Ni Excel Resin (Cytiva) previously equilibrated in 25 mM phosphate pH 8.0, 300 mM NaCl. Affinity resins were washed with 25 mM phosphate pH 8.0, 300 mM NaCl, and 40mM imidazole prior to elution with 25 mM phosphate pH 8.0, 300 mM NaCl and 300 to 500 mM imidazole. The RBD was concentrated and purified further on a Superdex 200 Increase 10/300 size-exclusion column (Cytiva) equilibrated in 50mM Tris pH 8.0 and 150mM NaCl. Fractions containing monomeric and monodisperse RBDs were flash frozen and stored at -80°C.

XBB.1.5, BQ.1.1, and EG.5 RBD constructs used for crystallization included a C-terminal 8xHis-Avi, Thrombin-TwinStrep-8xHis, or Thrombin-8xHisTag, respectively. Proteins were expressed similarly as XBB.1.5, BQ.1.1, and EG.5 RBD constructs used for SPR binding assays, as described above, except with the addition of 10 μM kifunensine. Cell culture supernatant was collected four days after transfection and supplemented with 10x PBS to a final concentration of 2.5x PBS. Protein was purified using a HisTALON Superflow cartridge (Takara) followed by buffer exchange into PBS using a Superdex 200 Increase 10/300 GL column (Cytiva).

The SARS-CoV-2 BA.2.86 Hexapro S ectodomain construct harbors its native signal peptide, BA.2.86 mutations (T19I, R21T, L24-, P25-, P26-, A27S, S50L, H69-, V70-, V127F, G142D, Y144-, F157S, R158G, N211-, L212I, V213G, L216F, H245N, A264D, I332V, G339H, K356T, S371F, S373P, S375F, T376A, R403K, D405N, R408S, K417N, N440K, V445H, G446S, N450D, L452W, N460K, S477N, T478K, N481K, V483-, E484K, F486P, R493Q, Q498R, N501Y, Y505H, E554K, A570V, D614G, P621S, H655Y, I670G, N679K, P681R, N764K, D796Y, S939F, Q954H, N969K, P1143L), Hexapro mutations (F817P, A892P, A899, A942P, K986P, V987P)(Hseih, 2020), a mutated furin cleavage site ($_{682}$RRARSV$_{687}$ to $_{682}$GSASSV$_{687}$), and a C-terminal foldon followed by an AviTag and an 8xHis tag. The SARS-CoV-2 BA.86 S ectodomain was expressed in Expi293 cells (Thermo) incubated at 37°C and 8% CO_2. Cells were transfected using Expifectamine293 (Thermo) following the manufacturer's protocol. Four days post-transfection, Expi293 cell supernatant was clarified by centrifugation at 4,121g for 30 minutes, supplemented with 25 mM phosphate pH 8.0, 300 mM NaCl. The supernatant was then passed over an His-Trap Excel column (Cytiva) previously equilibrated in 25 mM phosphate pH 8.0, 300 mM NaCl and washed with 20-40mL of buffer containing 25 mM phosphate pH 8.0, 300 mM NaCl, and 40mM Imidazole. S protein was eluted using 25 mM phosphate pH 8.0, 300 mM NaCl, and 300mM imidazole prior to being buffer exchanged to 50 mM Tris-HCl pH 8.0, 150 mM NaCl using a centrifugal filter device with a MWCO of 100 kDa. The S glycoprotein was subsequently run over a Superose 6 Increase 10/300 size-exclusion chromatography column (Cytiva) equilibrated in 50mM Tris pH 8.0 and 150mM NaCl and

the fractions containing monodisperse prefusion trimers were flash frozen and stored at -80°C.

For SPR binding measurements, recombinant human ACE2 (residues 19-615 from Uniprot Q9BYF1 with a C-terminal thrombin cleavage site-TwinStrep-10xHis-GGG-tag, and N-terminal signal peptide) was expressed in Expi293 cells at 37°C and 8% CO_2. Transfection was performed using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatant was collected six days after transfection, adjusted to a final concentration of 80 mM Tris-HCl pH 8.0, 100 mM NaCl, and then incubated with BioLock solution (IBA GmbH). ACE2 was purified using a StrepTrap HP column (Cytiva) followed by isolation of monomeric ACE2 by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva) pre-equilibrated in 20 mM Tris-HCl pH 7.5, 150 mM NaCl. Recombinant ACE2 used for BLI competition experiments (residues 19-615 from Uniprot Q9BYF1 with a C-terminal Avi-10xHis-GGG-tag, and N-terminal signal peptide) was expressed in Expi293 cells as described above and purified using a HisTrap excel column followed by buffer exchange using a Superdex 200 Increase 10/300 GL column (Cytiva) pre-equilibrated in PBS.

Production of VSV-based SARS-CoV-2 (and SARS-CoV-1 Urbani) S pseudotyped virus

To generate SARS-CoV-2 and SARS-CoV-1 Urbani S pseudotyped vesicular stomatitis virus (VSV) for use in neutralization assays employing VeroE6 cells, Lenti-X 293T cells were seeded in 15-cm dishes. The next day, cells were transfected with the plasmid

encoding for the SARS-CoV-2 spike variant (or SARS-CoV-1 Urbani) using TransIT-Lenti (Mirus Bio) according to the manufacturer's instructions. One day post-transfection, cells were infected with VSV(G*ΔG-luciferase) (Kerafast) at an MOI of 3 infectious units/cell. Viral inoculum was washed off after one hour and cells were incubated for another day at 37°C. The cell supernatant containing S pseudotyped VSV was collected at day 2 post-transfection, centrifuged at 1,000 x g for 5 minutes to remove cellular debris, aliquoted, and frozen at -80 °C.

VSV-based sarbecovirus (clade 1b, clade 1a, and clade 2) S pseudotyped virus production

To generate sarbecovirus S pseudotyped VSV for use in neutralization assays performed in HEK-293T-hACE2 and HEK-293T-R.alc.ACE2, constructs for membrane-anchored S glycoproteins from SARS-CoV-1 Urbani, BA.2.86 (WPL86459.1), GX-Pangolin (QIA48623.1), Khosta-1 (QVN46559.1), Khosta-2 (QVN46569.1), SARS-CoV-1 Civet007 (AAU04646.1), RaTG13delta21 (QHR63300.2), WIV1 (AGZ48828.1), RsSHC014 (AGZ48806.1), PRD-0038 (QTJ30153.1), PRD-0038-dm (harboring mutations of the SARS-CoV-2 equivalent positions K493Y/T498W) (QTJ30153.1),[39,40] and BtKY72 (APO40579.1) were codon optimized and synthesized by Genscript for mammalian cell expression, cloned in frame with a Kozak sequence to direct translation and harboring wild-type signal peptides. The last 21 residues were deleted,[73] except for SARS-CoV-1 Urbani, Civet007, GX-Pangolin, and WIV1. Khosta-1, Khosta-2, RaTG13, RsSHC014, PRD-0038, PRD-0038-dm (K493Y/T498W) S genes were synthesized with a triple Flag tag while the rest of the genes were synthesized with no tag. All the S genes were cloned

into the HDM vector[74] except for WIV1 S and SARS-CoV-1 Urbani S which were cloned into pcDNA3.1(-) and for GX-Pangolin S which was cloned into phCMV1.

Sarbecovirus S pseudotyped VSV were generated as previously described.[21] Briefly, HEK293T cells in DMEM supplemented with 10% FBS and 1% PenStrep and seeded in poly-D-lysine coated 10-cm dishes were transfected with a mixture of 24 µg of the corresponding plasmid encoding for: SARS-CoV-1 Urbani S, BA.2.86 S , WIV1 S, RaTG13 S, SARS-CoV-1 Civet007 S, Khosta-1, Khosta-2 S, GX-Pangolin S, RsSHC014 S, PRD-0038 S S, PRD-0038-dm S or BtKY72 S and 60 µl Lipofectamine 2000 (Life Technologies) in 3 ml of Opti-MEM, following manufacturer's instructions. After 5 h at 37°C, DMEM supplemented with 20% FBS and 1% PenStrep was added. The next day, cells were washed three times with DMEM and were transduced with VSVΔG-luc.[75] After 2 h, virus inoculum was removed and cells were washed five times with DMEM prior to the addition of DMEM supplemented with anti-VSV-G antibody [Il-mouse hybridoma supernatant diluted 1 to 25 (v/v), from CRL-2700, ATCC] to minimize parental background. After 18-24 h, supernatants containing pseudotyped VSV were harvested, centrifuged at 2,000 x g for 5 minutes to remove cellular debris, filtered with a 0.45 µm membrane, concentrated 10 times using a 30 kDa cut off membrane (Amicon), aliquoted, and frozen at -80°C until use.

Neutralization of SARS-CoV-2 (and SARS-CoV-1 Urbani) pseudoviruses

For neutralization of VSV-based SARS-CoV-2 (and SARS-CoV-1 Urbani) S pseudotyped viruses, Vero E6 cells were seeded into clear bottom white walled 96-well plates

(PerkinElmer) at 20,000 cells/well and cultured overnight at 37°C. The next day, 9-point 3-fold serial dilutions of mAbs were prepared in media. SARS-CoV-2 (or SARS-CoV-1 Urbani) pseudotyped VSVs were diluted at 0.1 MOI in media and added 1:1 to each mAb dilution. Virus:mAb mixtures were incubated for 1 hour at 37 °C. Media was removed from the Vero E6 cells and 50 µL of virus:mAb mixtures were added to the cells. One hour post-infection, 100 µL medium was added to all wells. After 20-24 hours incubation at 37 °C, medium was removed and 100 µL Steadylite plus reagent (diluted 1:1 in DPBS) was added to each well. The plates were incubated at room temperature for 10 minutes and luminescence was read on a plate reader. Two to three technical replicates were measured and at least two biological repeats were performed. For most neutralizations performed with the S2V29 mAb and for the majority of VIR-7229 neutralizations against single-site point mutations (but not VIR-7229 neutralizations against SARS-CoV-2 strains), experiments were performed with the addition of 100 ng/ml anti-VSV-G Ab. All data were normalized based on internal control RLU values (untreated cells for 100% neutralization and infected cells with or without anti-VSV-G antibody for 0% neutralization) and plotted with GraphPad Prism (version 10.1.2) using a nonlinear regression 4-parameters model.

Neutralization of sarbecovirus pseudoviruses

For pseudotyped VSV sarbecovirus neutralizations, HEK293T cells were transiently transfected with plasmids encoding for full-length human ACE2 or R. alcyone ACE2 following a previously described protocol(Tortorici, 2022).[75] Briefly, HEK293T cells at 90% confluency and seeded in poly-D-lysine coated 10-cm dishes were transfected with

a mixture of 8 µg of the corresponding plasmid encoding the ACE2 ortholog and 30 µl of Lipofectamine 2000 (Life Technologies) prepared in Opti-MEM according to the manufacturer's instructions. After 5 h at 37°C, cells were trypsinized, seeded into poly-D-lysine coated clear bottom white walled 96-well plates at 40,000 cells /well and cultured overnight at 37°C. For neutralizations, eleven 2-fold serial dilutions of SA55, S2K146, S2X259, Omi-42, S309, S2V29 or VIR-7229 IgGs were prepared in DMEM. 20 µl of the different sarbecovirus pseudotypes described above were added 1:1 (v/v) to each IgG and mixtures were incubated for 45-60 min at 37°C. After removing their media, transfected HEK293T cells were washed two times with DMEM and 40 µL of the mixture containing virus:IgG were added. Two hours later, 40 µL DMEM were added to the cells. After 17-20 h, 60 µL of One-Glo-EX substrate (Promega) were added to each well and incubated on a plate shaker in the dark. After 5-15 min incubation, plates were read on a Biotek Neo2 plate reader. (S309 neutralization of SARS-CoV-1 Urbani was performed with Vero-TMPRSS2 cells.) Measurements were made in duplicate with at least two biological replicates. Relative luciferase units were plotted and normalized in Prism (GraphPad): cells alone without pseudotyped virus were defined as 0% infection, and cells with virus only (no IgG) were defined as 100% infection.

Neutralization of authentic SARS-CoV-2 viruses

VeroE6 cells were seeded into flat bottom tissue culture 96-well plates at 20,000 cells/well and cultured overnight at 37°C. Twenty-four hours later, 9-point 1:4 serial dilutions of VIR-7229 were prepared in infection media (DMEM + 2% BSA) and each dilution was tested in triplicate per plate (range: 1,250 to 0.019 ng/mL final concentration). SARS-CoV-2

authentic virus stock was diluted in infection media for a final concentration of 200 plaque forming units per well (MOI 0.01). Antibody dilutions were added to virus and incubated for 30 minutes at 37°C. Media was removed from the cells, mAb-virus complexes were added, and cells were incubated at 37°C. At 18-30 hours post-infection (wild-type USA-WA1/2020 18h, Delta 30h, BA.1 24h, BA.2 30h, BA.5 18h, XBB.1.5 24h, XBB.1.6 30h, EG.5.1 24h, FL.1.5.1 30h), cells were fixed with 4% PFA for 30 minutes at RT, then washed 3 times with PBS to remove residual PFA. The cells were permeabilized with 100 µL of 0.25% Triton X-100 in PBS for 30 minutes at RT, followed by three washes with PBS. Cells were incubated with 50 µL of anti-SARS-CoV-2 nucleocapsid antibody (Sino Biologicals, 40143-R001) at 1:2000 for 1 hour at RT. Plates were washed three times with PBS and then incubated for 1 hour at RT with 50 µL/well of goat anti-rabbit IgG Alexa647 (Invitrogen, A-21245) secondary antibody at a final dilution of 1:1000 mixed with 2 µg/mL Hoechst dye. After washing 3 times with PBS, 200 µL of fresh PBS was added for imaging. Plates were imaged on a Cytation5 plate reader. Whole well images were acquired (12 images at 4X magnification per well) and nucleocapsid-positive cells were counted using the manufacturer's software.

Affinity determination by surface plasmon resonance (SPR)

Measurements were performed using a Biacore 8K or Biacore T200 instrument. CM5 chips with covalently immobilized anti-AviTag polyclonal antibody (GenScript, Cat #: A00674-40) were used to capture His-AviTag containing RBDs. Running buffer was 1x HBS-EP+ pH 7.4 (10 mM HEPES, 150 mM NaCl, 3 mM EDTA and 0.05% v/v Surfactant P20) (Cytiva, Cat #: BR100669). Experiments were performed with a 3-point 4-fold

dilution series of monomeric VIR-7229 Fab starting at 50nM (50, 12.5, 3.13 nM; or for JN.1, a 4-point 4-fold dilution series starting at 50 nM) or monomeric TwinStrep-His-tagged ACE2 starting at 300nM (300, 75, 18.75 nM).

Experiments were run as single-cycle kinetics with at least 2 replicates for each RBD ligand. Data were double reference-subtracted and fit to a binding model using the Biacore Insight software. The 1:1 binding model was used to determine the kinetic parameters. K_D, k_a, and k_d are reported in **Data S1** as the average of all replicates with the corresponding standard deviation. For VIR-7229 Fab binding to Shaanxi2011 RBD, a constant R_{max} calculated from the RBD capture level was applied to account for the low R_{max} from the default analysis.

Competition of Fab fragments and ACE2 for binding to RBD by biolayer interferometry (BLI)

Protein reagents were diluted in 10x Kinetics Buffer (Sartorius, Cat #:18-1105). The experiment was performed on an Octet Red instrument. Streptavidin biosensors (Sartorius, Cat #: 18-5019) were hydrated in water for 15 min before the experiment. Biotinylated His-Avi-tagged Wuhan (Wu-WT) RBD was immobilized on the sensors at 10 µg/mL for 10 s. Wu-WT RBD-immobilized sensors were then dipped into Kinetics Buffer for 60s to establish a baseline before being dipped into a 100 nM Fab solution for 600 s (association phase 1) and subsequently into a mixture of 100 nM of the same Fab plus 300 nM His-Avi-tagged ACE2 solution for 600 s (association phase 2).

Cell-surface mAb-mediated S_1 shedding

Expi-CHO cells were transiently transfected with plasmids encoding for Wuhan-Hu-1 or XBB.1.5 S. After 48h of incubation at 37°C, cells were collected, counted, and 90,000 cells/well were dispensed into a round-bottom 96 well plate (Corning). After a washing step with wash buffer (PBS supplemented with 0.01% BSA), cells were incubated with 15 µg/ml mAbs (diluted in wash buffer) for 180, 120, 60, 30 or 5 min at 37°C. After the incubation for the allotted time, cells were washed with ice-cold wash buffer and stained with 1.5 µg/mL Alexa Fluor647-labeled Goat Anti Human IgG secondary Ab (Jackson Immunoresearch) for 20 min on ice in the dark. Cells were then washed twice with cold wash buffer and analyzed using a ZE5 cytometer (Biorad) with acquisition chamber T= 4°C. Binding at each time point (MFI) was determined, normalizing to the MFI at 5 minutes time point, and data plotted using GraphPad Prism v. 9.1.1

Fc-mediated effector function
Determination of mAb-Dependent Activation of Human FcγRIIa and FcγRIIIa
Activation of human FcγRIIa (allele H131) and FcγRIIIa (high-affinity binding allele V158) was tested using validated, commercially available bioreporter assays in which CHO cells stably expressing SARS-CoV-2 Wuhan-Hu-1 spike protein (CHO-CoV-2-Spike) were used as the target antigen. Nine-point serial dilutions of mAbs (5,000 ng/ml to 0.076 ng/ml) were incubated with 12,500 (for FcγRIIIa) or 10,000 (for FcγRIIa) CHO-CoV-2-Spike cells per 96-plate well in a white, flat-bottom plate for 25 minutes at room temperature. Jurkat effector cells (Promega; Cat. Nr.: G7018 and G9995) expressing indicated FcγRs and stably transfected with an NFAT-driven luciferase gene were thawed, diluted in assay buffer, and added to the plate at an effector to target cell ratio of 6:1 for FcRγIIIa or 5:1 for FcγIIa. Control wells were also included that were used to

measure antibody-independent activation (containing target cells and effector cells but no antibody) and background luminescence of the plate (wells containing assay buffer only). Plates were incubated for 20 hours at 37°C with 5% CO2. Activation of human FcγRs in this bioassay results in the NFAT-mediated expression of the luciferase reporter gene. Luminescence was measured with a luminometer (Bio-Tek) after adding the Bio-Glo™ Luciferase Assay Reagent according to the manufacturer's instructions.

Determination of NK-Cell Mediated Antibody-Dependent Cellular Cytotoxicity

3,000 SARS-CoV-2 Wuhan-Hu-1 S CHO-K1 (HaloTag-HiBiT) cells per well were thaw and seeded in a white polypropylene 96 well round-bottom plate and incubated overnight in a 37°C, 5% CO2 humified incubator. The day after, serial dilutions of mAbs (serially diluted 5-fold in AIM-V Medium from 20,000 ng/ml to 0.26 ng/ml) were added to the plate already containing the cells (incubation for 10 min prior to mixing with NK cells). Wells without antibody or containing Digitonin (Sigma-Aldrich; Cat. Nr.: D141) at 100ug/ml were used as negative and positive controls. Finally, primary human NK cells freshly isolated from whole EDTA blood (Becton Dickinson AG, Cat. Nr.: 367525) from two previously genotyped donors expressing heterozygous affinity (F/V158) or high affinity (V/V158) FcγRIIIa using the MACSxpress NK isolation Kit (Miltenyi, Cat. Nr.: 130-127-695) following the manufacturer's instructions and resuspended in AIM-V medium were added as effectors at an effector-to-target ratio of 10:1 (30,000 cells/well) and the plate was incubated in a 37°C, 5% CO_2 humified incubator for 4 hours. ADCC results in the lysis of target cells which release HiBiT protein into the medium producing luminescence, which can be read with a luminometer after addition of NanoLuc HiBiT Extracellular Detection reagent (Promega; Cat. Nr.: N2421) according to the manufacturer's instructions. In brief, after substrate reconstitution, 70ul/well was added and incubated in the dark for 10

minutes and luminescence was then measured. The percent specific lysis was determined by applying the following formula:

(specific RLU – negative control RLU) / (positive RLU – negative control RLU) x 100.

Evaluation of sarbecovirus cross-reactivity via high-throughput yeast-display binding assays

The complete pipeline for measuring mAb breadth across the pan-sarbecovirus panel is described at:

https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29/blob/main/results/summary/summary.md

MAb binding via high-throughput FACS-seq was evaluated against a previously published pan-sarbecovirus panel of yeast-displayed RBDs(Starr, 2022; Lee, 2023) that was supplemented with additional newly described sarbecovirus and SARS-CoV-2 variants (SARS-CoV-2 Omicron BA.1, BA.2, BA.5, BQ.1.1, and XBB.1.5; RhGB03(Apaa, 2023); RhGB02-230(Tan, 2023); Rc-kw8, Rc-os20, and Rc-mk2(Murakami, 2022); and BtSY1-RtLS01 and BtSY2-RmCX02(Wang, 2022)).

The yeast-display RBD library was grown, induced for yeast-surface expression, and labeled with monoclonal antibody at 10,000, 400, 16, 0.64, 0.0256, and 0 ng/mL concentration for one hour at room temperature. Yeast were washed with PBS-BSA and labeled with secondary Myc-FITC antibody (Immunology Consultants CMYC-45F) and PE-conjugated goat anti-mouse-IgG (Jackson ImmunoResearch 109-115-098). Libraries were then partitioned, at each labeling concentration, into four bins of mAb binding on a BD FACSAria, collecting a minimum of 1 million RBD^+ cells per sample concentration across the four bins. Cells were grown post-sort, plasmid purified, N16 barcode amplified, and sequenced on an Illumina NextSeq. Raw Illumina sequencing data is available from

the NCBI Sequence Read Archive, BioProject PRJNA714677, BioSample SAMN41715061. Barcode reads were mapped to library barcodes, with raw counts found at: https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29/blob/main/results/counts/variant_counts.csv.

For each library barcode, an EC50 binding strength was derived from its distribution of sequence reads across sort bins. First, the strength of mAb binding to each barcode at each mAb dilution was determined as the simple mean bin from cell counts across integer-weighted bins. Any barcode with less than 2 cell counts at any single sample concentration or less than an average of 5 cell counts across all sample concentrations was eliminated from analysis. An EC50 metric was then calculated from the fit of a sigmoidal curve between mean bin (mAb binding) and mAb labeling concentration. Per-barcode EC50 calculation and representative titration curve-fits can be found at: https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29/blob/main/results/summary/compute_EC50.md, and per-barcode EC50 metrics are available at: https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29/blob/main/results/bc_mAb_EC50/bc_mAb_EC50.csv.

We then computed the per-variant EC50 as the robust mean of replicate barcodes linked with the identical RBD variant, by taking the mean per-barcode EC50 after trimming tails of the top and bottom 5% of EC50 values among the replicate barcodes. The final variant derivation can be found at: https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29/blob/main/results/summary/collapse_barcodes_lib61_SARSr-wts.md,

and final per-variant mAb-binding values are available at: https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29/blob/main/results/final_variant_scores/final_variant_scores_lib61.csv

Evaluation of escape mutants via yeast-display deep mutational scanning

Deep mutational scanning libraries for SARS-CoV-2 variants Wuhan-Hu-1, Omicron BA.2, Omicron BQ.1.1, and Omicron XBB.1.5 were described in prior publications, including library construction, library availability, and measurements of mutational impacts on RBD expression and ACE2-binding affinity(Starr, 2022; Taylor, 2023; Starr, 2022). New deep mutational scanning libraries for SARS-CoV-2 variants EG.5 and BA.2.86 were constructed and assayed for mutational impacts on expression and ACE2-binding affinity per these prior methods, and are available from GitHub in advance of future publication: https://github.com/tstarrlab/SARS-CoV-2-RBD_DMS_Omicron-EG5-FLip-BA286. These libraries consist of virtually all single amino acid changes in each RBD background in a yeast-surface display platform.

Duplicate yeast-display deep mutational scanning libraries were induced for RBD expression, and 5 OD*mL of yeast were incubated in 1 mL for one hour at room temperature with a concentration of mAb corresponding to the EC90 of the mAb for the respective yeast-displayed wildtype RBD determined from pilot isogenic binding assays. In parallel, for FACS gate setting, 0.5 OD*mL of the respective wildtype parental constructs were incubated in 100 µL of antibody at the matched EC90 concentration or 1/10 the EC90 concentration. Cells were washed, incubated with 1:100 FITC-conjugated

chicken anti-Myc antibody to label RBD expression and 1:200 PE-conjugated goat anti-human IgG to label bound antibody, and washed in preparation for FACS.

Antibody-escape cells in each library were selected via FACS on a BD FACSAria II or Cytek Aurora Cell Sorter. FACS selection gates were drawn to capture approximately 50% of yeast expressing the parental RBD control labeled at the 10x reduced antibody labeling concentration (see representative gating scheme in **Fig. S5A**). For each sample, 4 million RBD+ cells were processed on the sorter with collection of cells in the antibody-escape bin, which were expanded overnight, plasmid purified, and barcodes sequenced on an Illumina NextSeq. In parallel, plasmid samples were purified from 30 OD*mL of pre-sorted library culture and sequenced to establish pre-selection barcode frequencies. Barcode reads are available on the NCBI SRA, BioProject PRJNA770094, BioSample SAMN41694243.

Demultiplexed Illumina barcode reads were matched to library barcodes and associated RBD mutant from previously assembled barcode-variant lookup tables using dms_variants (version 0.8.9), yielding a table of counts of each barcode in each pre- and post-sort population, available at: https://github.com/tstarrlab/SARS-CoV-2-RBD_Omicron_MAP_S2V29/tree/main/results/counts.

The escape fraction of each barcoded variant was computed from sequencing counts in the pre-sort and antibody-escape populations via the formula:

$$E_v = F \times \frac{\left(\frac{n_v^{post}}{N_{post}}\right)}{\left(\frac{n_v^{pre}}{N_{pre}}\right)}$$

where F is the total fraction of the library that escapes antibody binding (e.g. annotated numbers in Fig. S4A), n_v is the counts of variant v in the pre- or post-sort samples with a pseudocount addition of 0.5, and N is the total sequencing count across all variants pre- or post-sort. These escape fractions represent the estimated fraction of cells expressing a particular variant that fall in the escape bin, which scales from 0 for a mutation that never causes sufficient loss of binding to drive cells into the antibody-escape bin, to 1 for a mutation that escapes binding >10-fold such that the variant falls into the antibody-escape bin defined by the control labeling gates. We applied computational filters to remove mutants with low pre-selection sequencing counts or highly deleterious mutations that escape antibody binding artefactually due to poor RBD surface expression, specifically mutants with orthogonally measured ACE2-binding impacts of <–3 (1000-fold loss of ACE2 binding) or expression scores of <–1.25, <–0.955, <–1.229, and <–1.25 for Wuhan-Hu-1, BA.2, BQ.1.1, and XBB.1.5, respectively, accounting for the variation in baseline expression levels of different wildtype variants. Final per-mutant escape fractions were computed as the average across barcodes within replicates, with the correlation between replicate library selections shown in Fig. S5B. Final escape fraction measurements averaged across replicates are available at: https://github.com/tstarrlab/SARS-CoV-2-RBD_Omicron_MAP_S2V29/tree/main/results/supp_data.

Polyreactivity to Hep2 cells

Assessment of polyreactivity was performed using HEp-20-10 (Euroimmun 1522-2010), an FDA-approved immunofluorescence test for reliable antinuclear antibody screening.

Test and control antibodies were diluted in PBS-Tween to a final concentration of 50 µg/mL and 25 µl of the antibody solution was applied on the test slide as described by the manufacturer's instructions. After 1h incubation at room temperature the test slide was washed 5 min in PBS-tween. Alexa488 anti hu IgG Fc gamma specific (109-545-098) was prepared at 3 µg/ml in PBS-Tween and 20 µl of the secondary antibody solution was added to the slides. After 1h-incubation at room temperature the slide was washed again for 5 min with PBS-tween, excess washing solution was removed and 10 µl glycerol/biochip was added to mount coverslip. Immunofluorescence images were acquired using the Automated Imaging Microplate Reader Cytation5 (Biotek).

Two antibodies previously extensively characterized for their unspecific binding profiles were used as positive and negative controls. An anti influenza A haemagglutinin antibody (FI6) was used as a positive control, while as a negative control an antibody targeting respiratory syncytial virus (RSV) F protein (MPE8) was used; the anti-RSV MPE8 has previously been tested on a large panel of human tissues and was shown not to display any significant non-specific binding.

Human tissue cross-reactivity

Assessments of cross-reactivity of S2V29 and VIR-7229 to a panel of human tissues (non-GLP) were performed by Labcorp at the Labcorp Early Development Laboratories in North Yorkshire, England. The relevant study numbers were 8510456 and 8520675, respectively.

To facilitate immunohistochemical detection, S2V29, VIR-7229, and the MGH2 negative control mAb were conjugated by Labcorp with AF488 using the commercial Alexa Fluor® 488 Protein Labelling Kit (ThermoFisher Scientific, A10235). The test system was preparations derived from histologically normal frozen human tissues, collected from one donor of each of these 39 tissues: Adrenal, Urinary Bladder, Blood Cells, Bone Marrow, Breast, Cerebellum (Brain), Cortex (Brain), Colon, Duodenum, Endothelium, Eye, Fallopian Tube, Gastric Antrum, Gastric Body, Heart, Ileum, Kidney, Liver, Lung, Lymph Node, Oesophagus, Ovary, Pancreas, Parotid, Peripheral Nerve, Pituitary, Placenta, Prostate, Skin, Spinal Cord, Spleen, Striated Muscle, Testis, Thymus, Thyroid, Tonsil, Ureter, Cervix (Uterus), and Endometrium (Uterus). Tissues, with the exception of blood cells, were cryo-sectioned. Blood cells were prepared as smears.

To generate positive and negative controls, CHO-K1 cells (ATCC® CRL-9618) previously engineered to stably express SARS-CoV-2 Spike protein on the cell surface (CHO-nCoV-2S cells) and parental CHO-K1 cells (CHO-K1 cells) were cultured for 6 days in DMEM/Ham's F-12 supplemented with 15 mM HEPES, stable Glutamine and 10% FBS. On day 6, cells were detached with Trypsin-EDTA, gently washed, and resuspended in PBS at a concentration of 10^6 cells/ml.

Positive control cell blocks were generated as a heterogenous sample of both positive-control and negative-control cells (to ensure signal is not saturated): CHO-nCoV-2S and CHO-K1 cells were mixed at a 7:3 ratio. For negative control cell blocks, only CHO-K1 cells were used. For each cell suspension, 3x1 ml were harvested by centrifugation, the supernatant was decanted, and the cells were resuspended in the remaining volume.

Finally, the cell suspensions were added to Shadon Disposable Base Mold prefilled with 1ml of Killik embedding medium (Bio-Optica, 05-9801) and mixed homogenously by gently swirling with a 1ml pipet tip. The molds were rapidly frozen by placing them on top of an acetone/dry ice bath and were immediately transferred to dry ice. The frozen molds were wrapped and store at -80°C until use.

Method development was performed by Labcorp to generate a suitable immunohistochemical staining method for use in the control titration and tissue titration. Findings for each tissue were individually graded and identified by cell type or structure where possible. In addition to grading intensity (I) and frequency (F), a text comment for each finding recorded, where possible, the staining pattern observed as either membranous, cytoplasmic or nuclear, or combinations thereof. The numerical score for intensity (I) indicates the highest intensity of positive staining observed.

In vivo efficacy evaluation using a Syrian hamster model

Male golden Syrian hamsters (Mesocricetus auratus; RjHan:AURA) of 5-6 weeks of age (average weight 60-80 grams) were purchased from Janvier Laboratories (Le Genest-Saint-Isle, France) and handled under specific pathogen-free conditions. The hamsters were treated intraperitoneally with VIR-7229-GH-rIgG2a (5 mg/kg, 1.5 mg/kg, 0.5 mg/kg, 0.17 mg/kg, or 0.06 mg/kg) or with the isotype control mAb MPE8v3-GH-rIgG2a (1.5 mg/kg), n=6 animals per group; an additional n=6 group was used for the mock-infection group. One day later, animals were anesthetized (intraperitoneal administration of ketamine [Imalgène 200 mg/kg] and xylazine [Rompun 2 mg/kg]) and

inoculated intranasally with 6x104 PFU of SARS-CoV2 Omicron/XBB.1.5 (GISAID ID: EPI_ISL_16353849, kindly provided by O. Schwartz and colleagues) or with 6x104 PFU of SARS-CoV2 Omicron/JN.1 (GISAID ID: EPI_ISL_18522058, provided by the National Reference Centre for Respiratory Viruses hosted by Institut Pasteur (Paris, France). Mock-infected animals received the physiological solution only. Infected and mock-infected hamsters were housed in separate isolators and were followed-up daily during four days at which the body weight and the clinical score were noted. On day 4 post-infection the animals were sacrificed and lungs harvested for evaluation of viral RNA load and viral titer.

Viral RNA load (RT-qPCR)

Frozen lungs fragments were weighted and homogenized with 1 mL of ice-cold DMEM (31966021, Gibco) supplemented with 1% penicillin/streptomycin (15140148, Thermo Fisher) in Lysing Matrix M 2 mL tubes (116923050-CF, MP Biomedicals) using the FastPrep-24™ system (MP Biomedicals) and the following scheme: homogenization at 4.0 m/s for 20 s, incubation at 4°C for 2 min, and new homogenization at 4.0 m/s for 20 s. The tubes were centrifuged at 10,000 x g for 1 min at 4°C. Afterwards, 125 µL of the tissue homogenate supernatant were mixed with 375 µL of Trizol LS (10296028, Invitrogen) and the total RNA was extracted using the Direct-zol RNA MiniPrep Kit (R2052, Zymo Research). The presence of SARS-CoV-2 RNA in these samples was evaluated by one-step RT-qPCR (Invitrogen 11732-020) in a final volume of 12.5 µL per reaction in 384-wells PCR plates using a thermocycler (QuantStudio 6 Flex, Applied Biosystems). Briefly, 2.5 µL of RNA were added to 10 µL of a master mix containing

6.25 µL of 2X reaction mix. 0.2 µL of $MgSO_4$ (50 mM), 0.5 µL of Superscript III RT/Platinum Taq Mix (2 UI/µL), and 3.05 µL of nuclease-free water containing the nCoV_IP2 primers (nCoV_IP2-12669Fw: 5'-ATGAGCTTAGTCCTGTTG-3'; nCoV_IP2-12759Rv: 5'-CTCCCTTTGTTGTGTTGT-3') at a final concentration of 400 nM, and the nCoV_IP2 probe (5'-FAM-AGATGTCTTGTGCTGCCGGTA-3'-TAMRA) at a final concentration of 200 nM. The amplification conditions were as follows: 55°C for 20 min, 95°C for 3 min, 50 cycles of 95°C for 15 s and 58°C for 30 s, and a last step of 40°C for 30 s. Viral load quantification (expressed as RNA copy number/g of tissue) was assessed by linear regression using a standard curve of six known quantities of RNA transcripts containing the *RdRp* sequence (ranging from 10^7 to 10^2 copies). The limit of detection is $1x10^2$ viral RNA copies/µL. This value is then converted to RNA copy number/g of tissue considering the weight of the homogenized lung fragments and the concentration of RNA in each sample.

Viral titer (TCID50)

Frozen lung fragments were weighted and homogenized with 1 mL of ice-cold DMEM supplemented with 1% penicillin/streptomycin (15140148, Thermo Fisher) in Lysing Matrix M 2 mL tubes (116923050-CF, MP Biomedicals) using the FastPrep-24™ system (MP Biomedicals), and the following scheme: homogenization at 4.0 m/s for 20 s, incubation at 4°C for 2 min, and new homogenization at 4.0 m/s for 20 s. The tubes were centrifuged at 10.000 x g for 2 min at 4°C and the supernatants collected. Lung homogenates were then serial diluted (1:10) in DMEM supplemented with 1% penicillin/streptomycin and 1 µg/mL of Trypsin-TPCK (4370285, Sigma-Aldrich) and then 100 µL of each dilution were added in a well of a 96 well-plate in six replicates. 100 µL

containing $8x10^4$ Vero-E6 cells were added in each well and the plates were incubated at 37°C and 5% CO_2 for 72 hours. The plates were then washed in PBS, stained with crystal violet (11778193, BD) for 15 minutes and washed again in PBS. Lysed wells were counted and the TCID50 was calculated using the TCID50 calculator (v2.1 - 20-01-2017_MB. available at: https://www.klinikum.uni-heidelberg.de/fileadmin/inst_hygiene/molekulare_virologie/Downloads/TCID50_calculator_v2_17-01-20_MB.xlsx). The limit of detection is $3.16x10^1$ plaque-forming units (PFU)/mL. The values of TCID50 in PFU/mL were then converted to PFU/100 mg of tissue considering the weight of the homogenized lung fragments.

Enzyme-linked immunosorbent assay (ELISA)

Ninety-six half area well-plates (Greiner 650001) were coated overnight at 4°C with 50 µL of SARS-CoV-2 Wuhan-Hu-1 RBD protein (nCoV-RBD-D-STREPH) prepared at 2 µg/mL in PBS pH 7.2. Plates were then washed three times with PBS 0.05% Tween 20 (PBS-T) and blocked with casein 1% in PBS during 2 hours. The plates were washed three times in PBS-T and 50 µL of diluted serum samples (1:10-1:20) were applied to each well, in duplicate. Standard curves were made with serial dilutions (1:3) of VIR-7229. The plates were incubated for 1 hour, washed six times with PBS-T and incubated again with an AP-labeled anti-hamster IgG (0.5 µg/mL) for 1 hour in the dark. Plates were then washed six times with PBS-T and 4-nitrophenyl phosphate substrate (pNPP, Sigma-Aldrich, 71768) was added. After 45 min incubation, absorbance at 405 nm was measured by a plate reader (Victor Nivo, Perkin Elmer).

Crystallization, data collection, structure determination and analysis

For both VIR-7229 Fab:RBD and S2V29 Fab:RBD crystallography, additional, non-competing Fabs (S309 or S2H97) were added during Fab:RBD complexation to support crystal formation. XBB.1.5 RBD was deglycosylated with EndoH (25,000 U/mg RBD) and mixed with a 1.1-fold molar excess of VIR-7229^{E} Fab and S309^{RK} Fab. BQ.1.1 and EG.5 RBDs were deglycosylated with EndoH (25,000 U/mg RBD) and the C-terminal purification tag was cleaved with thrombin (20 U/mg RBD). Deglycosylated and tagless BQ.1.1 and EG.5 RBDs were mixed with a 1.1-fold molar excess of S2H97 and either S2V29 or VIR-7229 Fabs, respectively. The complexes were purified on a Superdex 200 10/300 GL column pre-equilibrated with 20 mM Tris-HCl pH 8.0, 50 mM NaCl for BQ.1.1 complex or 25 mM Tris-HCl pH 7.5, 50 mM NaCl for XBB.1.5 and EG.5 complexes. Crystals of all three complexes were obtained by the sitting-drop vapor diffusion method at 20°C.

For XBB.1.5 RBD-VIR-7229^{E}- S309^{RK}, a total of 200 nL complex at 8 mg/ml was mixed with 200 nL mother liquor solution from the Morpheus protein crystallization screen(Gorrec, 2009) containing 0.1 M carboxylic acids (0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M sodium citrate tribasic dihydrate, 0.02 M potassium sodium tartrate tetrahydrate, 0.02 M sodium oxamate), 0.1 M buffer system 2 pH 7.5 (sodium HEPES, MOPS), and 30% precipitant mix 2 (20% v/v ethylene glycol, 10% w/v PEG 8000). Crystals were flash frozen in liquid nitrogen using the mother liquor solution as a cryoprotectant.

For EG.5 RBD-VIR-7229-S2H97, a total of 200 nL complex at 8 mg/ml was mixed with 200 nL mother liquor solution containing 0.1 M TRIS pH 8, 22% w/v PEG-MME 2000, and 20 mM $NiCl_2$. Crystals were flash frozen in liquid nitrogen using the mother liquor solution supplemented with 20% ethylene glycol as a cryoprotectant.

For BQ.1.1-S2V29-S2H97 crystals, a total of 200 nL complex at 7.7 mg/mL was mixed with 200 nL mother liquor containing 0.1 M Tris pH 8.5, 20% PEG-MME 2000, and 10 mM $NiCl_2$. Crystals were flash frozen in liquid nitrogen using the mother liquor solution supplemented with 20% ethylene glycol as a cryoprotectant.

For XBB.1.5 and EG.5 complexes, data were collected at beamline 14-1 at the Stanford Synchrotron Radiation Lightsource facility in Stanford, CA. Data were processed with the XDS software package(Kabsch, 2010) for final datasets of 2.41 Å in space group $P2_12_12_1$ and 1.90 Å in space group $P2_12_12_1$ for XBB1.5 RBD-VIR-7229^{E}-S309^{RK} and EG.5 RBD-VIR-7229-S2H97, respectively. For BQ.1.1-S2V29-S2H97, data were collected at beamline 8.2.1 at the Advanced Light Source in Berkeley, CA, and processed similarly to a final resolution of 1.67 Å in space $P2_12_12_1$.

The complex structures were solved by molecular replacement using Phaser(McCoy, 2007) from starting models consisting of RBD-S309 (PDB: 7R6W) or RBD-S2H97 (PDB: 7M7W) and homology models for the respective Fabs generated using the Molecular Operating Environment (MOE) software package (Chemical Computing Group, https://www.chemcomp.com). Subsequent rounds of model building and refinement were performed using Coot(Emsley, 2010), Refmac5(Murshudov, 2011), and Phenix(Liebschner, 2019). Validation was performed using Molprobity(Chen, 2010).

CryoEM sample preparation, data collection and data processing

For one dataset, BA.2.86 Hexapro S was incubated at 1 mg/ml with a 1.5 molar excess of VIR-7229 Fab during 30-45 seconds at room temperature. For the other two datasets, BA.2.86 Hexapro S precomplexed with a 1.5 molar excess of S309 Fab was incubated for 5-10 min with a 1.5 Fab molar excess of VIR-7229 Fab for 30-45 seconds. Three microlitres of the complexes were loaded onto freshly glow discharged R 2/2 UltrAuFoil (Electron Microscopy Sciences), M4-Au300-2.0/1.0 holey NiTi grids (Single Particle LLC) or C-Flat 2/2-4Cu-50 (Electron Microscopy Sciences) covered with a thin layer of manually added carbon before plunge-freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 6–6.5 s blot time or blot force of –1 and 4.5 s blot time for thin carbon grids, at 100% humidity and 21 °C. Data were acquired on a FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K3 direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon(Suloway, 2005) at a nominal magnification of 105,000× with a pixel size of 0.835 Å. The dose rate was adjusted to ~10 electrons per pixel per second, and each movie was fractionated in 100 frames of 40 ms per frame. A total of three datasets were collected to obtain the final structure with a defocus ranging between 0.8 and 2 µm, yielding a total of 29,989 micrographs. Data collected using sample vitrified on gold grids and NiTi grids comprised untilted data along with data collected with the stage tilted at 30° and 45° to circumvent particle preferential orientation(Tan, 2017) whereas data collected from the sample vitrified on C-Flat grids covered with thin carbon comprised untilted data only. For each dataset, movie frame alignment, estimation of the microscope contrast-transfer

function parameters, particle picking, and extraction were carried out using Warp(Tegunov, 2019). Particles were extracted with a box size of 400 pixels with a pixel size of 1.67 Å. Two rounds of reference-free 2D classification were performed in cryoSPARC(Punjani, 2017) to select well-defined particle images. After 2D classification, particles from three datasets were combined and an initial model was generated, using ab-initio reconstruction in cryoSPARC (Punjani, 2017), and used as reference for heterogenous 3D refinement. Particles belonging to classes with the best resolved S density were selected. To improve particle picking further, the Topaz picker was trained on Warp-picked particle sets belonging to the selected classes after heterogeneous 3D refinement. The particles picked using Topaz were extracted and subjected to two rounds of 2D-classification followed by heterogenous 3D refinement in cryoSPARC. The two different particle sets picked from Warp and Topaz were merged and duplicate particle picks were removed in cryoSPARC using a minimum distance cutoff of 90 Å. After two rounds of heterogeneous refinements, the particles belonging to the class with the best resolved RBD:VIR-7229 Fab density were selected and used to carry out a non-uniform refinement (NUR)(Punjani, 2020). Particles from the NUR were transferred from cryoSPARC to Relion using the pyem program package (Asarnow,) and subjected to the Bayesian polishing procedure(Zivanov, 2019) in Relion(Zivanov, 2018; Kimanius) during which particles were re-extracted with a box size of 512 pixels and a pixel size of 1.0 Å. After polishing, particles were subjected to 2D-classification followed by a heterogeneous refinement in cryoSPARC to select particles belonging to the class with the best-resolved RBD:VIR-7229 Fab density. NUR with per-particle defocus refinement yielded a final reconstruction of BA.2.86 S in complex with VIR-7229 Fab at 3.1 Å resolution comprising

314,440 particles (the S309 density was largely averaged out). To further improve the density at the RBD:VIR-7229 Fab interface, local refinement was performed using cryoSPARC with a soft mask comprising the RBD and the VIR-7229 variable domains yielding a reconstruction at 3.3 Å resolution enabling model building. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion(Chen, 2013; Rosenthal, 2003). Local resolution estimation was carried out using cryoSPARC.

Model building and refinement

RBD and VIR-7229 Fab complex models were built and refined by iterating between manual rebuilding in Coot(Emsley, 2010) and refinement in Rosetta(Frenz, 2019; Wang, 2016). Validation was done using Phenix and Molprobity(Liebschner, 2019; Chen, 2010). Figures were generated using UCSF ChimeraX(Pettersen,2004).

Molecular dynamics analysis of Fab:RBD and ACE2:RBD

The coordinates of VIR-7229:RBD (EG.5) and VIR-7229:RBD (XBB.1.5) were obtained from the present work, and ACE2:RBD (XBB.1.5) from PDB ID 8FXB (ACE2:RBD [XBB.1])(Addetia, 2023) with F486P mutagenesis and rotamer optimization (to change XBB.1 to XBB.1.5) performed in silico in MOE using ProteinBuilder (MOE 2022.02; https://www.chemcomp.com). Glycan coordinates were taken from previous work(Thomson, 2021). These models were prepared using QuickPrep (MOE 2022.02; https://www.chemcomp.com).

The three complexes were parameterized for molecular dynamics (MD) as previously described(Starr, 2022) using AMBER(Case, 2021) with the ff14SB protein force field(Maier, 2015), GLYCAM_06j-1 glycan force field(Kirschner, 2008), TIP3P water force field,(Jorgensenm 1983) and Joung and Cheatham ions force field(Joung, 2008).

For the VIR-7229:RBD (EG.5) and VIR-7229:RBD (XBB.1.5) complexes, the equilibration protocol followed that used in previous studies(Momont, 2023; Hauser, 2016). By seeding with different initial velocities, five independent 0.8 μs AMBER MD simulations (4.0 μs) were executed for each of the two complexes.

For ACE2:RBD (XBB.1.5), five independent trajectories (each with production simulation time 0.9 μs) were generated for a total of 4.5 μs. Equilibration and production MD were run with OpenMM 8.(Eastman) Equilibration was performed according to a multistage protocol as previously (Hauser, 2016) with the exception that the following atoms were left unrestrained in all stages of the equilibration protocol: the ACE, NME caps and missing loops in 8FXB. Energy minimization stages were performed using the OpenMM 8 LocalEnergyMinimizer with an energy tolerance of 10 kJ/mol. The molecular dynamics stages used the OpenMM 8 LangevinMiddleIntegrator(Fass, 2018; leimkuhler, 2013; leimkuhler, 2016). Hydrogen atom masses were set to 4 amu by transferring mass from connected heavy atoms, bonds to hydrogen were constrained, and center of mass motion was not removed. Pressure was controlled by a molecular-scaling Monte Carlo barostat with a pressure of 1 atmosphere, a temperature of 300 K, and an update interval of 50 steps. Non-bonded interactions were treated with the Particle Mesh Ewald method (Sun, 2022) using a real-space cutoff of 1.0 nm and an Ewald error tolerance of 0.00025, with grid spacing selected automatically. Long range anisotropic dispersion corrections

were applied to steric interactions(Shirts, 2007). A virtual bond was added between the first atoms of each protein chain to ensure that the chains are imaged together. Default parameters were used unless noted otherwise. The code for running equilibration and molecular dynamics is available at: https://github.com/choderalab/rbd-ace2-xbb15-simulations.

After excluding the first 2.5 ns of MD, trajectories were post-processed followed by scripted protein:protein contact analysis in MOE (CCG MOE 2022.02) as previously described(Momont, 2023). A contact is defined for residue:residue (or residue:glycan) distances within 5 Å; inter-residue contacts were evaluated over all MD frames (sampled every 10 ns). Fraction occupancy for each RBD:mAb or RBD:ACE2 residue:residue (or residue:glycan) pair was calculated as the percentage of MD frames where a contact was observed. Contact analysis for the static structures were performed on the models after QuickPrep.

Selection of SARS-CoV-2 mAb escape mutants by rVSV serial passaging

Propagation of replicating VSV-SARS-CoV-2 S chimeras (rVSV)

Wuhan-Hu-1 rVSV (GFP) was produced as described earlier(Case, 2020). Omicron rVSV-spike constructs were designed in-house and purchased from VectorBuilder (en.vectorbuilder.com) and propagated in Vero-TMPRSS2 cells. For stock production, 100mm dishes (Falcon Cat. 353003) of Vero-TMPRSS2 were infected at MOI = 0.03 in infection medium (DMEM Gibco Cat. #11995-040, 1% Pen/strep Gibco Cat. #15140-122, 2% FCS VWR Cat. #97068-085, 20mM HEPES pH 7.7 Gibco Cat. 15630080). Plates

were incubated at 34°C, 5% CO_2 for 1 hour, after which inoculum was removed and fresh media was added (same formulation). Plates were incubated at 34°C, 5% CO_2 for 72 hours. CPE was assessed visually, and virus containing media was removed, clarified and stored at -80°C for later use.

rVSV titration

Vero-TMPRSS2 cells were plated in a 12-well format (CellTreat Cat. 229111). A five-point curve of 10-fold rVSV-spike dilutions were prepared (starting at 1:100 dilution) in infection medium. Diluted virus was incubated with the cells for 1 hour, after which inoculum was removed and plates washed once with PBS (Gibco Cat. 10010023). An overlay of 1% methylcellulose (Sigma Aldrich Cat. M7027-250G) was added to each well, and plates were incubated at 34°C for 24 hours. Overlays were removed and plates fixed in 4% PFA for 30 minutes, and washed 3 times with PBS. Cells were permeabilized with 0.25% Triton X-100 (Sigma Aldrich Cat. X100-100ML) and stained with mouse anti-VSV-N primary antibody (clone 10G4, Kerafast Cat. EB0009). Plates were washed 3 times with PBS and stained with horseradish peroxidase-conjugated anti-mouse secondary (Jackson ImmunoResearch, Cat. 115-035-062) for 30 minutes at room temperature. Plates were washed 4 times and developed using True Blue substrate (SeraCare Cat. 5510-0030) for 15 minutes. A final wash step with ultrapure H_2O (Gibco Cat. 10977015) was used to inactivate the enzyme and foci were counted manually to calculate infectious titers.

rVSV Serial Passaging

To produce an even monolayer in a 12-well format, 2.5E5 Vero-TMPRSS2 cells were plated in 2mL of complete DMEM (10% FCS). After 24 hours, rVSV-spike aliquots were thawed and diluted to infect cells at MOI = 2 after neutralization. Antibody dilutions (7-point, 4-fold, final maximum concentration = 20 mg/mL) were prepared in 2 mL deepwell plates (Nunc Cat. 260251) and diluted rVSV-spike was added. Antibody and virus mixture was incubated at 37°C for 1 hour. After neutralization, virus was added to cells and incubated at 37°C for 1 hour, with gentle rocking every 15 minutes. Inoculum was removed, wells were washed once with PBS, and fresh media was added, containing matched concentrations of antibody. Plates were incubated at 37°C for 72 hours, and percent cytopathic effect (CPE) was assessed via inverted light microscope. The criteria used for selection was: maximum antibody treatment at which CPE exceeds 20%. Supernatants from the selected well were collected, clarified and stored at -80°C for later use. Supernatants from the no-antibody positive control well were also collected for passaging.

For subsequent passages, Vero-TMPRSS2 were plated as before. Passage 1 virus was diluted 1:5 in media containing antibody (diluted as before) and neutralized as above. Infection, selection, and collection were performed identically for each passage. Passaging was ceased after complete escape (CPE >20% at 20 mg/mL) or after 7-10 passages had elapsed. All passaging experiments were performed in duplicate.

rVSV Resistance Mutant Sequencing

Using Trizol (Invitrogen Cat. 15596026) phase extraction followed by column cleanup (Macherey-Nagel Cat. 740983.50), RNA was extracted from resistant rVSV-

spike samples. RNA was reverse transcribed using Protoscript II reverse transcriptase (NEB Cat. M0368L) and cDNA for the SARS-CoV-2 spike gene was amplified in two fragments using the following primers (IDT) and KAPA HiFi HotStart ReadyMix (Roche Cat. 07958935001): ATTGCCACTAGTCTCTAGTC & CAAGAACAACAGCCCTTGAG, CTTTACAAGGGAACGATTGAGC & ATCGGAAGAGAATTGAATTTCC. PCR products were cleaned up using NucleoSpin Gel and PCR Clean-up (Macherey-Nagel Cat. 740609.250) and purity visualized via gel electrophoresis (Invitrogen Cat. A42100). Samples were submitted to MCLab (www.mclab.com) for Sanger sequencing, or Primordium (primordiumlabs.com) for Nanopore sequencing. Results were analyzed with Snapgene software (www.snapgene.com).

Plaque-based selection of SARS-CoV-2 mAb escape mutants by rVSV

Production of VSV-SARS-CoV-2 S chimeras

Recovery of recombinant VSV was performed as described(Whelan, 1995). Briefly, BSRT7/5 cells(Buchholz, 1999) were inoculated with vaccinia virus vTF7-3(Fuerst, 1986) and subsequently transfected with T7-expression plasmids encoding VSV N, P, L, and G, and an antigenomic copy of the viral genome. Cell culture supernatants were collected at 72 h, clarified by centrifugation (5 min at 1,000 x g), and filtered through a 0.22 mm filter. Virus was plaque-purified on Vero CCL81 cells in the presence of 25 mg/mL of cytosine arabinoside (Sigma-Aldrich), and plaques in agarose plugs were amplified on Vero CCL81 cells. Viral stocks were amplified on MA104 cells at an MOI of 0.01 in Medium 199 containing 2% FBS and 20 mM HEPES pH 7.7 at 34 °C. Viral supernatants were

harvested upon extensive cytopathic effect and clarified of cell debris by centrifugation at 1,000 x g for 5 min. Aliquots were maintained at -80 °C.

Selection of monoclonal antibody resistant mutants (MARMs)

Replicating VSV-SARS-CoV-2 S chimeras were used to select for MARMs as previously described(Case, 2020;Liu, 2021). In brief, MARMs were recovered by plaque isolation on Vero E6 cells (ATCC, CRL-1586) with the indicated monoclonal antibody in the overlay. The concentration of monoclonal antibody in the overlay was determined by neutralization assays at a MOI of 100. Escape clones were plaque-purified on Vero cells (ATCC, CCL-81) in the presence of monoclonal antibody, and plaques in agarose plugs were amplified on MA104 cells (a gift from H. B. Greenberg (Stanford School of Medicine)) with the monoclonal antibody present in the medium. Viral stocks were amplified on MA104 cells at an MOI of 0.01 in Medium 199 containing 2% FBS and 20 mM HEPES pH 7.7 (Millipore Sigma) at 34 °C. Viral supernatants were collected upon extensive cytopathic effect and clarified of cell debris by centrifugation at 1,000g for 5 min. Aliquots were maintained at −80 °C. Viral RNA was extracted from VSV-SARS-CoV-2 mutant viruses using RNeasy Mini kit (Qiagen), and S was amplified using OneStep RT–PCR Kit (Qiagen). The mutations were identified by Sanger sequencing (Genewiz). Their resistance was verified by subsequent virus infection in the presence or absence of antibody. In brief, Vero cells were seeded into 12-well plates for overnight. The virus was serially diluted using DMEM and cells were infected at 37 °C for 1 h. Cells were cultured with an agarose overlay in the presence or absence of monoclonal antibody at 34 °C for 2 days. Plates were scanned on a biomolecular imager and expression of eGFP is shown at 48 h after infection.

Cell-cell fusion between VeroE6-TMPRSS2 and SARS-CoV-2 expressing cells

Split GFP cell-cell fusion assay

To evaluate the effect of point mutations in the XBB.1.5 and EG.5 S backgrounds in fusion, we used the split GFP cell-cell fusion assay. Briefly, BHK21 cells transfected with the different S point mutants and VeroE6-TMPRSS2 stably expressing GFP1-10 and GFP 11, respectively, were utilized to quantify cell-cell fusion over 18 hours. BHK21-GFP1-10 cells were cultured in 6-well cell-culture grade plates and incubated overnight in DMEM containing 10% FBS and 8 µg/ml puromycin. Next day, BHK21-GFP1-10 cells were transfected with 4ug of DNA encoding one S construct, Opti-MEM and lipofectamine2000 (Invitrogen Cat# 11668027) following manufacture instructions and VeroE6-TMPRSS2 GFP11 cells in DMEM containing 10% FBS, 8 µg/ml puromycin and 4 µg/ml blasticidin were seeded in a black, glass bottom 96-well plate at ~36,000 cells per well and incubated overnight. After ~20 hr incubation, transfected BHK21-GFP1-10 cells were washed three times with FluoroBrite DMEM (Gibco Cat# A1896701) and incubated with 500uL of Gibco cell dissociation buffer for 10 minutes at 37 °C with agitation to facilitate resuspension of cells. Cells were then passed through cell-strainer capped tubes (Falcon Cat# 08-771-23) to eliminate aggregates and diluted using FluoroBrite DMEM to have ~60,000 cells per mL. VeroE6-TMPRSS2-GFP11 cells were carefully washed with FluoroBrite DMEM three times and 150 uL of transfected BHK21-GFP1-10 cells were added on top and co-cultured for a period of ~18 hours in the Cytation7 plate imager. Data were analyzed with GraphPad PRISM 10 to quantify GFP+ areas.

Flow cytometry for normalization of expression

Flow cytometry was utilized to quantify S expression levels on the cell surface of BHK21-GFP1-10 cells. BHK21-GFP1-10 cells under puromycin selection were cultured in 6 well cell-culture grade plates and incubated overnight. After ~20hrs, the BHK21-GFP1-10 cells were washed with unenriched DMEM and 2.5 mL of fresh DMEM with 10% FBS was added to cells prior to transfection step. To match conditions with the cell-cell fusion assay, cells were transfected with 4ug of S plasmid and lipofectamine2000 in Opti-MEM and incubated for ~38 hours. The BHK21-GFP1-10 cells were then washed with Flow Stain buffer and 500 uL of Gibco cell dissociation buffer was added to remove cells from bottom of the 6 well plates. Cells were resuspended with Flow Stain buffer (eBioscience Cat# 00-4222-26), filtered through a cell-strainer capped tube, diluted to ~1x10^7 cells per mL. and added to a pyramid bottom 96-well plate. Cells in the plate were centrifuged and the supernatant discarded. Cells were resuspended in Flow Stain Buffer containing 5 µg/mL of S2L20 mAb(McCallum, 2021) and incubated at 4 °C for ~25 minutes. Cells were washed 3 times with 100 µl of flow stain buffer before adding 0.5 µg of Anti-IgG Fc-PE Secondary Antibody per well (Thermo Fisher Scientific Cat#12-4998-82). After ~25 minutes at 4 °C, cells were washed 3X with Flow Stain Buffer. To fix the cells, 100uL of 2% paraformaldehyde was added to each well and incubated at 4 °C for 15 minutes. Cells were then washed 2 times with Flow Stain Buffer prior to being resuspended in 50 µL of fresh Flow Stain Buffer. Cells were transferred into tubes with 450 µL of Flow Stain Buffer, kept on ice, shielded from direct light prior to being counted on BD Biosciences' FACSymphony. All results were analyzed on BD FlowJo 10.8.2. For

normalization of expression and cell-cell fusion, total percentage of fusion after 18 hours was divided by mean fluorescence intensity (MFI).

Prevalence analysis in GISAID database

The viral sequences and the corresponding metadata were obtained from GISAID EpiCoV project (https://www.gisaid.org/). Analysis was performed on sequences submitted to GISAID up to May X, 2024. S protein sequences were either obtained from GISAID download page or, for the most recently submitted sequences, from the genomic sequences with exonerate 2.4.0–haf93ef1_3 (https://quay.io/repository/biocontainers/exonerate?tab=tags) using protein to DNA alignment with parameters -m protein2dna –refine full –minintron 999999 –percent 20 and using accession NC_045512.2 as a reference. Multiple sequence alignment of all spike proteins was performed with mafft 7.508—hec16e2b_0 (https://quay.io/repository/biocontainers/mafft?tab=tags&tag=7.508--hec16e2b_0) with parameters --mapout --auto --op 4.5 --reorder --keeplength --addfragments using the same reference as above. The --mapout parameter was used to retrieve insertions. S sequences that were <80% (1019/1273) of the canonical protein length were discarded. To identify each mutation prevalence, missingness (or ambiguous amino acids) was taken into account in both nominator and denominator. Per week prevalence of each mutation was then calculated to get the temporal trend.

Bioinformatic analysis of intra-individual SARS-CoV-2 genomic variability

We used low frequency viral variants occurring intra-individual as a proxy to estimate the replication error rate of the SARS-CoV-2 RNA polymerase, i.e. how frequently is a variant sampled in the absence of pressure in the SARS-CoV-2 genome.

This strategy has the caveat that it could also capture polymerase errors introduced during the library preparation and sequencing errors. The former should however be orders of magnitude less common than the viral polymerase error rate, and therefore should introduce limited noise, while the latter can be minimized with stringent quality controls filtering. We analyzed raw data (fastq) rather than consensus sequences in order to identify intra-individual variation. 1,763 SARS-CoV-2 samples (paired fastq files) collected from June 2022 to Dec 2023 were downloaded from the SRA database using fastq-dump of sratoolkit (version 3.0.7) [parameters: *–split-files*]. Only samples sequenced with Illumina were selected for this analysis, in order to limit the technological error rate. Variants were called against the consensus of the respective sample using an in-house pipeline that leverages trimmomatic v.0.39(Bolger, 2014), bwa-mem v.0.7.17(Li, 2011)), lofreq v.2.1.5(Wilm, 2012), and bcftools v.1.10.2(Danecek, 2021). As mutation rates have been shown to be context dependent, i.e. are influenced by the nucleotide(s) adjacent to the variant(di Iulio, 2018), we computed trimer variation rates. The frequency of each trimer variant (N=192, or 4^3*3, i.e each possible trimer varying to 3 alternate trimers based on the middle nucleotide) was calculated for each sample, as follows: The numerator is the number of times a trimer variant (e.g. TTG > TGG) event is seen in a sample, and the denominator is the number of times the respective trimer (e.g. TTG) is seen in the viral genome consensus sequence of the corresponding sample. The minimum number of read coverage is set at 50. In order to minimize the potential technological error rate interference, a variant event was considered for each mutation with allelic frequency >1% that passed all of the following criteria: (i) at least 2 reads carrying the mutation, (ii) a Phred quality score of ³27 (corresponding to an error rate

~0.02%) at the mutation position, (iii) an average read quality score of ³27 for the reads carrying the mutation, (iv) a read quality score of ³30 for ³50% of the reads carrying the mutation and (v) no positional bias in the read (e.g. the mutation always being located at the same position in the read). The frequency obtained for each trimer variant is then aggregated across individuals (extracting the mean and standard error). The threshold of 1% was empirically chosen to minimize the possibility of sequencing errors (the higher the threshold, the less likely we would capture technical errors), while also minimizing the capture of variants under pressure (the higher the threshold, the more likely the variant would be under pressure and detectable at higher proportion in the sample).

QUANTIFICATION AND STATISTICAL ANALYSIS

Description of the analysis of neutralization assays can be found in Method Details sections, “Neutralization of sarbecovirus pseudoviruses” and “Neutralization of authentic SARS-CoV-2 viruses.” Final IC50 values are the geometric mean of per-replicate measurements, with exact values of *n* given in **Data S1**.

Description of the analysis of SPR binding assays can be found in Method Details section “Affinity determination by surface plasmon resonance (SPR),” analyzed via Biacore Insight software. Final K_D values were computed as the mean and standard deviation across replicates, with exact values of *n* given in Data S1.

Description of the analysis of yeast-display pan-sarbecovirus binding assays can be found in Method Details section “Evaluation of sarbecovirus cross-reactivity via high-throughput yeast-display binding assays.” The final EC50 reported is the robust mean (eliminating top and bottom 5% of values) across internally replicated barcodes linked to

each RBD library variant. Full quantitative analysis pipeline is available from GitHub: https://github.com/tstarrlab/SARSr-CoV_mAb-breadth_S2V29.

Statistical analysis of *in vivo* hamster protection studies is described in the **Fig. 2** legend, with exact *n* defined in Method Detail section, “In vivo efficacy evaluation using a Syrian hamster model.” Significance was established via comparison of median values via ANOVA non-parametric Kruskal-Wallis test followed by Dunn’s multile comparison test.

Description of the analysis of yeast-display deep mutational scanning escape mapping can be found in Method Details section “Evaluation of escape mutants via yeast-display deep mutational scanning. The final escape fraction is the mean of experimental duplicates. Full quantitative analysis pipeline is available from GitHub: https://github.com/tstarrlab/SARS-CoV-2-RBD_Omicron_MAP_S2V29.

10.8 Figures and Tables

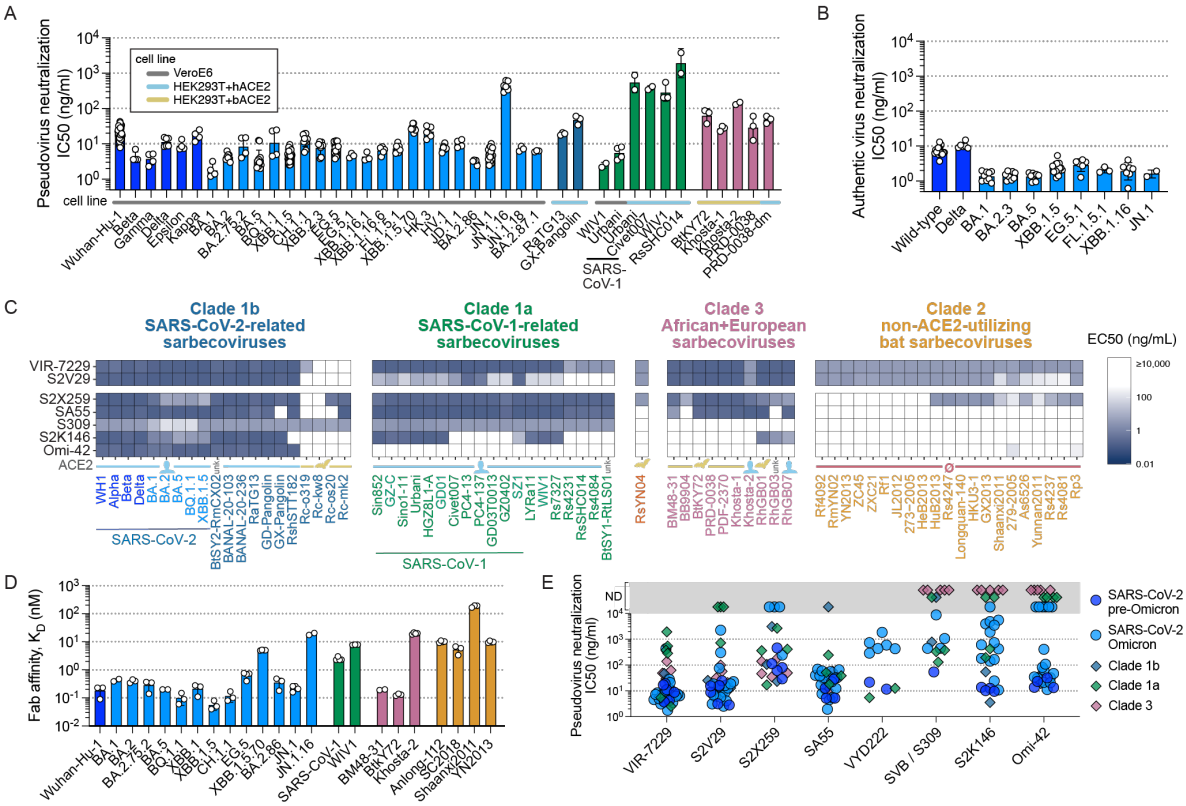


Figure 10.1: VIR-7229 is a potent, pan-sarbecovirus neutralizing mAb.

(A) VIR-7229-mediated neutralization of SARS-CoV-2 variant or sarbecovirus pseudoviruses. Bar color denotes sarbecovirus clade, as in panel C. Horizontal lines beneath bars denote cell line employed in neutralization assay: VeroE6 (gray), HEK-293T+human ACE2 (cyan), HEK-293T+bat (*R. alcyone*) ACE2 (yellow). PRD0038-dm refers to PRD0038 harboring the K482Y/T487W RBD mutations (SARS-CoV-2 numbering 493Y/498W), which allow for entry using human ACE2(Starr, 2022; Lee, 2023). SARS-CoV-1 Urbani and WIV1 experiments were run in two different assay conditions, see Methods. See also **Data S1**. See **Fig. S2** for neutralization mechanisms of action.

(B) VIR-7229-mediated neutralization of authentic SARS-CoV-2 virus, performed with VeroE6 cells. Wild-type virus isolate has the same S haplotype as Wuhan-Hu-1. See also **Data S1**.

(C) Breadth of VIR-7229 and comparator mAbs binding to a yeast-displayed panel of sarbecovirus RBDs spanning the known phylogenetic diversity. Colored line below the graph, denoted by “ACE2,” indicates whether a sarbecovirus has been found to bind to or enter cells via human ACE2 (blue), bat but not human ACE2 (yellow), no ACE2 (pink), or unknown (unk.). See **Fig. 4A** for phylogenetic relationships and clade definitions. See **Data S1** for full sequences, phylogeny, and alignment.

(D) VIR-7229 Fab fragment binding affinity measured by SPR. Bar color denotes sarbecovirus clade, as in panel C. SARS-CoV-1 RBD is the Urbani strain. See also **Data S1**.

(E) Overview of pseudovirus neutralization by comparator mAbs, colored by sarbecovirus clade as in panel A. Data points within the gray bar represent neutralization not detected (ND), i.e. IC50 >10,000 ng/ml. See **Fig. S1** and **Data S1** for data separated by strain.

Virology and clinical endpoints on day 4 after SARS-CoV-2 XBB.1.5 or JN.1 infection of Syrian hamsters prophylactically administered with VIR-7229 produced with hamster Fc or 1.5 mg/kg (mpk) isotype-matched control antibody. See also **Data S3**.

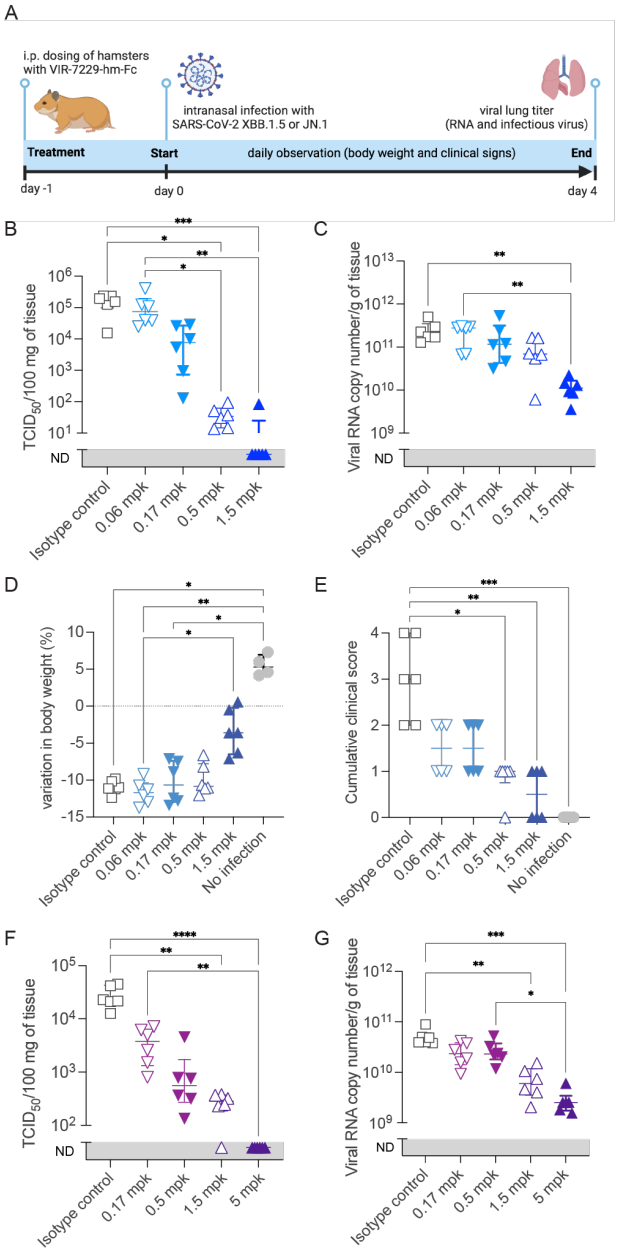


Figure 10.2: In vivo efficacy of VIR-7229.

(A) Experiment outline.

(B) Infectious viral lung titers for XBB.1.5 infection. ND: not detected. The limit of detection expressed as $TCID_{50}$/100 mg of tissue is not uniform as it depends on lung tissue weight and sample dilution.

(C) Lung viral RNA load for XBB.1.5 infection. ND: not detected.

(D) Variation in body weight relative to day 0 for XBB.1.5 infection.

(E) Cumulative clinical score for XBB.1.5 infection (0-4): ruffled fur, slow movements, apathy, absence of exploratory activity.

(F) Infectious viral lung titers for JN.1 infection. Same comments as panel B.

(G) Variation in body weight relative to day 0 for JN.1 infection. ND: not detected.

(B-G) X-axis indicates dose of VIR-7229-hmFc or 1.5 mpk isotype control. Median ± interquartile range is shown; significance is based on ANOVA non-parametric Kruskal-Wallis test followed by Dunn's multiple comparison test, *p<0.05, **p< 0.01, *** p < 0.001.

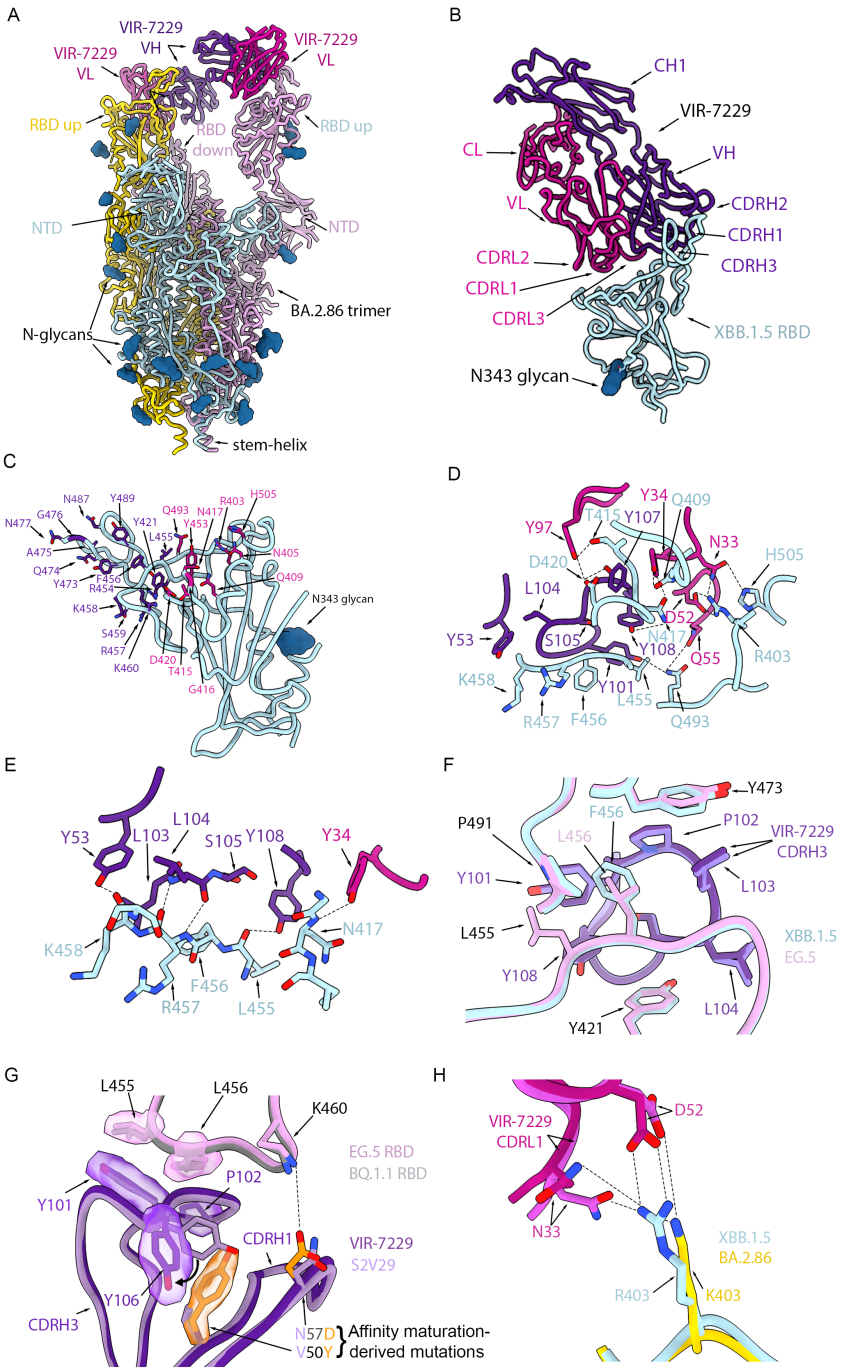


Figure 10.3: Structural basis for VIR-7229 pan-sarbecovirus neutralization.

(A) Ribbon diagram of the cryoEM structure of the BA.2.86 S ectodomain trimer (cyan, pink and gold) in complex with two VIR-7229 Fabs (purple and magenta) with N-linked glycans rendered as blue surfaces. See also **Fig. S3**.
(B) Ribbon diagram of the VIR-7229-bound XBB.1.5 RBD crystal structure. The bound S309 Fab is omitted for clarity. The N343 glycan is rendered as a blue surface. See also **Fig. S4**.

(C) XBB.1.5 RBD (cyan) with VIR-7229 epitope residues shown as sticks and colored according to the (dominant) Fab interacting chain. RBD residues 420, 453, 455, 460 and 493 interact with the VIR-7229 heavy and light chains and were colored based on the chain with which they bury the greatest surface area.
(D) Zoomed-in view of select interactions formed between VIR-7229 and the XBB.1.5 RBD. Hydrogen bonds and salt bridges are indicated with black dash lines.
(E) Zoomed-in view of hydrogen-bonds (black dash lines) formed between VIR-7229 and the RBD backbone.
(F) Superposition of the VIR-7229-bound XBB.1.5 RBD (cyan RBD, dark purple mAb) and VIR-7229-bound EG.5 RBD (pink RBD, light purple mAb) showing accommodation of the F456L residue mutation.
(G) Superposition of the VIR-7229-bound EG.5 RBD (pink RBD, dark purple mAb) and S2V29-bound BQ.1.1 RBD (gray RBD, light purple mAb). The two CDRH3 residues differing between S2V29 and VIR-7229 (V50Y and N57D) are highlighted in orange. Select residues from the VIR-7229:EG.5 RBD structure are also shown as semi-transparent surfaces colored according to the sticks.
(H) Superposition of the VIR-7229-bound XBB.1.5 RBD (cyan RBD, dark pink mAb) and VIR-7229-bound BA.2.86 S (gold RBD, bright pink mAb) structures highlighting the conservation of electrostatic interactions (dashed lines) at the epitope/paratope interface due to the BA.2.86 R403K mutation. The D52 side chain is weakly resolved in the BA.2.86 S cryoEM density and was therefore not modeled.

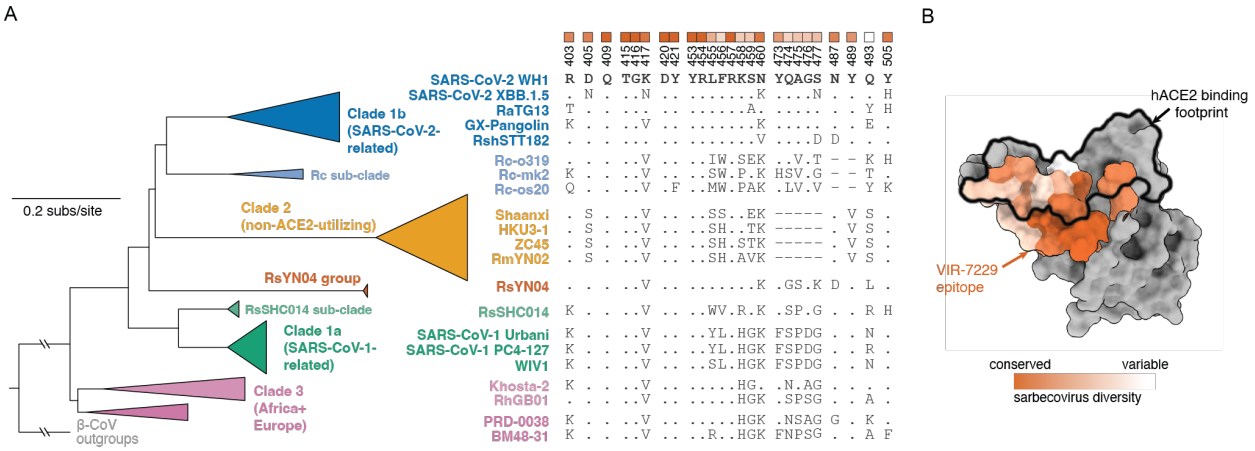


Figure 10.4: VIR-7229 epitope encompasses sarbecovirus diversity.

(A) Cartoon-collapsed sarbecovirus phylogeny (left) with multiple sequence alignment of select sarbecovirus representatives (right) illustrating variation at VIR-7229 epitope positions. RBD numbering is relative to SARS-CoV-2. Dots indicate the SARS-CoV-2 Wuhan-Hu-1 identity. Heatmap at top of alignment illustrates extent of variation (white) or conservation (orange) across the entire sarbecovirus alignment, matched to the structural mapping in panel B. See **Data S1** for full uncollapsed phylogeny and alignment and for full sequences.
(B) Sarbecovirus conservation of the VIR-7229 epitope mapped to SARS-CoV-2 Wuhan-Hu-1 RBD structure (PDB 6M0J). ACE2 binding footprint is illustrated as a black outline.

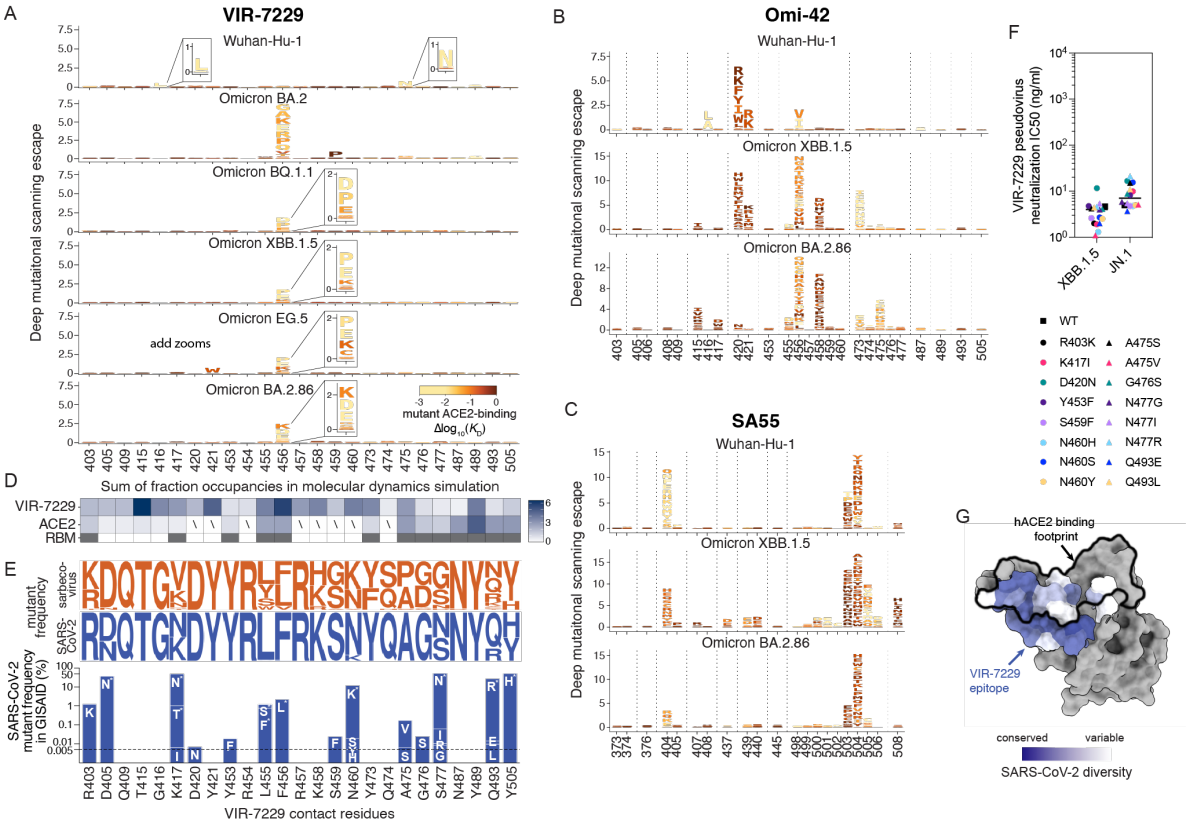


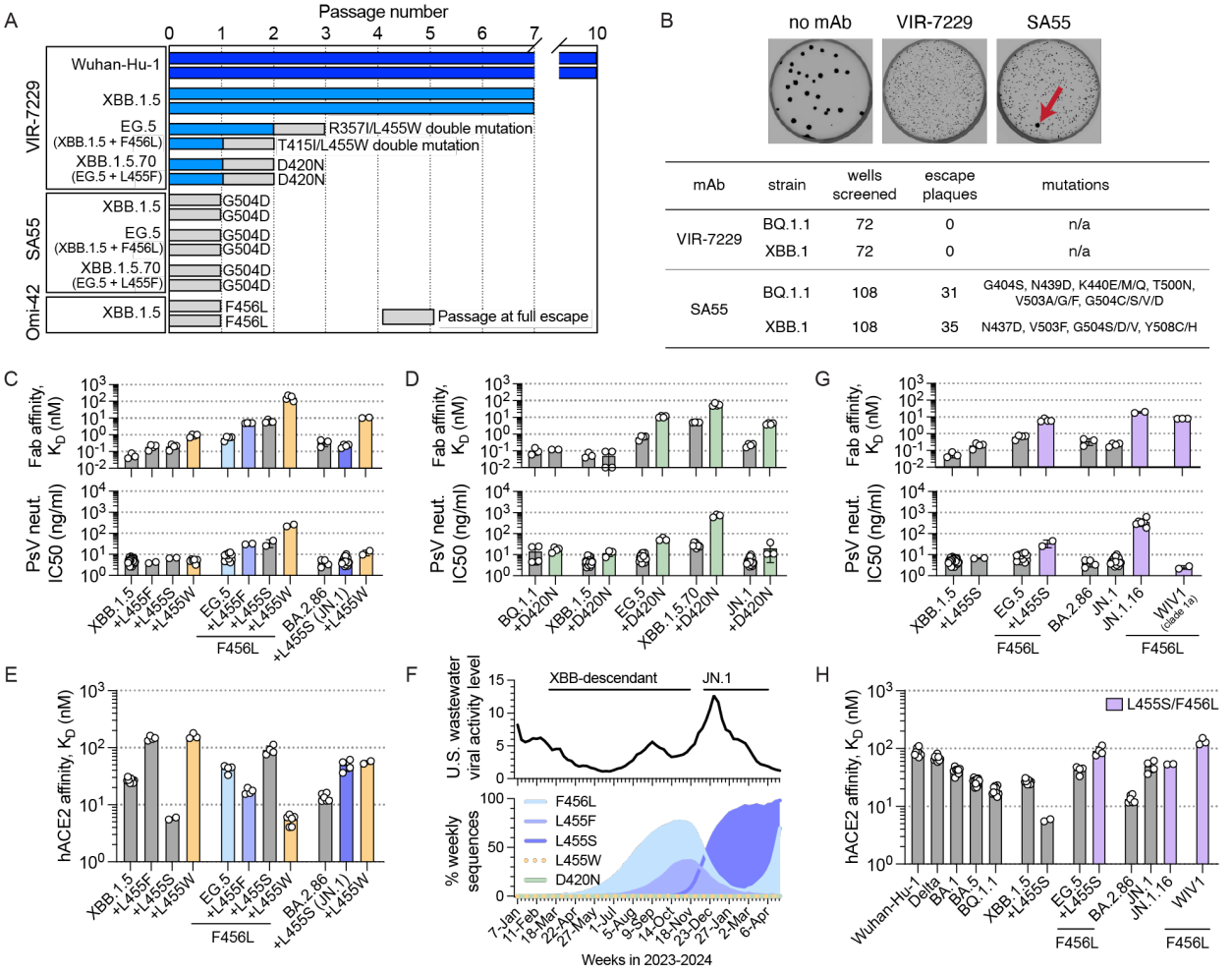
Figure 10.5: VIR-7229 has high tolerance for SARS-CoV-2 epitope variation.

(A-C) Complete elucidation of mutations in the Wuhan-Hu-1, BA.2, BQ.1.1, XBB.1.5, EG.5, and BA.2.86 RBDs that enable escape from VIR-7229 (A), Omi-42 (B), or SA55 (C) binding using a yeast-display deep mutational scanning method. Height of letter is proportional to mutant escape. Letters are colored by their orthogonally measured impacts on ACE2-binding affinity, where lighter yellow indicates increasingly deleterious effects on receptor binding (scale bar, bottom-right). See also **Figs. S5** and **S6**.

(D) Summary of ³4.0-μs MD simulations of XBB.1.5 RBD bound to VIR-7229 or ACE2. Boxes are the number of persistent VIR-7229 or ACE2 contacts at each RBD position in the VIR-7229 epitope, expressed as the fraction occupancy for each VIR-7229 or ACE2 contact across the MD simulation, added together for each RBD position. See panels A or E for RBD position annotations. (Slash indicates no contact formed, i.e. sum of fraction occupancy <0.1.) Full glycans were modeled into the RBD:ACE2 X-ray structure; some ACE2 contacts are glycan-mediated, see **Fig. S4F**. The third row indicates RBM residues (gray boxes), defined as RBD:ACE2 protein:protein contacts (within 5 Å) in the X-ray structure. See also **Fig. S4** and **Data S4**.

(E) Top, logoplots illustrating the frequency of amino acid variation in VIR-7229 epitope residues across human-ACE2-utilizing sarbecovirus sequences (orange) and SARS-CoV-2 sequences available on GISAID from May 8, 2024 (blue). Bottom, barplot illustrating SARS-CoV-2 mutant frequencies (log scale) for all mutants with >0.005% occurrence in GISAID (up to May 8, 2024). VIR-7229 neutralization of each of these mutations was validated via neutralization of single mutants introduced into XBB.1.5 and

JN.1 pseudovirus (**panel F**) or presence of a mutation in a circulating variant that is neutralized (**Fig. 1A**), with the latter mutations labeled with asterisk.
(F) VIR-7229-mediated neutralization of SARS-CoV-2 epitope variants with >0.005% frequency in GISAID (**panel E**), tested on the XBB.1.5 and JN.1 backgrounds. See also **Data S1**.
(G) SARS-CoV-2 conservation of the VIR-7229 epitope mapped to SARS-CoV-2 Wuhan-Hu-1 RBD structure (PDB **6M0J**). ACE2 binding footprint is illustrated as a black outline.

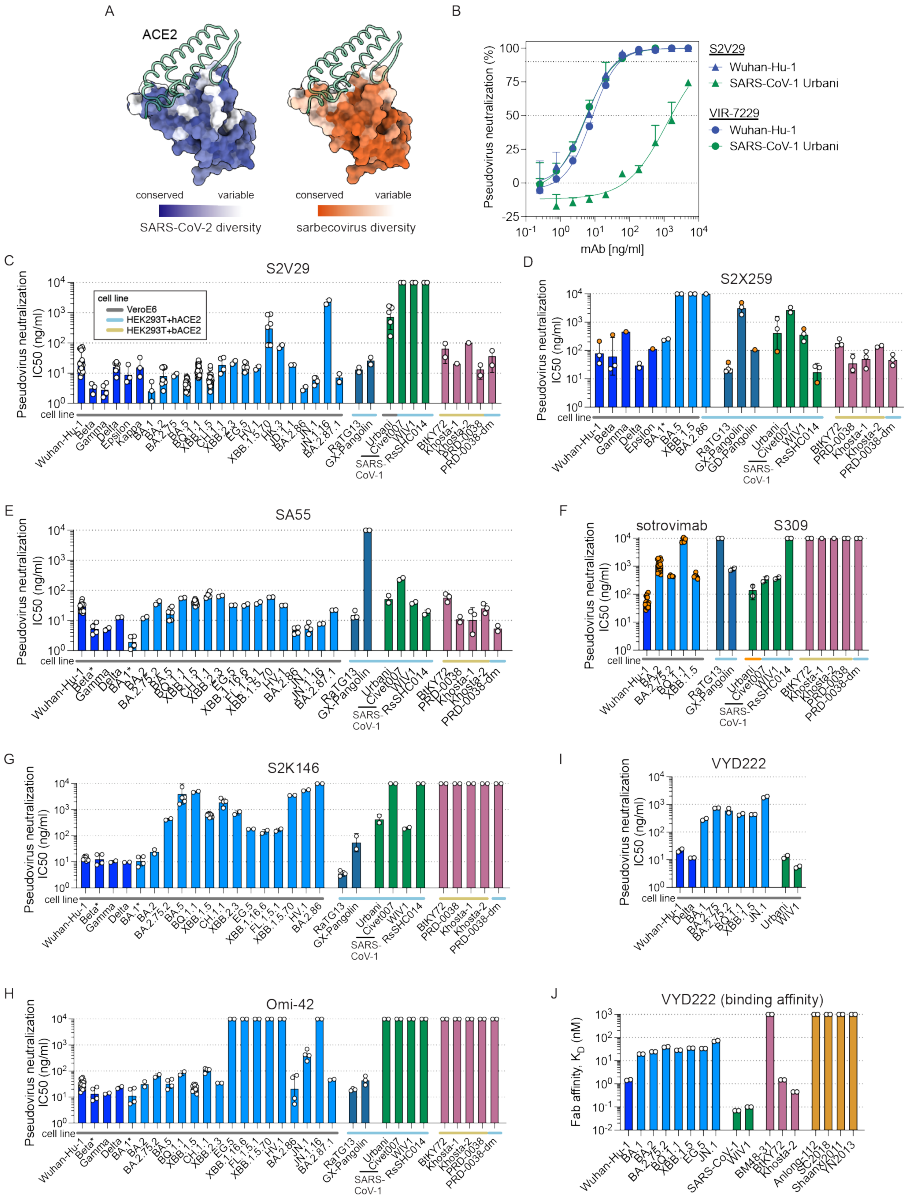


mAb	strain	wells screened	escape plaques	mutations
VIR-7229	BQ.1.1	72	0	n/a
	XBB.1	72	0	n/a
SA55	BQ.1.1	108	31	G404S, N439D, K440E/M/Q, T500N, V503A/G/F, G504C/S/V/D
	XBB.1	108	35	N437D, V503F, G504S/D/V, Y508C/H

Figure 10.6: VIR-7229 exhibits a high barrier to viral escape.

(A) Serial passaging of Wuhan-Hu-1 and XBB.1.5 rVSV in the presence of mAb did not result in escape from VIR-7229 (experiment terminated after 10 and 7 passages, respectively) whereas XBB.1.5 rVSV escaped from comparator mAbs (SA55 and Omi-42) after a single passage. EG.5 rVSV escaped from VIR-7229 after two to three passages, utilizing a double mutation, and from a comparator mAb (SA55) after a single passage. XBB.1.5.70 rVSV escaped from VIR-7229 after two passages and from a comparator mAb (SA55) after a single passage. Two independent replicates were performed for each experiment. See also **Fig. S6** and **Data S5**.

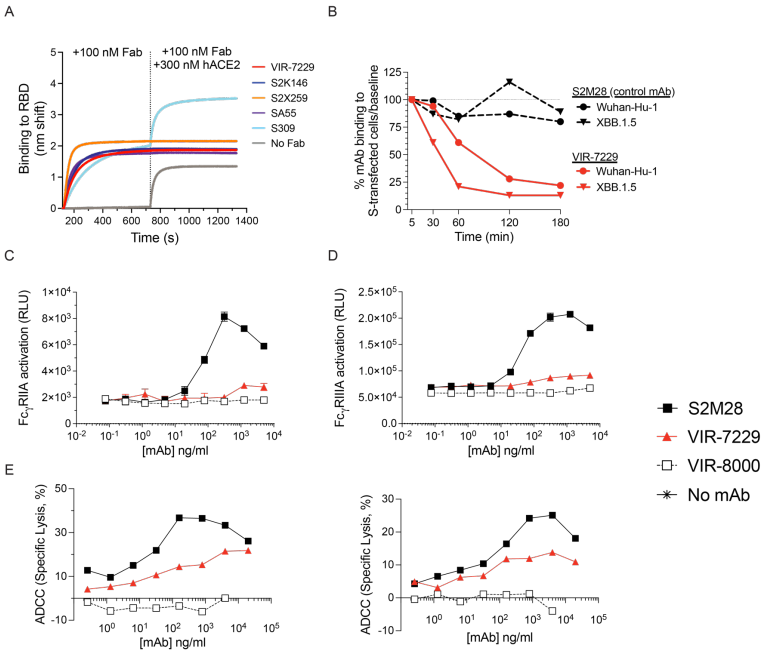
(B) Plaque-based selection of BQ.1.1 and XBB.1 rVSV escapes was performed with VIR-7229 and comparator mAb SA55. Zero escape plaques were observed in 72 independent selections for VIR-7229 whereas 31 and 35 escape plaques, respectively, were observed in 108 independent selections for SA55. Representative images from BQ.1.1 selection are shown, red arrow indicates escape plaque. See also **Data S5**.
(C) Impact of mutations at RBD position 455 on VIR-7229 Fab fragment binding affinity measured by SPR (top) and on VIR-7229-mediated pseudovirus neutralization (bottom). For EG.5+L455F neutralization, the pseudotyped VSV also carries the non-RBD Q52H S mutation (i.e. HK3 strain haplotype). Colored bars correspond to mutations plotted in panel F. See also **Data S1**.
(D) Impact of the D420N mutation on VIR-7229 Fab fragment binding affinity measured by SPR (top) and on VIR-7229-mediated pseudovirus neutralization (bottom). See also **Data S1**.
(E) Impact of mutations at RBD position 455 on ACE2 affinity measured by SPR. Colored bars correspond to mutations plotted in panel F. See also **Data S1**.
(F) Top – SARS-CoV-2 viral activity level in U.S. wastewater, January 2023 – April 2024 (cdc.gov). Bottom – Frequency of SARS-CoV-2 S mutations as percentage of weekly sequences deposited in GISAID, January 2023 – April 2024. L455W and D420N frequencies are too low to be visible. Greater than 99% of L455F and L455W mutations co-occur with F456L, i.e. occur in EG.5 and derivative strains; >98% of L455S mutations are in the BA.2.86 or JN.1 background.
(G) Impact of L455S +/- F456L on VIR-7229 Fab fragment binding affinity measured by SPR (top) and on VIR-7229-mediated pseudovirus neutralization (bottom). See also **Data S1**.
(H) Impact of L455S +/- F456L mutations on ACE2 affinity measured by SPR. See also **Data S1**.



Supplemental Figure 10.1: Potency and breadth of comparator mAbs. Related to Fig. 1.

(**A**) Sequence variation among SARS-CoV-2 sequences in GISAID (blue, left) or sarbecoviruses (orange, right) mapped to the SARS-CoV-2 Wuhan-Hu-1 RBD structure (PDB: 6M0J).
(**B**) Pseudovirus neutralization plotted versus mAb concentration for VIR-7229 and parent mAb S2V29, for Wuhan-Hu-1 and SARS-CoV-1, utilizing VeroE6 cells.
(C-I) Neutralization of SARS-CoV-2 variant or sarbecovirus pseudoviruses mediated by S2V29 parent mAb and six comparator mAbs. Bar colors and horizontal lines as in **Fig. 1A**; orange horizontal line for S309 indicates Vero-TMPRSS2 cells. S309, parent mAb of sotrovimab(Pinto, 2020), and SA55(SA55, 2023) were isolated from SARS-CoV-1 survivors, the latter after receipt of a vaccine based on SARS-CoV-2 Wuhan;

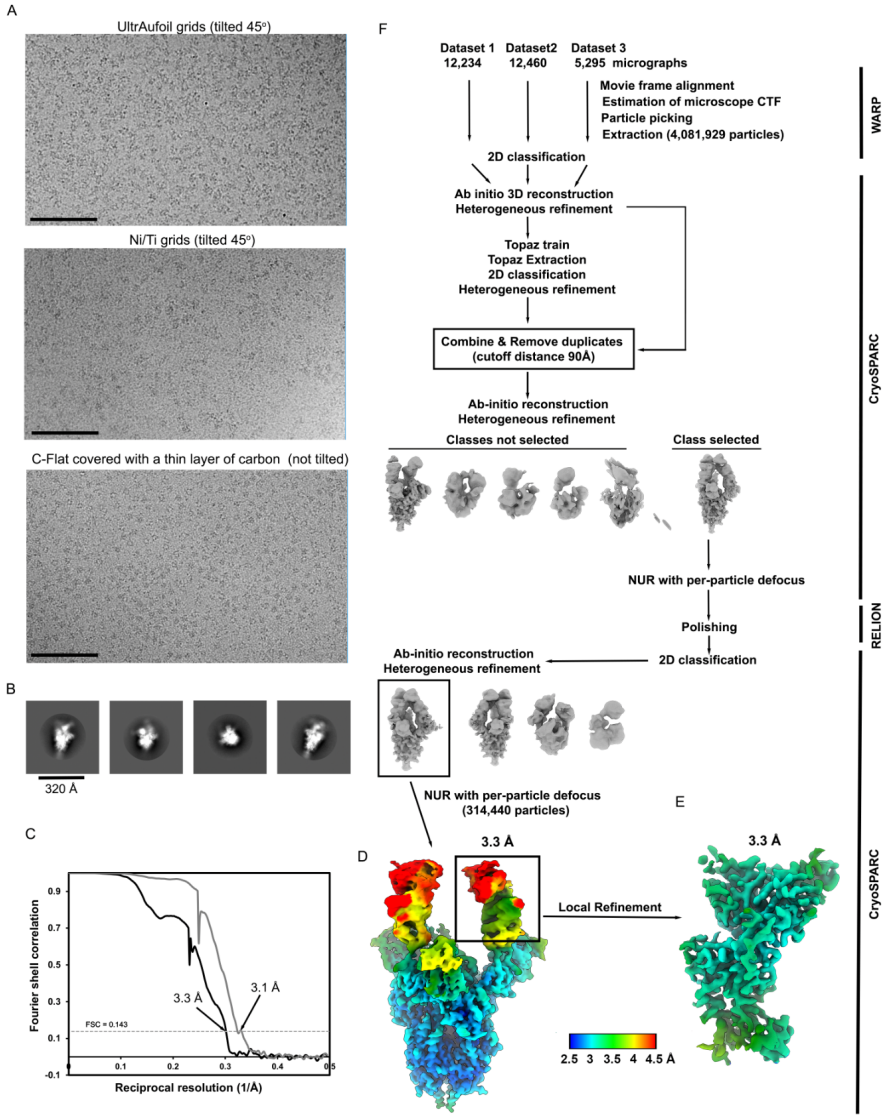
S2K146(Park, 2022) and S2X259(Tortorici, 2021) were isolated from SARS-CoV-2 patients after a pre-Omicron infection; Omi-42(Nutalai, 2022), like S2V29, was isolated from vaccinees following a breakthrough Omicron infection; VYD222 was affinity-matured from the ADI-55688 mAb(Walker, 2024; FDA, 2024; Branyn West). Asterisk indicates maximum neutralization plateaus at ~90% due to entry mediated by VSV-G (i.e. pseudovirus preparations with reduced titer). Orange data points are previously published: S2X259(Tortorici, 2021) and sotrovimab(Addetia, 2023). SARS-CoV-2 strains which completely escape Omi-42 all share the F456L mutation. See also **Data S10.1**.
(J) VYD222 Fab fragment binding affinity to recombinant RBDs measured by SPR. Bar color denotes sarbecovirus clade, as in **Fig. 10.1C**. SARS-CoV-1 RBD is the Urbani strain. See also **Data S10.1**.



Supplemental Figure 10.2: VIR-7229 mechanisms of action. Related to Figs. 1 and 2.

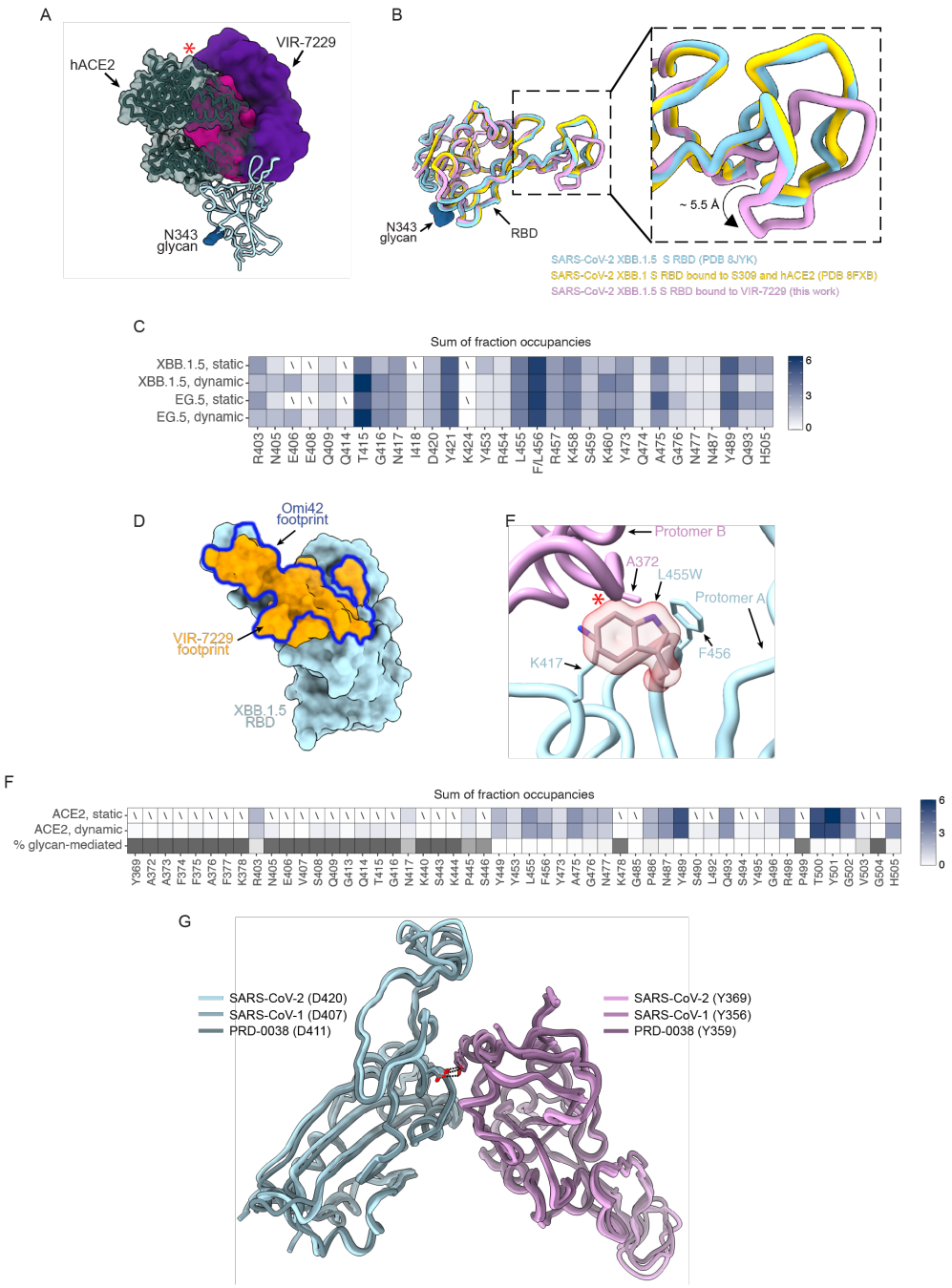
(**A**) VIR-7229 Fab fragment competes with monomeric ACE2 for binding to Wuhan-Hu-1 RBD, as measured by bio-layer interferometry (BLI). All comparator mAbs also compete with ACE2, with the exception of S309.
(**B**) VIR-7229 efficiently triggers S_1 shedding from Wuhan-Hu-1 and XBB.1.5 SARS-CoV-2 S transiently expressed on the surface of Expi-CHO cells, whereas anti-NTD control mAb S2M28 does not.
(C-D) Activation of human FcγRIIa (**C**) and FcγRIIIa (**D**) was evaluated using a bioreporter assay. Target cells were CHO stably expressing SARS-CoV-2 Wuhan-Hu-1 S and effector cells were Jurkat expressing the indicated FcγR and engineered with a NFAT-mediated luciferase reporter to reflect activation of human FcγRs. Data points show means ± SD of duplicates.
(**E**) Antibody-dependent cell cytotoxicity (ADCC; NK-cell mediated) was evaluated using freshly isolated cells from two previously genotyped donors (FcγRIIIa): heterozygous

(F/V158; left) or homozygous high-affinity (V/V158; right). Target cells had surface expression of SARS-CoV-2 Wuhan-Hu-1 S and intracellular expression of HiBiT; ADCC was measured using NanoLuc HiBiT Extracellular Detection Reagent.



Supplemental Figure 10.3: Cryo-EM data processing and validation of VIR-7229-bound BA.2.86 S. Related to Fig. 3.

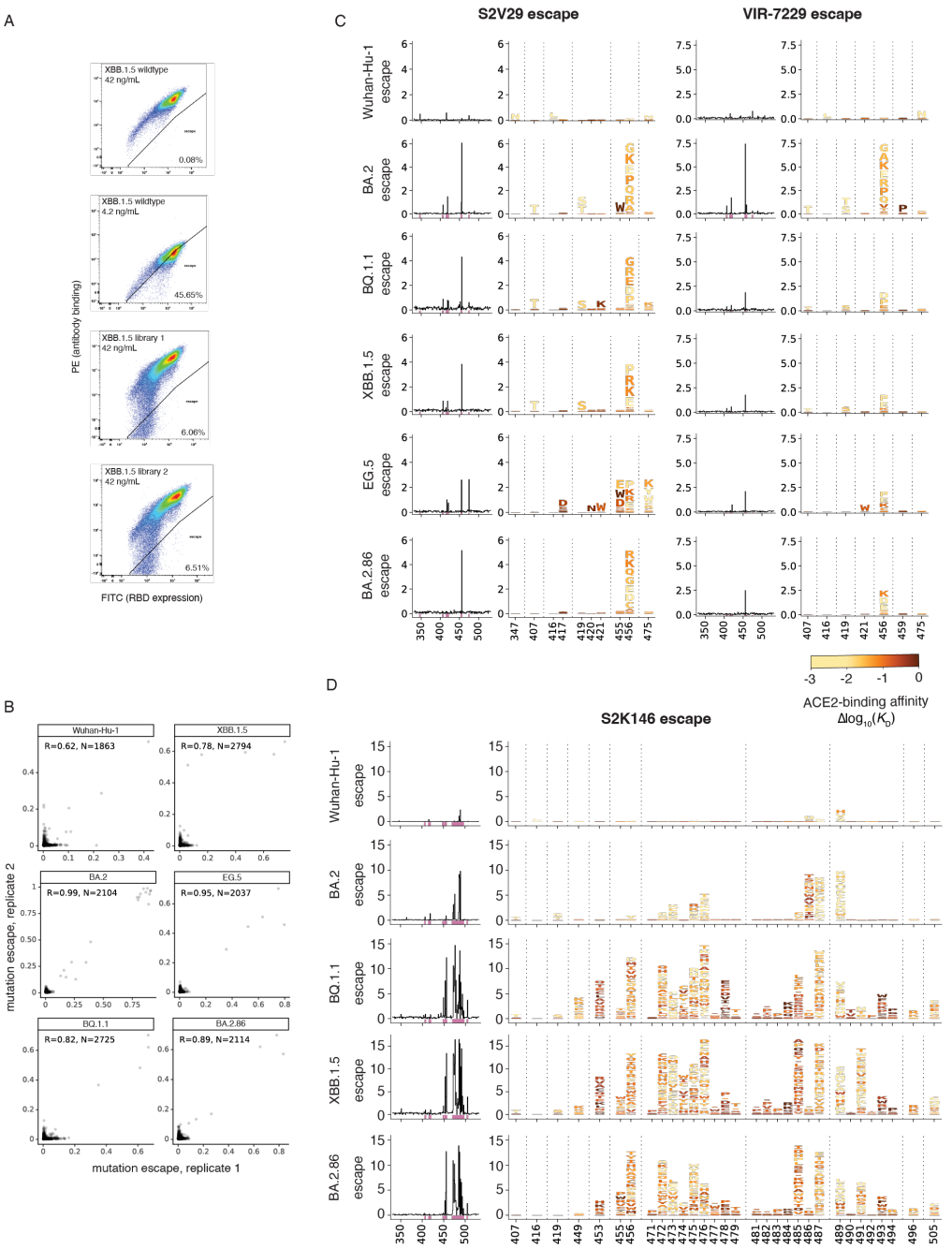
(**A, B**) Representative electron micrographs (**A**) and class averages (**B**) of BA.2.86 S in complex with VIR-7229 Fab. Scale bars, 100 nm (**A**).
(**C**) Gold-standard Fourier shell correlation curves for the S trimer bound to two VIR-7229 Fabs (black line) and the locally refined reconstruction of an RBD and VIR-7229 variable domains (grey line).
(**D, E**) Local resolution map for the S trimer bound to two VIR-7229 Fabs (**D**) and the locally refined reconstruction of an RBD and VIR-7229 variable domains (**E**).
(**F**) Cryo-EM data processing flowchart.



Supplemental Figure 10.4: Structural analysis of VIR-7229:RBD and ACE2:RBD. Related to Figs. 3 and 5.

(**A**) Superposition of VIR-7229-bound and human-ACE2-bound (PDB: 6M0J) SARS-CoV-2 RBD structures; steric clash between VIR-7229 and ACE2 is indicated with a red asterisk.
(**B**) Ribbon diagram of VIR-7229 Fab-bound XBB.1.5 RBD indicating a conformational change of residues 473 to 489 relative to apo XBB.1.5 RBD (PDB:8JYK) and ACE2-bound XBB.1 RBD (PDB: 8FXB; S309 Fab also bound).

(C) Summary of MD simulations of XBB.1.5 RBD or EG.5 RBD bound to VIR-7229 Fab (dynamic) as compared to analysis of X-ray structures (static). Boxes indicate the sum of fraction occupancies of VIR-7229 contacts to each RBD residue, as in **Fig. 5B**. Contacts in the X-ray structure are treated as 100% occupancy. Contacts in the MD simulation beyond the static epitope are shown if the sum of fraction occupancies is ³0.1. Slash indicates no contact formed. See also **Data S4**.
(D) To illustrate the similarity of VIR-7229 and Omi-42 epitopes, VIR-7229 epitope is showed in orange on XBB.1.5 RBD structure and Omi-42 binding footprint is depicted as a blue outline.
(E) In silico modeling of L455W in a fully closed SARS-CoV-2 S structure (PDB 7K43; with S2M11 Fab bound) indicating expected clashes with a neighboring protomer. All energetically favored rotamers are sterically incompatible with the closed structure due to clashes with either the same or the neighboring protomer; one of the three most prevalent rotameric configurations was selected for visualization purposes. The two SARS-CoV-2 S protomers are shown in cyan and pink with the modeled L455W side chain highlighted as a red semi-transparent surface and steric clash indicated with a red asterisk.
(F) Summary of MD simulation of XBB.1.5 RBD bound to ACE2 (dynamic) as compared to analysis of X-ray structure (static). Boxes indicate the sum of fraction occupancies of ACE2 contacts to each RBD residue; full glycans were modeled on the RBD:ACE2 structure, MD contacts may be glycan mediated, as indicated in the third row (darker gray indicates a larger percentage glycan-mediated). Contacts in the MD simulation beyond the static epitope are shown if the sum of fraction occupancies is ³0.1. Slash indicates no contact formed. See also **Data S4**.
(G) Zoomed-in view of the interface between two closed RBDs within a S trimer emphasizing the inter-protomer hydrogen bond formed between SARS-CoV-2 D420 and Y369 (PDB 6ZGE), clade 1a SARS-CoV-1 D407 and Y356 (PDB 7ZH1), or clade 3 PRD-0038 D411 and Y359 (PDB 8U29).

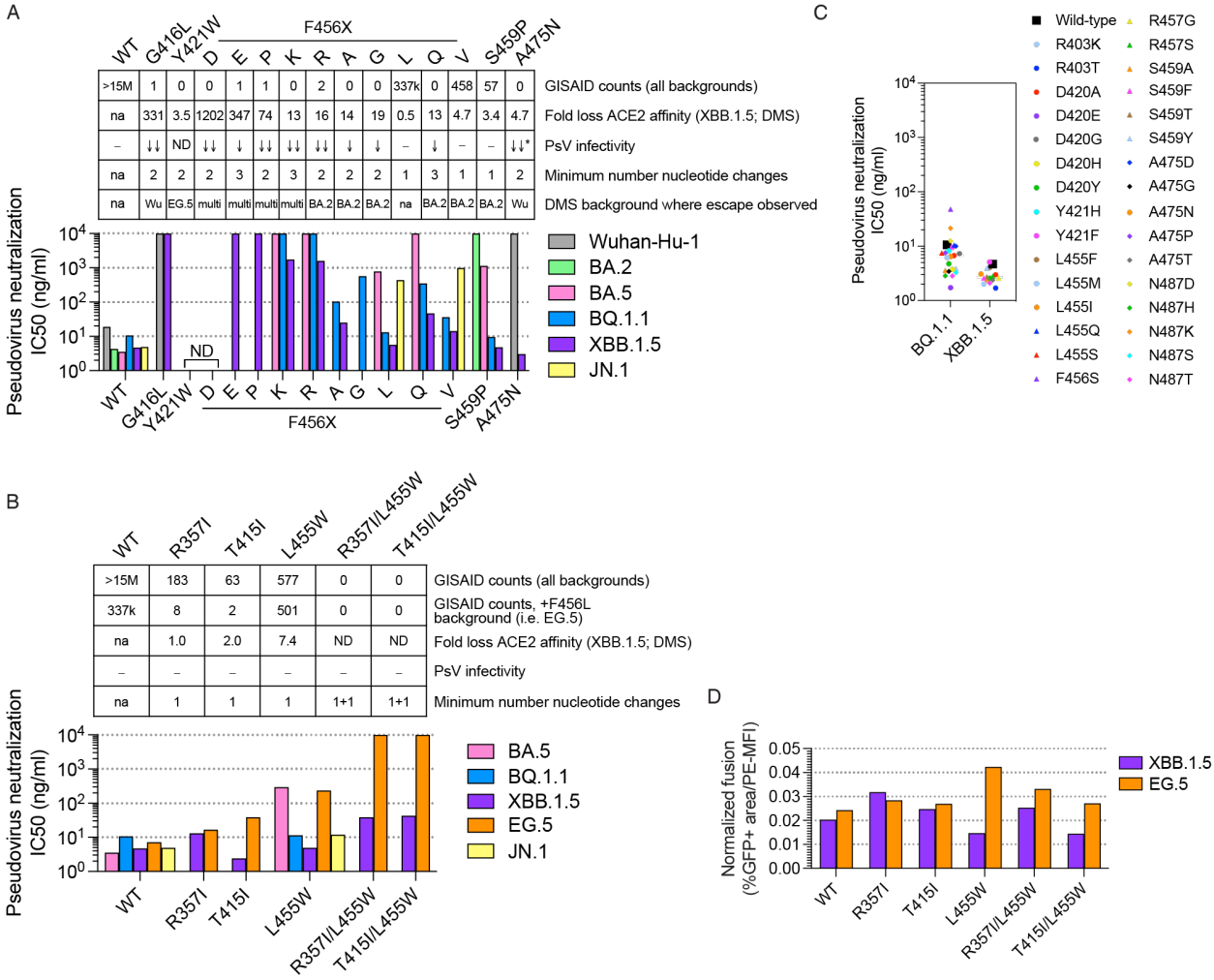


Supplemental Figure 10.5: Deep mutational scanning profiling of VIR-7229, S2V29, and S2K146 escape mutations. Related to Fig. 5.

(A) Representative FACS gates used to identify mutations that escape antibody binding. An antibody-escape gate was drawn that captures approximately 50% of the cells in the respective wildtype control labeled at 0.1x the library selection antibody concentration. The “escape fraction” represents the fraction of cells of a mutant genotype that fall into this antibody-escape FACS gate.

(B) For each experiment with VIR-7229, the correlation in the per-mutation escape fraction between duplicate library selections.

(C) Full deep mutational scanning escape profiles of parental mAb S2V29 (left) and VIR-7229 (right), compared to **Fig. 5A** that illustrates VIR-7229 epitope profiles. For each experiment, lineplots (left) show the total escape at each site in the RBD, with sites of strong escape annotated with pink indicators. Logoplots (right) illustrate mutation-level escape fraction at sites of strong escape, with mutations colored according to mutational impact on ACE2-binding affinity. Note, mutations to T/S at sites 407 and 419 introduce N-linked glycosylation motifs due to the presence of N405 and N417 in the Omicron (but not Wuhan-Hu-1) variants.
(D) Deep mutational scanning escape profiling of the comparator mAb S2K146 (details as in (C)) illustrating a broadening of the functional epitope over evolutionary time, likely due to erosion of S2K146 affinity across variant evolution.



Supplemental Figure 10.6: Validation and fitness of VIR-7229 escape mutations. Related to Figs. 5 and 6.

(A) VIR-7229-mediated neutralization of SARS-CoV-2 pseudoviruses carrying mutations that were observed as DMS binding escapes in at least one strain background, plus F456L. Mutations were tested in different strain backgrounds, as indicated by bar color. Mutations are annotated by: total counts in the GISAID database as of May 8, 2024, fold-change reduction in ACE2 affinity in the XBB.1.5 background (evaluated by DMS[55], available from https://tstarrlab.github.io/SARS-CoV-2-RBD_DMS_Omicron-XBB-BQ/RBD-heatmaps/), impact on pseudovirus infectivity (- is typical range, ‾ is on the low end of typical, ‾‾ is a large decrease; assessed qualitatively by evaluating XX following infection by virus with normalized MOI, and comparison to the unmutated backbone), minimum number of nucleotide changes required for the mutation to occur, and the RBD background where the mutation was observed as a binding escape in the DMS experiment. Asterisk indicates that the reduced infectivity observed for A475N is only observed in the Wuhan-Hu-1 background. ND: Not determined; na: not applicable. See also **Data S1**.
(B) VIR-7229-mediated neutralization of SARS-CoV-2 pseudoviruses carrying mutations that were observed during the EG.5 rVSV resistance selection experiment. Strain background is indicated by bar color. For L455W, JN.1 is also the BA.2.86 background (JN.1 = BA.2.86+L455S). Mutations are annotated as in panel A, as well as by counts in the +F456L background in the GISAID database as of May 8, 2024. See also **Data S1**.
(C) VIR-7229-mediated neutralization of SARS-CoV-2 pseudoviruses carrying a set of epitope mutations below 0.005% frequency in GISAID but accessible by a single nucleotide change from wild-type sequences. Mutations were tested in either the BQ.1.1 or XBB.1.5 background; some mutations were tested in both backgrounds. See also **Data S1**.
(D) Cell-cell fusion between VeroE6-TMPRSS2 and SARS-CoV-2 S-expressing cells, normalized by SARS-CoV-2 S expression. The same panel of SARS-CoV-2 S variants were tested as in panel B.

Table S 10.1: CryoEM data collection and refinement statistics

	BA.2.86 S:VIR-7229 PDB EMD	BA.2.86 S:VIR-7229 Local refinement RBD:VH/VL PDB 9ASD EMD-43813
Data collection and processing		
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	54	54
Defocus range (μm)	0.8-2.0	0.8-2.0
Pixel size (Å)	0.835	0.835
Symmetry imposed	C1	C1
Initial particle images (no.)	4,081,929	4,081,929

Final particle images (no.)	314,440	314,440
Map resolution (Å)	3.1	3.3
FSC threshold	0.143	0.143
Validation		
MolProbity score	1.1	1.13
Clashscore	1.84	1.64
Poor rotamers (%)	0.19	0.85
Ramachandran plot		
Favored (%)	97.15	96.58
Allowed (%)	2.68	3.42
Disallowed (%)	0.17	0

Table S10.2: Crystallographic data collection and refinement statistics.

	EG.5 RBD/ VIR-7229/S2H97 PDB 9ATM	XBB.1.5 RBD/ VIR-7229/S309 PDB 9AU1	BQ.1.1 RBD/ S2V29/S2H97 PDB 8S6M
Wavelength	1.195 Å	1.195 Å	1.00 Å
Resolution range	46.78 - 1.9 (1.92 - 1.9)	44.98 - 2.41 (2.45 - 2.41)	47.35 - 1.67 (1.69 - 1.67)
Space group	P 21 21 21	P 21 21 21	P 21 21 21
Unit cell	48.728 149.698 167.078 90 90 90	73.726 112.525 170.314 90 90 90	148.175 48.099 166.526 90 90 90
Total reflections	1277227 (33897)	734265 (36747)	941473 (16037)
Unique reflections	97295 (3168)	55433 (2722)	138930 (4582)
Multiplicity	13.1 (10.7)	13.3 (13.5)	6.8 (3.5)
Completeness (%)	99.91 (99.69)	99.88 (99.82)	99.96 (100.00)
Mean I/sigma I	12.8 (0.6)	13.2 (0.5)	15.5 (0.7)
Wilson B-factor	36.44	60.65	29.32
Reflections used in refinement	97295 (3017)	55433 (2599)	138930 (4349)
Reflections used for R-free	4870 (151)	2757 (123)	6987 (233)
R-work	0.1939 (0.4012)	0.2268 (0.4255)	0.1839 (0.3597)
R-free	0.2273 (0.4276)	0.2574 (0.4369)	0.2027 (0.3562)
CC(work)			
CC(free)			

Number of non-hydrogen atoms	8689	8190	8673
macromolecules	7977	8078	7844
ligands	36	42	59
solvent	676	70	770
Protein residues	1070	1070	1067
RMS(bonds)	0.007	0.008	0.006
RMS(angles)	0.87	0.96	0.87
Ramachandran favored (%)	98.19	95.72	98.08
Ramachandran allowed (%)	1.71	3.9	1.83
Ramachandran outliers (%)	0.10	0.38	0.1
Rotamer outliers (%)	0.47	2.36	0.73
Clashscore	2.18	3.89	1.9
Average B-factor	53.96	76.42	45.76
macromolecules	54.16	76.40	45.78
ligands	70.71	87.15	68.20
solvent	50.70	71.79	43.77

Table S10.3: Bioinformatic analysis of intra-individual SARS-CoV-2 genomic variability.

Nucleotide trimer change	Amino acid change	Genome-wide trimer mean mutation rate [%]	Standard error	Prevalence [%] of variant trimer at residue 455 in all backgrounds (GISAID)[1]	Prevalence [%] of variant trimer at residue 455 in F456L background (GISAID)[1]
TGT > TTT	L455F	0.0394	0.0086	0.40	18.51
TGC > TTC	L455F	0.0271	0.0053	0.25	12.65
TTG > TCG	L455S	0.0153	0.0014	0.31	0.04

TGT > TCT	L455F	0.0119	0.0014	0.02	1.19
TGC > TCC	L455F	0.0076	0.0015	4.1e-05	2.1e-03
A**TT** > AGT	L455V	0.0034	6.0e-04	3.4e-04	6.9e-04
TTG > TGG	L455W	0.0023	4.0e-04	3.7e-03	0.16
A**TT** > AAT	L455M	0.0021	4.0e-04	1.4e-04	0

The table depicts the in silico estimated mutation rate for all the nucleotide changes (in the context of their surrounding nucleotides) that can lead to an amino acid change at residue 455. Mean mutation rate and standard error are estimated from whole viral genomes of N=1763 samples (see details in Methods). Note, if the nucleotide change does not occur in the second nucleotide of the 455 amino acid codon, the trimer is not representative of the codon (e.g. for L455F, the variant occurs on the 3rd nucleotide of the 455 codon, the trimer therefore represents the last 2 nucleotides of the 455 codon and the first nucleotide of the 456 codon; for L455V and L455M, the variant occurs on the 1st nucleotide of the 455 codon, the trimers therefore represent the last nucleotide of the 454 codon and the first two nucleotides of the 455 codon). Nucleotides that are part of the wild-type Leucine codon are in bold. See also **Data S5**.

[1]GISAID prevalence is for sequences deposited up to January 17, 2024. As these results are based on the consensus sequences (as opposed to intra-individual variability), the result reflects viral fitness, immune pressure, transmission rate, as well as collection and submission biases (e.g. geographical).

CHAPTER 11. STRUCTURAL BASIS OF SARS-COV-2 OMICRON IMMUNE EVASION AND RECEPTOR ENGAGEMENT.

At the time of this next story we were two years into the COVID-19 pandemic, and several variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) had arisen that show increased infectivity or evade immunity. The Omicron variant of concern had 37 mutations in the spike protein. In this story I detail the omicron mutations which fall largely in the two domains targeted by neutralizing antibodies, the receptor-binding domain (RBD), and the N-terminal domain (NTD). I describe our structures of the viral spike bound to S309, a therapeutic antibody that maintains neutralizing activity against Omicron, and show how Omicron retains high-affinity binding to ACE2 while greatly reducing binding to other therapeutic antibodies.

Adapted From: Matthew McCallum *et al.* Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement.*Science***375**,864-868(2022).DOI:10.1126/science.abn8652.

11.1 Chapter Introduction

Although sequential COVID-19 waves have swept the world, no variants have accumulated mutations and mediated immune evasion to the extent observed for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Omicron (B.1.1.529) variant of concern (VOC). This variant was first identified in late November 2021 in South Africa and was quickly designated a VOC by the World Health Organization (*Viana, 2022*). Omicron has spread worldwide at a rapid pace relative to previous

SARS-CoV-2 variants (*McCallum, 2021; McCallum 2021*). The Omicron spike (S) glycoprotein, which promotes viral entry into cells (*Walls, 2020; Wrapp, 2020*), harbors 37 residue mutations in the predominant haplotype relative to Wuhan-Hu-1 S (*Walls, 2020*), whereas SARS-CoV-2 Alpha and Delta VOC display ~10 substitutions (*McCallum, 2021; Collier, 2021*). The Omicron receptor-binding domain (RBD) and N-terminal domain (NTD) contain 15 and 11 mutations, respectively, which lead to severe dampening of plasma-neutralizing activity in previously infected or vaccinated individuals (*Cameroni, 2021; Cao, 2021; Liu, 2021; Planas, 2021; VanBlargan, 2021*). Although the Omicron RBD harbors 15 residue mutations, it binds to the human ACE2 entry receptor with high affinity and can efficiently recognize mouse ACE2 (*Cameroni, 2021*). As a result of this antigenic shift, the only authorized or approved therapeutic monoclonal antibodies (mAbs) with neutralizing activity against Omicron are S309 (sotrovimab parent) and the COV2-2196/COV2-2130 cocktail (cilgavimab/tixagevimab parents). Even these mAbs had reduced potency (by a factor of 2 to 3 and by a factor of 12 to 200, respectively), according to pseudovirus or authentic virus assays (*Cameroni, 2021; Cao, 2021; Liu, 2021; Planas, 2021; VanBlargan, 2021*). This extent of evasion of humoral responses has important consequences for therapy and prevention of both the current pandemic and future pandemics.

11.2 Results

To define the molecular mechanisms involved in Omicron immune evasion and altered receptor recognition, we determined cryo–electron microscopy (cryo-EM) structures of the prefusion-stabilized SARS-CoV-2 Omicron S ectodomain trimer bound to S309 and S2L20 (NTD-specific mAb) Fab fragments (**Fig. 11.1, Fig. S11.1, and**

table S11.1) and the x-ray crystal structure of the Omicron RBD in complex with human ACE2 and the Fab fragments of S309 and S304 at 2.85 Å resolution (**Table S11.2**). S309 recognizes antigenic site IV (*Pinto, 2020*), whereas S304 binds to site IIc (*Piccoli, 2020*) and was used to assist crystallization. Furthermore, we used surface plasmon resonance (SPR) to evaluate the binding of clinical mAbs to the Omicron RBD and S ectodomain trimer.

Three-dimensional classification of the cryo-EM data revealed the presence of two conformational states with one (45% of selected particles) or two (55% of selected particles) RBDs in the open conformation, for which we determined structures at 3.1 Å and 3.2 Å resolution, respectively (**Fig. 11.1, A and B, figs. S11.1 and S11.2, and Table S11.1**). The larger fraction of open RBDs, relative to the apo (*Walls, 2020; Wrapp, 2020*) and S309-bound (*Pinto, 2020*) Wuhan-Hu-1 S ectodomain trimer structures, could result from the Omicron mutations, the prefusion-stabilizing mutations (*Hsieh, 2020; Olmedillas, 2021*), or S2L20 binding. Focused classification and local refinement of the S309-bound RBD (domain B) and of the S2L20-bound NTD (domain A) were used to account for their conformational dynamics and led to improved local resolution of these regions (to 3.0 and 3.3 Å resolution, respectively).

Whereas most VOCs have only a few mutations beyond the NTD, RBD, and furin cleavage site regions, the Omicron spike harbors eight substitutions outside of these areas: T547K, H655Y, N764K, D796Y, N856K, Q954H, N969K, and L981F, which could all be modeled in the map (**Fig. 11.1, A and B, and Fig. 11.2**). Three of these mutations introduce interprotomer electrostatic contacts between the S_2 and S_1 subunits: N764K binds Gln^{314}(in domain D), Ser^{982} binds T547K (in domain C of protomers with closed

RBDs), and N856K binds Asp^{568} and Thr^{572} (in domain C, the former residue is closer to N856K in protomers with closed RBDs) (**Fig. 11.2**) (*Walls, 2016; Tortorici, 2019*). Furthermore, N969K forms interprotomer electrostatic contacts with Q755, and L981F improves intra-protomer hydrophobic packing in the prefusion conformation (**Fig. 11.2**). The latter mutation is close to the prefusion-stabilizing 2P mutations (K986P and V987P) used in all three vaccines deployed in the US (**Fig. 11.2**). Consistent with recent studies (*Cong, 2021; Meng, 2021; Sato, 2022*), S_1 shedding might be reduced by enhanced interactions between the S_1 and S_2 subunits in Omicron S along with prefusion stabilization, or by altered processing at the S_1-S_2 cleavage site due to the N679K and P681H mutations. Dampened S_1 subunit shedding might enhance the effector function activity of vaccine- or infection-elicited Abs along with that of therapeutic mAbs (*Tortorici, 2021*) that retain affinity for Omicron S.

The Omicron NTD carries numerous mutations, deletions (del), and an insertion (ins) including A67V, del69-70, T95I, G142D, del143-145, del211, L212I, and ins214EPE (**Fig. 11.1C and Fig. S11.3**). Many of these mutations have been described in previously emerged VOCs: del69-70 was found in Alpha, T95I was present in Kappa and Iota, and G142D was present in Kappa and Delta. T95I, del211, L212I, and ins214EPE are outside the NTD antigenic supersite but in the vicinity of the epitope targeted by the P008_056 mAb, which suggests that these mutations could putatively modulate recognition of similar mAbs or have another functional relevance. Although the region comprising del143-145 is weakly resolved in the map, it is expected to alter antibody recognition as a result of the introduced sequence register shift, whereas G142D is incompatible with binding of several potent NTD-neutralizing mAbs (such as

S2X333) because of steric hindrance (**Fig. 11.1D**) (*McCallum, 2021; McCallum, 2021*). Moreover, del143-145 is reminiscent of Alpha del144, which was also isolated as an escape mutation in the presence of mAb S2X333 and led to viral breakthrough in a hamster challenge model (McCallum, 2021). These data suggest that G142D and del143-145 account for the observed SARS-CoV-2 Omicron evasion from neutralization mediated by a panel of NTD mAbs (*Cameroni, 2021; Liu, 2021*).

The RBD is the main target of plasma-neutralizing activity in convalescent and vaccinated individuals and comprises several antigenic sites recognized by neutralizing Abs with a range of neutralization potencies and breadth (*Pinto, 2020; Piccoli, 2020; Tortorici, 2021; Stamatatos, 2021; Starr, 2021; Barnes, 2020; Jette, 2021; Barnes, 2020; Zost, 2020; Martinez, 2021; Dong, 2021; Greaney, 2021; Hansen, 2020; Rappazzo, 2021; Wec, 2020; Park, 2022; Bowen, 2021*) (**Fig. 11.3A**). Our structures provide a high-resolution blueprint of the residue substitutions found in this variant (**Fig. 11.3B**) and their impact on binding of clinical mAbs (**Table 11.1**). Several individual mutations or subsets of mutations occurring in the Omicron RBD have been reported to affect neutralizing antibody binding or neutralization (*Corti, 2021*). The K417N, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y. and Y505H mutations are part of antigenic site I, which is immunodominant in previous variants (*Piccoli, 2020; Starr, 2021*). K417N, E484A, S477N, and Q493R would lead to loss of electrostatic interactions and steric clashes with REGN10933 whereas G446S would lead to steric clashes with REGN10987, consistent with the dampened binding to the Omicron RBD and S trimer (**Fig. 11.3, C and D, Fig. S11.3, and table S11.3**) and with previous analyses of the impact of individual mutations on neutralization by each of these two

mAbs (*Liu, 2021; Tada, 2021; Yuan, 2021; Wang, 2021*). Moreover, N440K was reported to dampen REGN10987 neutralization severely (*Liu, 2021*). Reduced binding of the Omicron RBD to COV2-2196 and COV2-2130, relative to the Wuhan-Hu-1 RBD, likely results from T478K [based on Delta S (McCallum, 2021)], Q493R, and putatively S477N for COV2-2196, as well as G446S and E484A for COV2-2130 (**Fig. 11.3, E and F**, **Fig. S11.4, and table S11.3**). Integrating these data with neutralization assays suggests that although each point mutation only imparts a small reduction of COV2-2196– or COV2-2130–mediated neutralization (*Liu, 2021*), the constellation of Omicron mutations leads to more pronounced loss of activity (*Cameroni, 2021; Cao, 2021; Liu, 2021; Planas; VanBlargan, 2021*). E484A abrogates electrostatic interactions with LY-CoV555 heavy and light chains, whereas Q493R would prevent binding through steric hindrance (**Fig. 11.3G, Fig. S11.4, and table S11.3**), as supported by neutralization data (*Liu, 2021*). K417N is expected to negatively affect the constellation of electrostatic interactions formed between the Omicron RBD and LY-CoV16 heavy chain, thereby abolishing binding (**Fig. 11.3H, Fig. S11.4, and table S11.3**) and neutralization of single-mutant S pseudoviruses (*Liu, 2021; Wang, 2021; Starr, 2021*)). Furthermore, S477N and Q493R have been shown to dampen binding of and neutralization mediated by LY-CoV16 (*Liu, 2021; Starr, 2021*). Finally, K417N, E484A, and Q493R hinder CT-P59 engagement through a combination of steric hindrance and remodeling of electrostatic contacts, thereby preventing binding (**Fig. 11.3I, Fig. S11.4, and tables S11.1 and S11.3**).

The SARS-CoV-2 Omicron G339D and N440K mutations are within or near antigenic site IV, which is recognized by the S309 mAb (*Pinto, 2020*). Nonetheless,

relative to Wuhan-Hu-1 pseudovirus or Washington-1 authentic virus, S309 undergoes only a factor of 2 to 3 reduction of neutralizing activity against Omicron (*Cameroni, 2021;Liu, 2021; Planas. 2021; VanBlargan, 2021*). The Lys side chain introduced by the N440K substitution points away from the S309 epitope and does not affect binding. The Asp side chain introduced by the G339D substitution does not interfere with the S309 epitope, although not all rotamers are compatible with mAb binding (**Fig. S11.2**). This finding likely explains the similarly moderate reduction of S309 potency against the single G339D S mutant (*Liu, 2021* or the full constellation of Omicron S mutations (*Cameroni, 2021;Liu, 2021; Planas. 2021; VanBlargan, 2021*). The modest reduction of the Omicron RBD binding to S309 (**Fig. 11.3J, Fig. S11.4, and table S11.3**) mirrors the reduced neutralization potency of this VOC (by a factor of 2 to 3 relative to ancestral viruses) and concurs with deep-mutational scanning analysis of individual mutations on S309 recognition (*Starr, 2021*). Overall, the S309 binding mode remains unaltered by the Omicron mutations, including recognition of the N343 glycan (**Fig. S11.5**).

The Omicron RBD is structurally similar to the Wuhan-Hu-1 RBD, and both structures can be superimposed with an RMSD of 0.8 Å over 183 aligned Cα residues [as compared to PDB 6m0j (*Lan, 2020*)]. However, the region comprising residues 366 to 375, which harbors the S371L/S373P/S375F substitutions, deviates markedly from the conformation observed for the Wuhan-Hu-1 RBD, irrespective of the presence of bound linoleic acid (*Walls, 2020; Lan, 2020; Toelzer, 2020*). Although this region is weakly resolved in the cryo-EM and x-ray structures, the conformation adopted in the latter structure is incompatible with binding of some cross-reactive site II mAbs such as S2X35, consistent with our observation of dampened binding (**Fig. S11.6**). We therefore

propose that these mutations participate in rendering this region of the RBD dynamic and mediate immune evasion from some site II mAbs.

We recently reported that the SARS-CoV-2 Omicron RBD binds human ACE2 with a factor of ~2.4 enhanced affinity relative to the Wuhan-Hu-1 RBD (*Cameroni, 2021*). Our crystal structure of the human ACE2-bound Omicron RBD elucidates how the constellation of RBD mutations found in this VOC affect receptor recognition (**Fig. 11.4A and table S11.2**). The N501Y mutation alone enhances ACE2 binding to the RBD by a factor of 6 relative to the Wuhan-Hu-1 RBD, as reported for the Alpha variant (*Collier, 2021*), likely as a result of increased shape complementarity between the introduced Tyr side chain and the ACE2 Tyr^{41} and Lys^{353} side chains (**Fig. 11.4B**). Omicron S residue Tyr^{501} and ACE2 residue Tyr^{41} form a T-shaped π-π stacking interaction, as previously observed for an N501Y-harboring S structure in complex with ACE2 (*Zhu, 2021*). The K417N mutation dampens receptor recognition by a factor of ~3 (*2*, *6*, *39*, *45*) likely through loss of a salt bridge with ACE2 Asp^{30}(**Fig. 11.4C**). The Q493R and Q498R mutations introduce two new salt bridges with $Glu^{35}$and Glu^{38}, respectively, replacing hydrogen bonds formed with the Wuhan-Hu-1 RBD and thereby remodeling the electrostatic interactions with ACE2 (**Fig. 11.4D**). Both of these individual mutations were reported to reduce ACE2 binding avidity slightly by deep-mutational scanning studies of the yeast-displayed SARS-CoV-2 RBD (*Starr, 2020*). Finally, S477N leads to formation of new hydrogen bonds between the introduced Asn side chain and the ACE2 Ser^{19} backbone amine and carbonyl groups (**Fig. 11.4E**). Collectively, these mutations have a net enhancing effect on binding of the Omicron RBD to human ACE2 relative to Wuhan-Hu-1, which suggests that structural epistasis

enables immune evasion while retaining efficient receptor engagement. The large number of Omicron mutations in the immunodominant receptor-binding motif likely explains a substantial proportion of the loss of neutralization by convalescent and vaccine-elicited polyclonal antibodies, and is in line with the known plasticity of this subdomain (*Starr, 2021*).

11.3 Discussion

Although the N501Y mutation has previously been described as enabling some SARS-CoV-2 VOCs to infect and replicate in mice, the Alpha and Beta variant RBDs only weakly bound mouse ACE2 (*Shuai, 2021; Pan, 2021*). The SARS-CoV-2 Omicron RBD, however, interacts more strongly with mouse ACE2 than do the Alpha and Beta variant RBDs when evaluated side-by-side (**Fig. S11.7A**) and can use mouse ACE2 as an entry receptor for S-mediated entry (*Cameroni, 2021; Hoffmann, 2022*). We propose that the Q493R mutation plays a key role in enabling efficient mouse ACE2 binding, which occurs through formation of a new electrostatic interaction with the Asn^{31} side chain amide (Lys^{31} in human ACE2); this is supported by in silico modeling based on our human ACE2-bound crystal structure (**Fig. S11.7B**). These findings concur with the emergence and fixation of the Q493K RBD mutation upon serial passaging in mice to yield a mouse-adapted virus designated SARS-CoV-2 MA10 (*Leist*, *2020*).

This work defines the molecular basis for the broad evasion of humoral immunity exhibited by SARS-CoV-2 Omicron and underscores the SARS-CoV-2 S mutational plasticity as well as the importance of targeting conserved epitopes in the design and development of vaccines and therapeutics. The S309 mAb (the parent of sotrovimab) neutralizes Omicron with one-half to one-third the potency with which it neutralizes

Wuhan-Hu-1 or Washington-1, whereas the seven other clinical mAbs or mAb cocktails undergo reduction of neutralizing activity of one to two orders of magnitude or greater. Furthermore, some Omicron isolates (≈9%) harbor the R346K substitution, which, in conjunction with N440K (present in the main haplotype), leads to escape from C135 mAb-mediated neutralization (*Barnes, 2020*; *Weisblum, 2020*). R346K does not affect S309 whether in isolation or in the context of the full constellation of Omicron mutations; hence, mAbs targeting antigenic site IV can be differently affected by Omicron (*Cameroni, 2021; Liu, 2021; Starr, 2020*). Whereas C135 was identified from a SARS-CoV-2 convalescent donor (*Barnes, 2020*), S309 was isolated from a subject who recovered from a SARS-CoV infection in 2003 (*Pinto, 2020*); the latter strategy increased the likelihood of finding mAbs that recognize epitopes that are mutationally constrained throughout sarbecovirus evolution. The identification of broadly reactive mAbs that neutralize multiple distinct sarbecoviruses, including SARS-CoV-2 variants, paves the way for designing vaccines that elicit broad sarbecovirus immunity (*Walls, 2021; Walls, 2020; Arunachalam, 2021; Martinez, 2021; Cohen, 2021*). These efforts offer hope that the same strategies that contribute to solving the current pandemic will prepare us for possible future sarbecovirus pandemics.

11.4 Methods

Production of recombinant spike glycoprotein

The SARS-CoV-2 S Omicron ectodomain contains the Omicron specific mutations A67V, Δ69-70, T95I, G142D/Δ143-145, Δ211/L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K,

D796Y, N856K, Q954H, N969K, and L981F, in addition to VFLIP stabilizing mutations F817P, A892P, A899P, A942P, V987P, Y707C, and T883C. The SARS-CoV-2 S HexaPro ectodomain was previously described (*Hsieh, 2020*). Both S trimers were produced in 200 mL cultures of Expi293F Cells (ThermoFisher Scientific) grown in suspension using Expi293 Expression Medium (ThermoFisher Scientific) at 37°C in a humidified 8% CO2 incubator rotating at 130 rpm. Cells grown to a density of 2.5 million cells per mL were transfected using the ExpiFectamine 293 Transfection Kit (ThermoFisher Scientific) and cultivated for 2-4 days at which point the supernatant was harvested. S ectodomains were purified from clarified supernatants using a Cobalt affinity column (Cytiva, HiTrap TALON crude), washing with 20 column volumes of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl or 20 mM sodium phosphate pH 8 and 100 mM NaCl and eluted with a gradient of 500 mM imidazole. The S ectodomain was then concentrated using a 100 kDa centrifugal filter (Amicon Ultra 0.5 mL centrifugal filters, MilliporeSigma), residual imidazole was washed away by consecutive dilutions in the centrifugal filter unit with 20 mM Tris-HCl pH 8.0 and 150 mM NaCl, and finally concentrated to 1 mg/mL before use immediately after purification. The various RBDs used for BLI were based on a previously reported construct (*Walls, 2020*) and produced and biotinylated as previously described (*McCallum, 2021; Collier, 2021; Starr, 2020*).

CryoEM sample preparation and data collection

100 µL of 1 mg/mL SARS-CoV-2 S B.1.1.529 ectodomain was incubated with 40 µl 3.4 mg/mL S309 Fab for 10 min at 37°C in 150 mM NaCl and 20 mM Tris-HCl pH 8 and then 2.2 µL of 67 mg/mL S2L20 Fab was added and the mixture was incubated for 15 min at 37°C. Unbound Fab was then washed away with six consecutive dilutions in

400 µL of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl over a 100 kDa centrifugal filter (Amicon Ultra 0.5 mL centrifugal filters, MilliporeSigma). The complex was concentrated to 3.5 mg/mL and 3 µL was immediately applied onto a freshly glow discharged 2.0/2.0 UltraFoil grid (84) (200 mesh), plunge frozen using a vitrobot MarkIV (ThermoFisher Scientific) using a blot force of −1 and 6.0 s blot time at 100% humidity and 23°C.

Data were acquired using the Leginon software (*Suloway, 2005*) to control a FEI Titan Krios transmission electron microscope equipped with a Gatan K3 direct detector and operated at 300 kV with a Gatan Quantum GIF energy filter. The dose rate was adjusted to 3.75 counts/super- resolution pixel/s, and each movie was acquired in 75 frames of 40 ms with a pixel size of 0.843 Å and a defocus range comprised between −0.2 and −2.0 µm.

CryoEM data processing

Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking and extraction (with a downsampled pixel size of 1.686 Å and box size of 256 pixels2) were carried out using Warp (*Tegunov, 2019*). Reference-free 2D classification was performed using cryoSPARC (*Punjani, 2017*) to select well-defined particle images. 3D classification with 50 iterations each (angular sampling 7.5˚ for 25 iterations and 1.8˚with local search for 25 iterations) were carried out using Relion without imposing symmetry (*Zivanov, 2018; Scheres, 2012*). 3D refinements were carried out using non- uniform refinement in cryoSPARC (*Punjani, 2020*) before particle images were subjected to Bayesian polishing using Relion *(Zivanov, 2019*) during which particles were re-extracted with a box size of 512 Å at a pixel size of 0.843 Å. Next, 86 optics groups were defined based on the beam tilt angle used for data

collection. Another round of non-uniform refinement in cryoSPARC was then performed concurrently with global and per-particle defocus refinement. For focused classification of the NTD, particles were symmetry-expanded and 3D classified in Relion without alignment using a mask that encompasses the NTD and the S2L20 VH/VL region. Focused classification of the RBD and the S309 VH/VL region was done on one of the two closed RBDs found in the asymmetric reconstruction with a single RBD open as described for the NTD/S2L20 but without symmetry expansion. Particles in well-formed 3D classes were then used for local refinement in cryoSPARC.

Reported resolutions are based on the gold-standard Fourier shell correlation of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (*Rosenthal, 2003; Chen, 2013*).

CryoEM model building and analysis

UCSF Chimera (*Pettersen, 2004*) and Coot (*Emsley, 2010*) were used to fit atomic models of S2L20, S309, and SARS-CoV-2 S (PDB 7SOB) into the cryo-EM maps. The model was then refined and rebuilt into the map using Coot (*67*), Rosetta (*Wang, 2016; Frenz, 2019*), Phenix (*Liebschner, 2019*), and ISOLDE (*Croll, 2018*). Model validation and analysis used MolProbity (*Chen, 2010*), EMRinger (*Barad, 2015*), Phenix (*Liebschner, 2019*), and Privateer (*Agirres, 2015*). Figures were generated using UCSF ChimeraX (Goddard, 2018)**.**

Monoclonal antibodies

Antibody VH and VL sequences for mAbs COV2-2130 (PDB ID 7L7E), COV2-2196 (PDB ID 7L7E, 7L7D), REGN10933 (PDB ID 6XDG), REGN10987 (PDB ID 6XDG) and ADI-58125 (PCT application WO2021207597, seq. IDs 22301 and 22311)

were subcloned into heavy chain (human IgG1) and the corresponding light chain (human IgKappa, IgLambda) expression vectors respectively and produced in transiently transfected ExpiCHO-S cells (Thermo Fisher, #A29133) at 37°C and 8% CO2. Cells were transfected using ExpiFectamine. Transfected cells were supplemented 1 day after transfection with ExpiCHO Feed and ExpiFectamine CHO Enhancer. Cell culture supernatant was collected eight days after transfection and filtered through a 0.2 µm filter. Recombinant antibodies were affinity purified on an ÄKTA Xpress FPLC device using 5 mL HiTrap™ MabSelect™ PrismA columns followed by buffer exchange to Histidine buffer (20 mM Histidine, 8% sucrose, pH 6) using HiPrep 26/10 desalting columns. Antibody VH and VL sequences for LY-CoV555, LY-CoV016, and CT-P59 were obtained from PDB IDs 7KMG, 7C01 and 7CM4, respectively and mAbs were produced as recombinant IgG1 by ATUM. S309 was produced by WuXi Biologics (China). Recombinant S304 and S309 Fabs were produced by ATUM.

Recombinant protein production

SARS-CoV-2 RBD proteins for SPR binding assays (residues 328-531 of S protein from GenBank NC_045512.2 with N-terminal signal peptide and C-terminal thrombin cleavage site-TwinStrep-8xHis-tag) were expressed in Expi293F (Thermo Fisher Scientific) cells at 37°C and 8% CO2.Transfections were performed using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatants were collected three days after transfection and supplemented with 10x PBS to a final concentration of 2.5x PBS (342.5 mM NaCl, 6.75 mM KCl and 29.75 mM phosphates). RBDs were purified using cobalt-based immobilized metal affinity

chromatography followed by buffer exchange into PBS using a HiPrep 26/10 desalting column (Cytiva) for Wuhan-Hu-1 protein, or a Superdex 200 Increase 10/300 GL column (Cytiva) for Omicron protein. SARS-CoV-2 Omicron RBD for crystallization (residues 328-531, with N-terminal signal peptide and 'ETGT', and C-terminal 8xHis-tag) was expressed similarly as described above in the presence of 10 μM kifunensine. Cell culture supernatant was collected four days after transfection and supplemented with 10x PBS to a final concentration of 2.5x PBS. Protein was purified using a HiTrap TALON crude cartridge followed by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva).

ACE2 for crystallization (residues 19-615 from Uniprot Q9BYF1 with a C-terminal thrombin cleavage site-TwinStrep-10xHis-GGG-tag, and N-terminal signal peptide) was expressed in ExpiCHO cells in the presence of 10 μM kifunensine at 37°C and 8% CO2. Transfection was performed using the ExpiCHO transfection kit (Thermo Fisher Scientific). Cell culture supernatant was collected eight days after transfection and supplemented to a final concentration of 80 mM Tris-HCl pH 8.0, 100 mM NaCl, and then incubated with BioLock (IBA GmbH) solution. hACE2 was purified using a 5 mL StrepTrap HP column (Cytiva) followed by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva) pre-equilibrated in PBS.

SPR binding measurements of IgG

SPR binding measurements were performed using a Biacore T200 instrument with a CM5 sensorchip covalently immobilized with Cytiva Human Antibody Capture Kit (RBD binding) or Cytiva His Capture Kit (S binding). Running buffer was Cytiva HBS-EP+ (pH 7.4). All measurements were performed at 25 °C. RBD analyte concentrations

were 3.1, 12.5, and 50 nM, run as single-cycle kinetics. IgG analyte concentrations for S binding experiments were 11, 33, 100, and 300 nM, run as single-cycle kinetics. For RBD binding, double reference-subtracted data were fit usingBiacore T200 Evaluation software (version 3.1) to a 1:1 binding model except two datasets (LY-CoV016 and REGN10987 binding to Wuhan-Hu-1 RBD) were fit to a Heterogeneous Ligand model ("biphasic fit") due to an artefactual kinetic phase with very slow dissociation that often arises when RBD is an analyte. For the Heterogeneous Ligand fits, the lower affinity of the two KD values reported by the fit is taken as the KD (the two KD values are separated by at least two orders of magnitude). Binding of the Omicron RBD to mAbs has been carried out once (n=1) except for S309 which has been performed three times (n=3). Wuhan-Hu-1 RBD binding data are representative of at least two (n=2) for S309, CT-P59, LY-CoV555, LY-CoV16, REGN10987 and REGN10933 mAbs and once for all others (n=1). S binding data were evaluated qualitatively, rather than quantitatively, due to the avidity of the bivalent analyte precluding the use of a 1:1 binding model, and due to the anticipated differential impact of avidity on binding in the SPR experiment as compared to the biological context (e.g. S density and orientation on the CM5 matrix of the chip surface likely does not recapitulate S density and orientation on the surface of a virus or infected cell).

Crystallization, data collection, structure determination, and analysis

Prior to forming the SARS-CoV-2 Omicron RBD-ACE2-S304-S309 complex, recombinant SARS-CoV-2 Omicron RBD was digested with EndoH (New England Biolabs, 21 units/μg RBD) at 4°C overnight. Recombinant hACE2 protein was digested using EndoH (New England Biolabs, 25 units/μg ACE2) and thrombin (Sigma-Aldrich, 1

unit/75 µg ACE2) at 4°C overnight. RBD was mixed with a 1.2-fold molar excess of EndoH deglycosylated hACE2, and a 1.1-fold molar excess of S304 Fab and S309 Fab. The complex was purified on a Superdex 200 10/300 GL column pre-equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl. Crystals of the SARS-CoV-2 Omicron RBD-hACE2-S304-S309 complex were obtained at 20°C by sitting drop vapor diffusion. A total of 200 nL of the complex at 6.5 mg/mL were mixed with 200 nL mother liquor solution containing 0.1 M NDSB-256, 20% v/v ethylene glycol, 10% w/v PEG 8000, and 0.1 M Tris (base)/bicine pH 8.5.

Data were collected at the Molecular Biology Consortium beamline 4.2.2 at the Advanced Light Source synchrotron facility in Berkeley, CA and processed with the XDS software package yielding a final dataset of 2.85 Å in space group P21. The SARS-CoV-2 Omicron RBD-hACE2- S304-S309 complex structure was solved by molecular replacement using Phaser (*McCoy, 2007*) from starting models consisting of RBD-S304-S309 (PDB: 7JX3) and hACE2 (PDB: 6m0j). Several subsequent rounds of model building and refinement were performed using Coot (*Emsley, 2010*), ISOLDE (*Croll, 2018*), Refmac5 (*Murshudov, 2011*), Phenix (*Liebschner, 2019*) and MOE (https://www.chemcomp.com), to arrive at a final model of the quaternary complex.

11.5 Figures and Tables

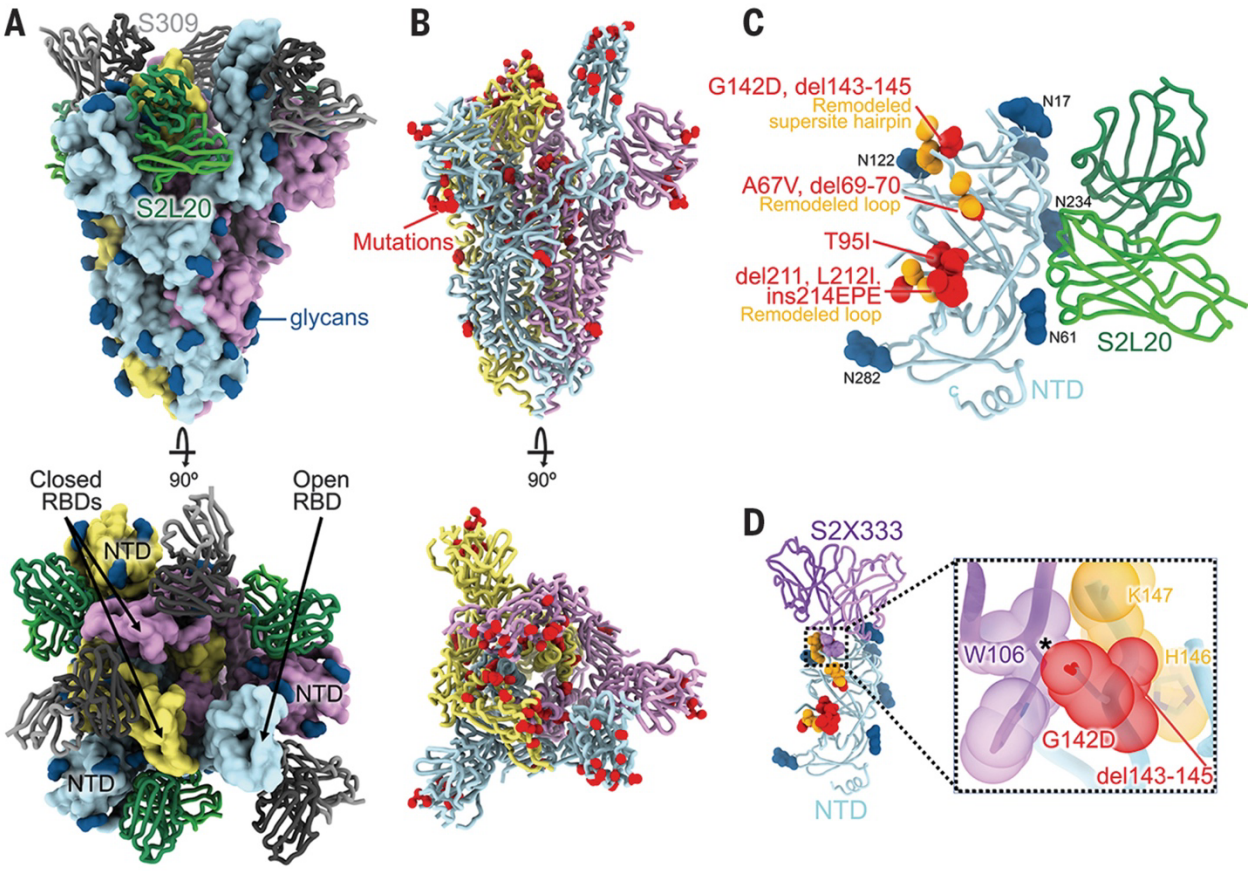


Figure 11.1: Cryo-EM structure of the SARS-CoV-2 Omicron S trimer reveals a remodeling of the NTD antigenic supersite.

(**A**) Surface rendering in two orthogonal orientations of the Omicron S trimer with one open RBD bound to the S309 (gray) and S2L20 (green) Fabs shown as ribbons. The three S protomers are colored light blue, pink, and gold. N-linked glycans are shown as dark blue surfaces. (**B**) Ribbon diagrams in two orthogonal orientations of the S trimer with one open RBD. Omicron residues mutated relative to Wuhan-Hu-1 are shown as red spheres (except D614G, which is not shown). (**C**) The S2L20-bound Omicron NTD with mutated, deleted, or inserted residues rendered or indicated as red spheres. Segments with notable structural changes are shown in orange and labeled. (**D**) Zoomed-in view of the Omicron NTD antigenic supersite overlaid with the S2X333 Fab [used here as an example of a prototypical NTD-neutralizing mAb (22)], highlighting the binding incompatibility; the modeled clash between S2X333 Trp^{106} and NTD G142D is indicated with an asterisk.

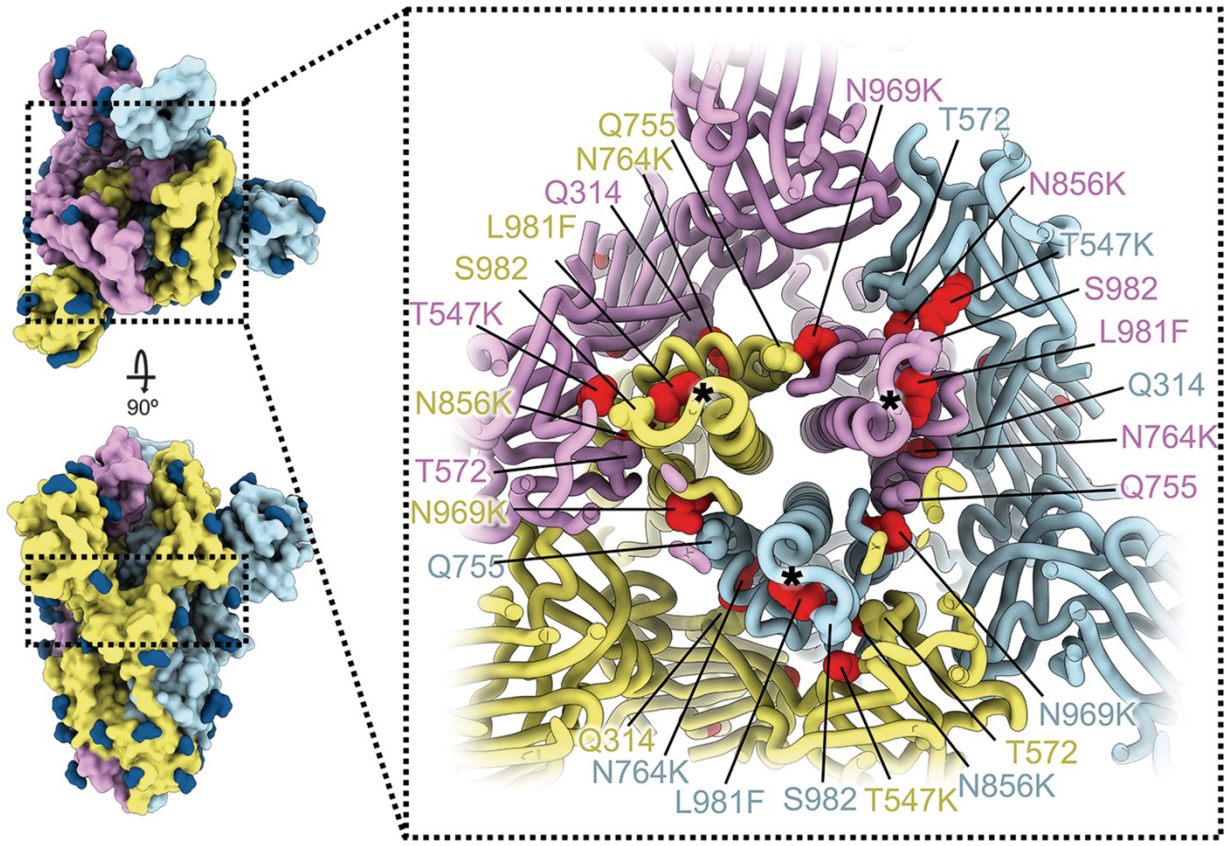


Figure 11.2: SARS-CoV-2 Omicron S mutations outside the NTD and RBD.

Ribbon diagram shows a cross section of the Omicron S glycoprotein (the location of this slice on the S trimer is indicated at left). Mutated residues T547K, N764K, N856K, N969K, and L981F are shown as red spheres, whereas the residues they contact are shown as spheres colored as the protomer to which they belong. Black asterisks show the positions of residues involved in the prefusion-stabilizing 2P mutations (K986P and V987P) used in all three vaccines deployed in the US. The three S protomers are colored light blue, pink, and gold. N-linked glycans are shown as dark blue surfaces.

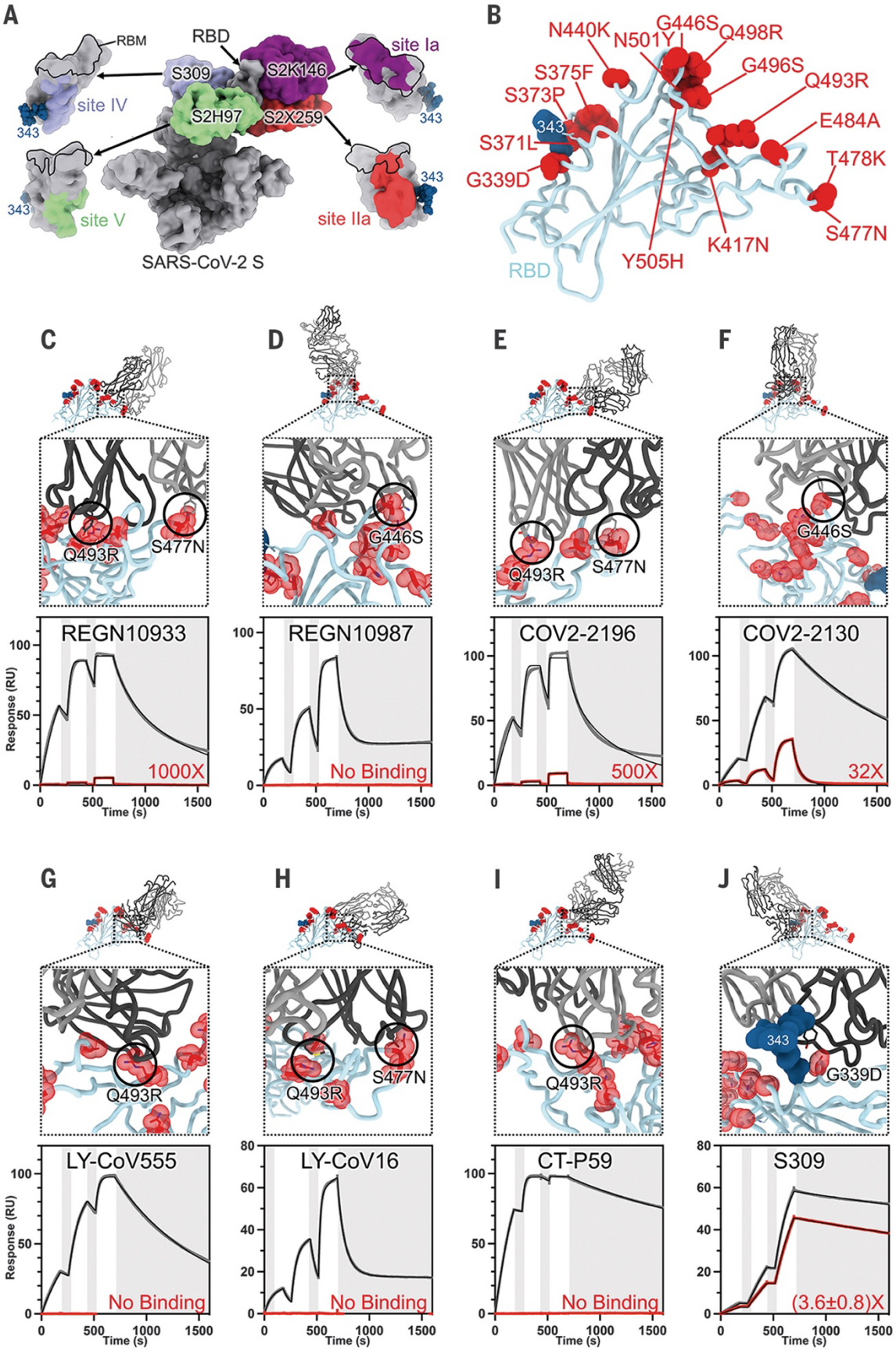

A
RBM
RBD
S309
S2K146
site Ia
site IV
S2H97
S2X259
343
343
343
343
site V
site IIa
SARS-CoV-2 S
B
N440K
G446S
N501Y
Q498R
G496S
S375F
S373P
Q493R
343
E484A
S371L
T478K
G339D
S477N
K417N
Y505H
RBD
C
D
E
F
Q493R
S477N
G446S
G446S
Q493R
S477N
G446S
REGN10933
REGN10987
COV2-2196
COV2-2130
1000X
No Binding
500X
32X
Response (RU)
Time (s)
G
H
I
J
Q493R
Q493R
S477N
Q493R
343
G339D
LY-CoV555
LY-CoV16
CT-P59
S309
No Binding
No Binding
No Binding
(3.6±0.8)X

Figure 11.3: SARS-CoV-2 Omicron RBD mutations promote escape from a panel of clinical mAbs.

(**A**) RBD antigenic map as determined in (*13*). (**B**) Ribbon diagram of the RBD crystal structure, with residues mutated relative to the Wuhan-Hu-1 RBD shown as red spheres. The N343 glycan is rendered as blue spheres. (**C** to **J**) Zoomed-in views of the Omicron RBD (blue) superimposed on structures of clinical mAbs (gray). Selected residues that interfere with the following mAbs are circled: (C) REGN10933, (D) REGN10987, (E) COV2-2196, (F) COV2-2130, (G) LY-CoV555, (H) LY-CoV16, (I) CT-P59, and (J) S309, which does not clash with G339D. All panels were rendered with the crystal structure except (J), which was generated with the cryo-EM model. Binding of the Wuhan-Hu-1 (gray line) or Omicron (red line) RBD to the corresponding mAb was evaluated using surface plasmon resonance (single-cycle kinetics) and is shown underneath each structural superimposition. White and gray stripes are association and dissociation phases, respectively. The thin black line is a fit to a kinetic model. The decrease in affinity between Wuhan-Hu-1 and Omicron binding is indicated in red. Results are consistent with immunoglobulin G binding to S ectodomains (Fig. S3).

mAbs based on structural analyses.

	REGN10933	REGN10987	COV2-2196	COV2-2130	LY-CoV555	LY-CoV016	CT-P59	S309	ADI-58125	Total GISAID counts**	Omicron counts
G339D										30360	29448
S371L										28467	28201
S373P										28897	28246
S375F										28753	28214
K417N	X					X	X			50411	7342
N440K										17194	7396
G446S		X		x						8566	7528
S477N	x		X			x				100335	28275
T478K			X							3,579,418	28431
E484A	X			X	X		X			30299	28275
Q493R	X		X	x	X	x	X			28921	28420
G496S							X			29090	28374
Q498R							X			28746	28385
N501Y										1,300,039	28410
Y505H									X*	28,589	28228
PDB ID	6XDG	6XDG	7L7D	7L7E	7KMG	7C01	7CM4	This study	n/a		

* For ADI-58125, the impact on binding of C, N, and S substitutions is shown at position Y505 based on mutagenesis studies Belk, J., Deveau, L. M., Rappazzo, C. G., Walker,

L. & Wec, A. WO2021207597 - COMPOUNDS SPECIFIC TO CORONAVIRUS S PROTEIN AND USES THEREOF. *ADAGIO THERAPEUTICS, INC.* (2021).
**As of December 24, 2021; excluded entries with >5% Ns.

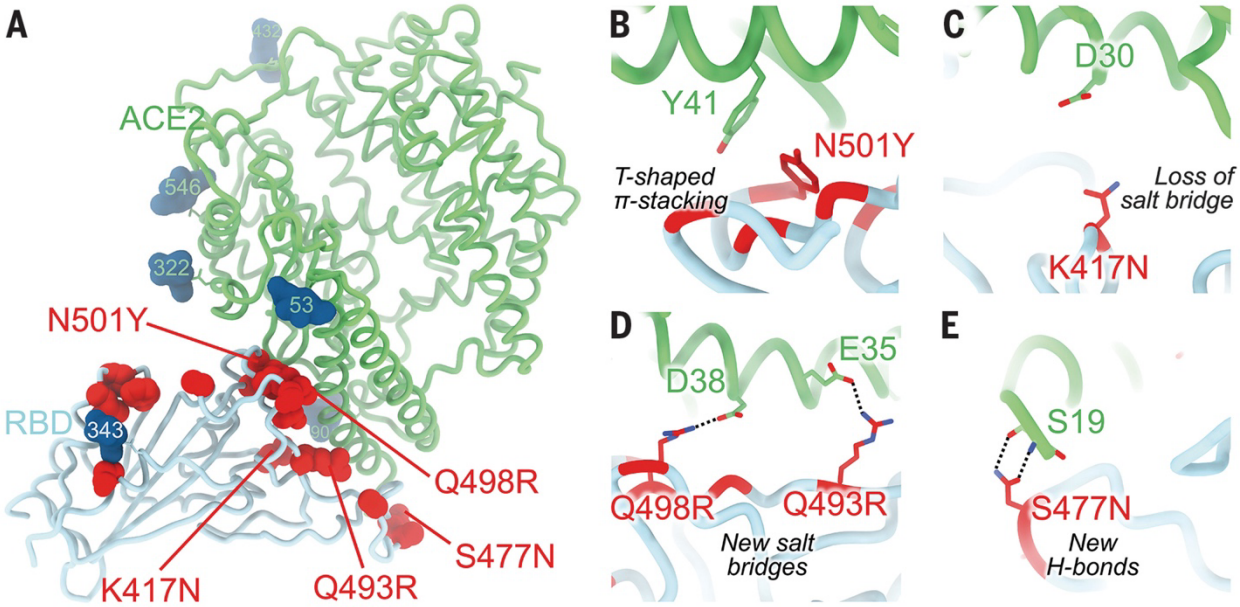
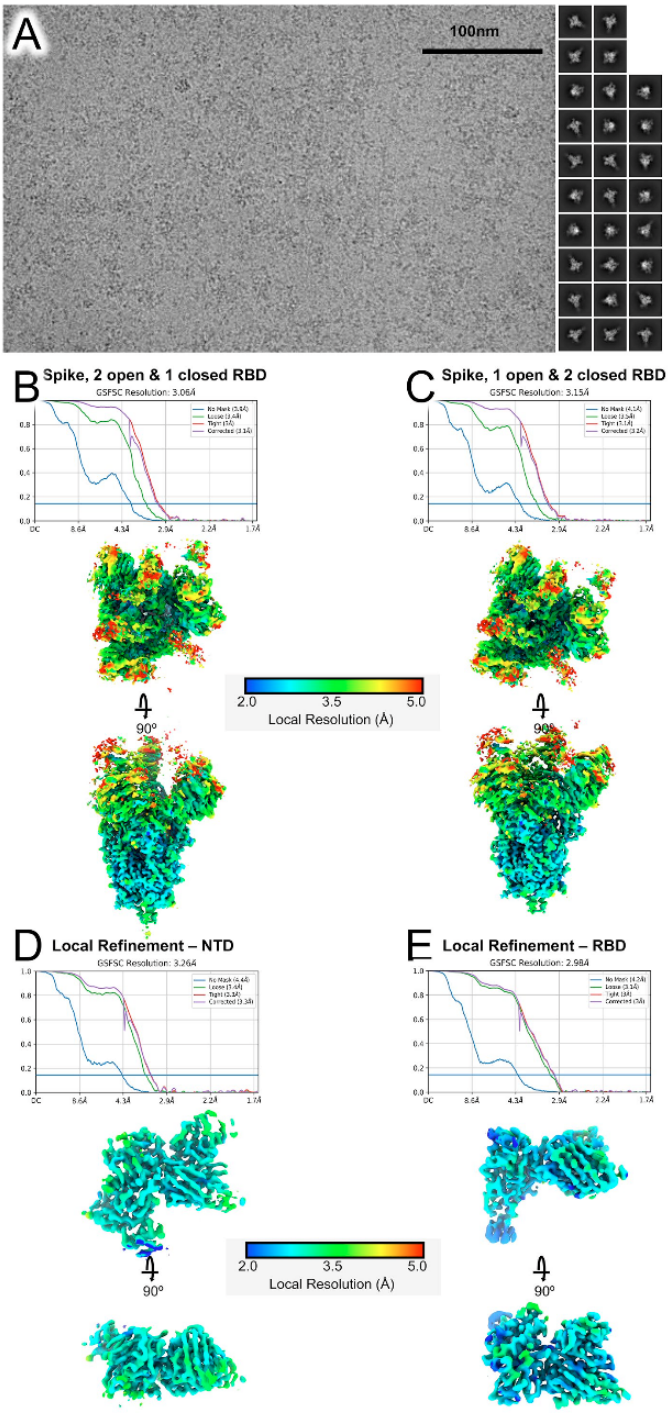


Figure 11.4: Molecular basis of human ACE2 recognition by the SARS-CoV-2 Omicron RBD.

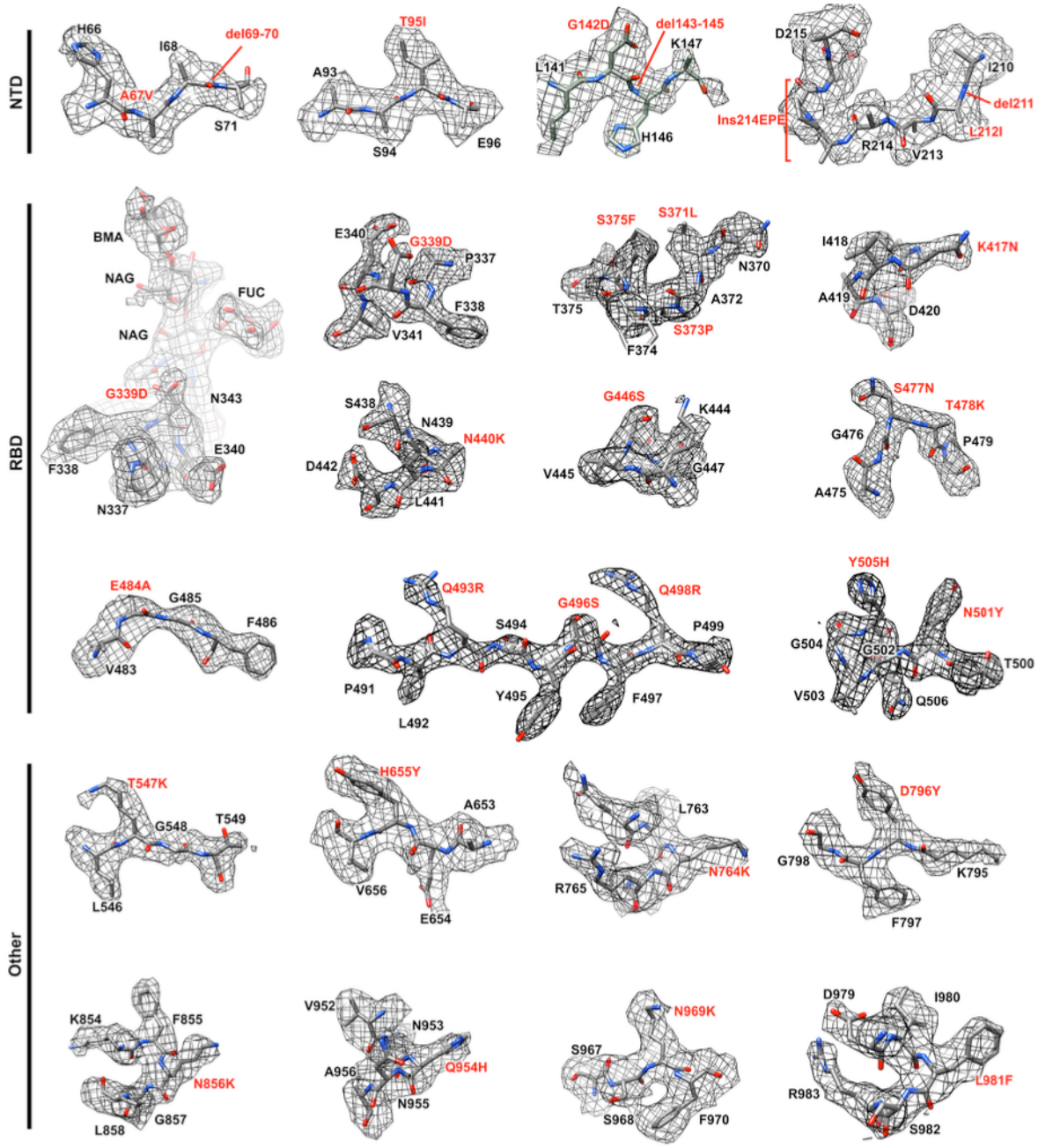
(**A**) Ribbon diagram of the crystal structure of the Omicron RBD in complex with the ACE2 ectodomain. The S309 and S304 Fab fragments are not shown for clarity. (**B** to **E**) Zoomed-in views of the RBD-ACE2 interface. highlighting modulation of interactions due to introduction of the N501Y (B), K417N (C), Q493R/Q498R (D), and S477N (E) residue substitutions.



Supplemental Figure 11.1: CryoEM data processing and validation.

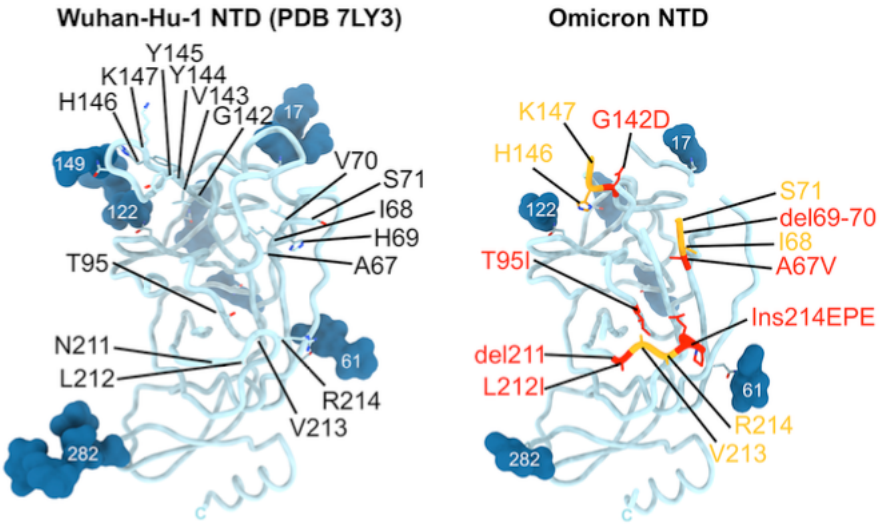
(A) Representative electron micrograph (left, scale bar: 100 nm) and 2D class averages (right) are shown for the indicated particlesembedded in vitreous ice. (**B-F**) Gold-standard Fourier shell correlation curves with the 0.143 cutoff indicated by a horizontal blue line (top) and unsharpened maps colored by local resolutioncalculated

using cryoSPARC (bottom) for whole reconstructions (B, C, and D) and the locally refined reconstructions of NTD- or RBD-bound Fab variable domains (E and F).



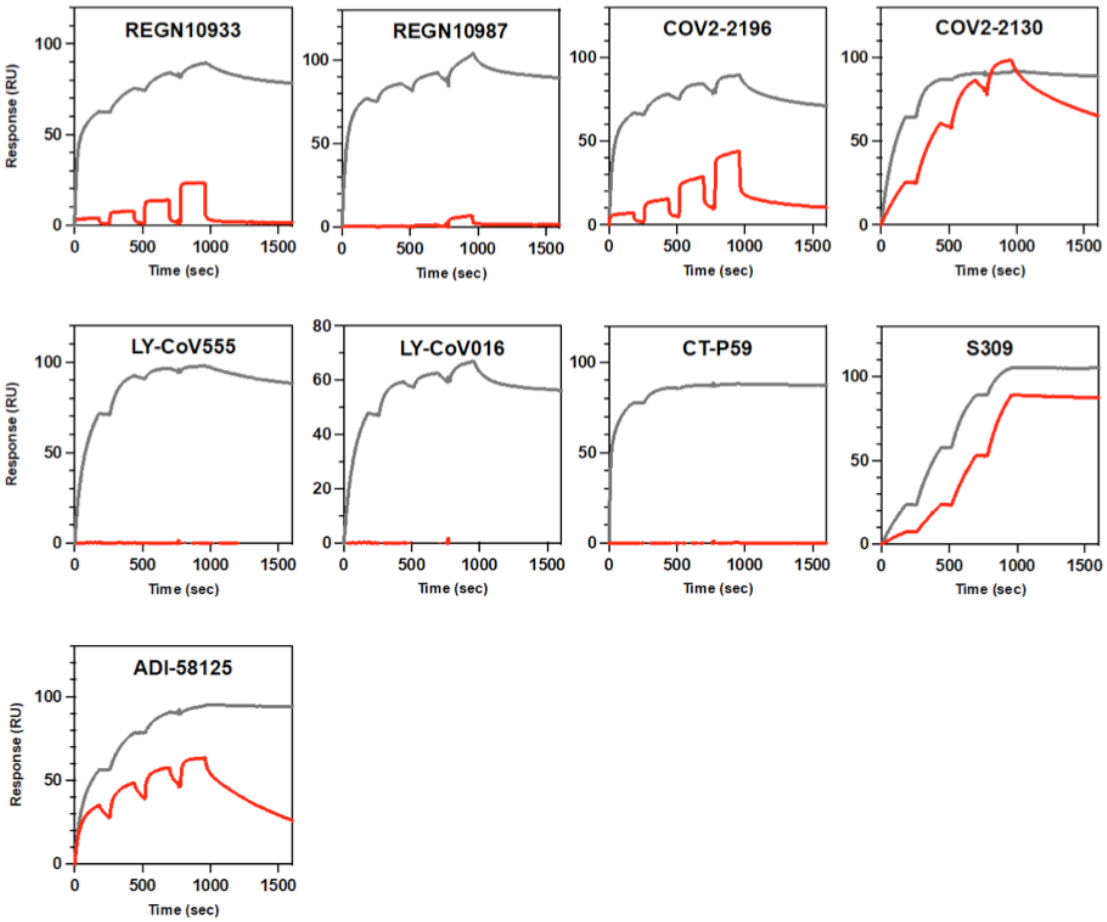
Supplemental Figure 11.2: Density resolving the Omicron S mutations.

The sharpened cryoEM map (NTD and other S regions) or 2Fo-Fc X-ray diffraction electron density (RBD) are rendered as mesh with the corresponding model shown as sticks (carbon: grey, nitrogen: blue, oxygen: red). The N343 glycan density corresponds to the cryoEM map.



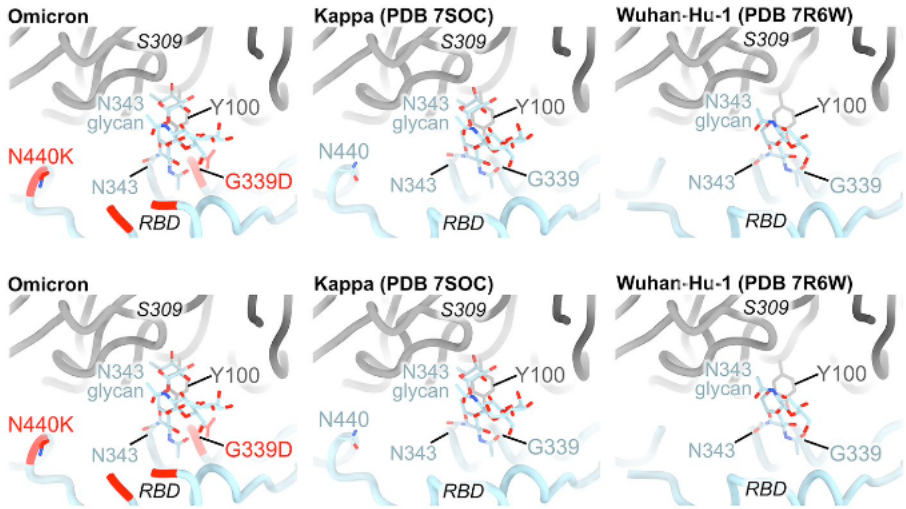
Supplemental Figure 11.3: Comparison of Omicron and Wuhan-Hu-1 NTDs. Omicron.

Mutated residues shown in red as sticks. The glycans are shown as dark blue surfaces.



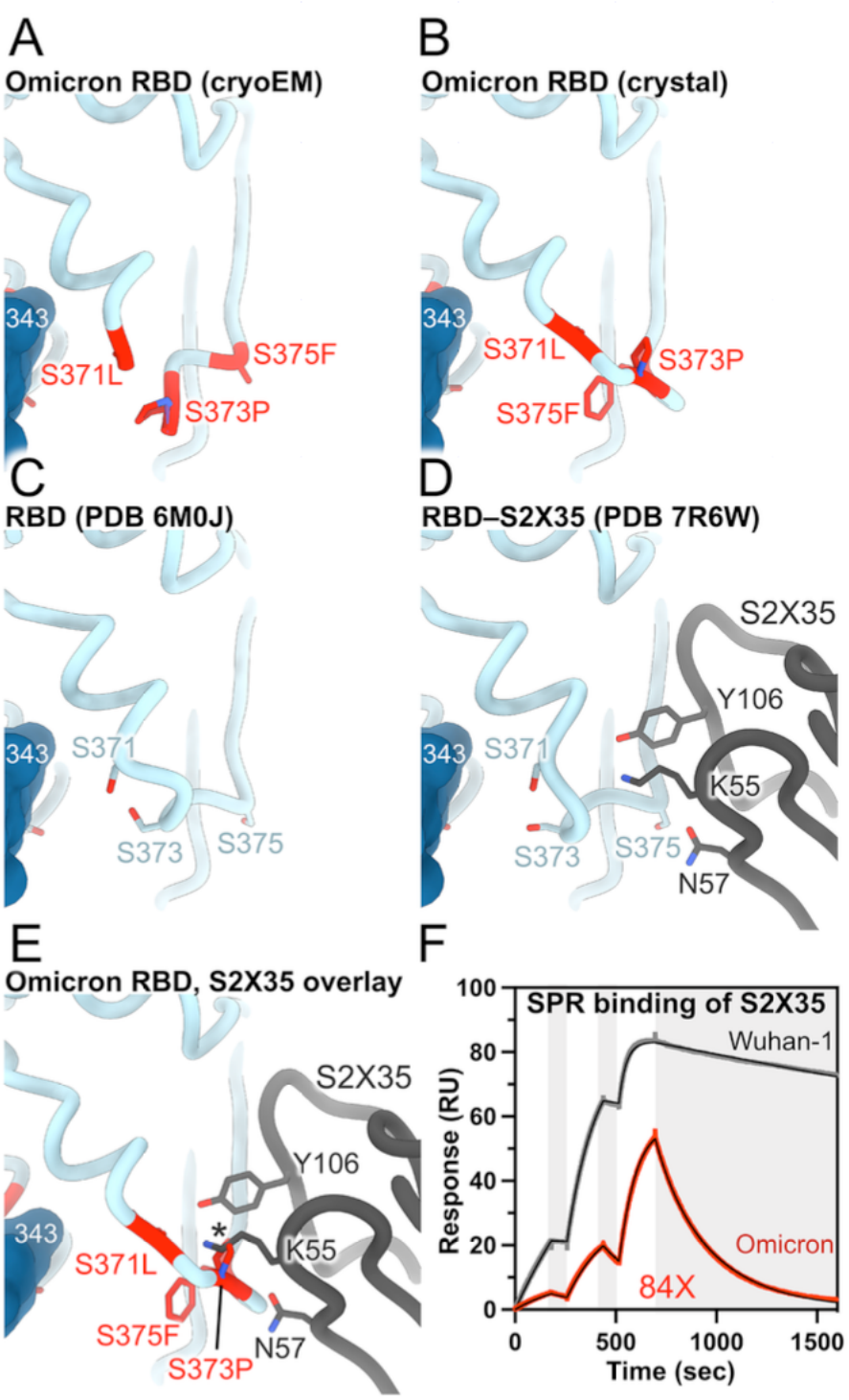
Supplemental Figure 11.4: SARS-CoV-2 Omicron mutations promote escape from clinical mAb binding in the context of the full S ectodomain trimer.

Binding of clinical-stage IgG to surface-immobilized S ectodomain, either Wuhan-Hu-1 (gray line) or Omicron (red line), was evaluated using single-cycle kinetics surface plasmon resonance (SPR) measurements. White and gray stripes are association and dissociation phases, respectively. Data reflect binding avidity due to the bivalent IgG in the mobile phase, and due to the required high density of S attached to the sensor chip matrix.



Supplemental Figure 11.5: Structural basis for S309 binding to the Omicron, Kappa and Wuhan-Hu-1 RBD.

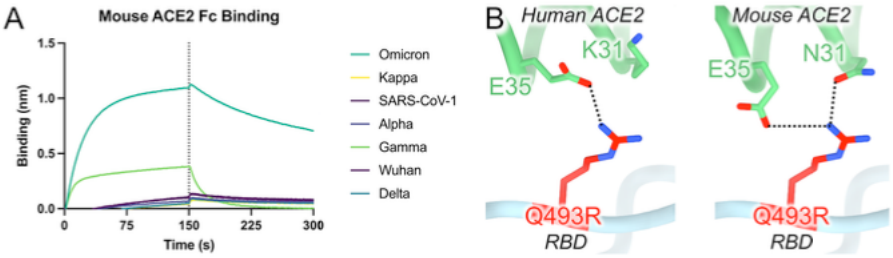
Zoomed-in view of the cryoEM structure of the S309 Fab fragment (black/grey) bound to the Omicron, Kappa (*2*) or Wuhan-Hu-1 (*12*) RBD (blue ribbon) with Omicron mutated residues shown in red as sticks. The N343 glycan is shown as sticks.



Supplemental Figure 11.6: Comparison of the region comprising residues 366 to 375 across structures, and their impact on S2X35 binding.

(A-D) Zoomed-in view of the Omicron RBD from the cryoEM structure (A) and crystal structure (B), as well as the Wuhan-1 RBD (C, PDB 6m0j) (*42*), and Wuhan-1 RBD

bound to S2X35 (D, PDB 7r6w) (*24*). (**E**) Overlay of the RBD from (B) with S2X35 from (D) showing clashing between S373P and K55 (asterisk). (**F**) Binding of the Wuhan-Hu-1 (gray line) or Omicron (red line) RBD to S2X35 mAb was evaluated using surface plasmon resonance (single-cycle kinetics). White and gray stripes are association and dissociation phases, respectively. The black line is a fit to a kinetic model. The decrease in affinity between Wuhan-Hu-1 and Omicron binding is indicated in red.



Supplemental Figure 11.7: The SARS-CoV-2 Omicron RBD recognizes mouse ACE2 efficiently.

A, Biolayer interferometry binding analysis of 1μM dimeric mouse ACE2-Fc to biotinylated SARS-CoV-2 variant RBDs immobilized at the surface of streptavidin (SA) biosensors. The vertical dashed lines indicate the transition between association and dissociation phases. B, In silico model of the interface between the Omicron RBD and mouse ACE2 (right, using PDB 7fdk) based on the crystal structure of the human ACE2-bound Omicron RBD (left).

Table S 11.1 CryoEM data collection, processing and model refinement statistics.

	Omicron S 2 open 1 closed RBDs	Omicron S 2 closed/1 open RBDs		
Data collection and processing				
Magnification (nominal)	130,000	130,000		
Voltage (kV)	300	300		
Electron exposure (e^{-}/Å^2)	63	63		
Defocus range (μm)	0.3-2.0	0.3-2.0		
Pixel size (Å)	0.843	0.843		
Processing Type	Global	Global	Local (RBD)	Local (NTD)
Symmetry imposed	C1	C1	C1	C1
Initial particle images (no.)	482,905	482,905	224,478	224,478
Final particle images (no.)	55,316	74,826	155,712	94,770
Map resolution (Å)	3.2	3.1	3.0	3.3
FSC threshold	0.143	0.143	0.143	0.143
Map sharpening *B* factor (Å^2)	59	63	76	88
Refinement				
Initial model used (PDB code)	-	7SOB	7SOC	7SOD
Model resolution (Å)	-	3.2	3.3	3.4
FSC threshold	-	0.5	0.5	0.5
Model composition				
Nonhydrogen atoms	-	28,354	3,105	3,621
Protein residues	-	4443	408	475
Glycan residues	-	52	4	6
B factors (Å^2)				
Protein	-	46	20	24
Glycans	-	63	23	29
R.m.s. deviations				
Bond lengths (Å)	-	0.01	0.01	0.01
Bond angles (°)	-	1.4	1.1	1.1
Validation				
MolProbity score	-	0.98	1.0	0.9
Clashscore	-	1.49	1.7	0.7
Rotamer outliers (%)	-	0.17	0.0	0.3
Ramachandran plot				
Favored (%)	-	97.55	97	97
Allowed (%)	-	2.45	3	3
Outliers (%)	-	0	0	0
EMRinger score	-	3.91	5.2	5.5
Data Availability				
EMDB	EMD-25993	EMD-25992	EMD-25990	EMD-25991
PDB	-	7TM0	7TLY	7TLZ

Table S 11.2: Crystallographic data collection and refinement statistics.

	Omicron RBD/ hACE2/S309/S304 PDB 7TN0
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	78.28, 183.65, 194.55
α, β, γ (°)	90.00, 95.98, 90.00
Resolution (Å)	48.91-2.85 (2.90-2.85)
R_{merge}	0.343 (2.676)
R_{pim}	0.153 (1.184)
$I / \sigma I$	5.6 (0.7)
CC(1/2)	0.975 (0.209)
Completeness (%)	96.7 (98.4)
Redundancy	5.8 (5.8)
Refinement	
Resolution (Å)	48.91-2.85
No. reflections	121127
R_{work} / R_{free}	23.03/26.79
No. atoms	
Protein	24486
Ligand/ion	221
Water	187
B-factors	
Protein	63.5
Ligand/ion	86.65
Water	42.64
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.658
Ramachandran	
Favored (%)	97.6
Outliers (&)	0

*Values in parentheses are for highest-resolution shell.

Table S 11.3: Kinetic parameters and fitting details of the SPR RBD binding measurements.

1:1 Binding Model							
IgG	**RBD**	**ka (1/Ms)**	**kd (1/s)**	**KD (M)**	**Rmax (RU)**	**Capture level (RU)**	
S309	Wu-Hu-1	1.49E+05	1.32E-04	8.87E-10	73.77	210.5	
S309	Omicron	7.15E+04	1.90E-04	2.67E-09	82.01	211.8	
ADI-58125	Wu-Hu-1	1.15E+06	7.17E-08[a]	< 1E-11[a]	113.6	231	
ADI-58125	Omicron	5.42E+06	0.2497	4.61E-08	83.81	230.3	
COV2-2130	Wu-Hu-1	4.26E+05	8.38E-04	1.97E-09	109.5	251.7	
COV2-2130	Omicron	9.30E+05	0.05897	6.34E-08	79.31	246.7	
COV2-2196	Wu-Hu-1	1.14E+07	0.01244	1.09E-09	100.7	240.5	
COV2-2196	Omicron	9.04E+05	0.441	4.88E-07	100[b]	239.7	
CT-P59	Wu-Hu-1	5.54E+06	3.65E-04	6.60E-11	97.99	236.8	
CT-P59	Omicron	No binding				236.9	
LY-CoV016	Wu-Hu-1	See Biphasic fit results, below				196.1	
LY-CoV016	Omicron	No binding				195	
LY-CoV555	Wu-Hu-1	7.92E+05	0.001174	1.48E-09	100.7	242.1	
LY-CoV555	Omicron	No binding				243.6	
REGN10933	Wu-Hu-1	5.29E+06	0.003645	6.89E-10	93.53	260	
REGN10933	Omicron	6.16E+05	0.5073	8.24E-07	94[b]	259.6	
REGN10987	Wu-Hu-1	See Biphasic fit results, below				231.6	
REGN10987	Omicron	No binding				231.5	
S2X35	Wu-Hu-1	5.70E+05	1.53E-04	2.68E-10	83.6	230.1	
S2X35	Omicron	2.43E+05	5.43E-03	2.24E-08	82.5	229.4	
Biphasic fit (Heterogeneous Ligand Model), binding to Wuhan-Hu-1 RBD[c]							
ka1 (1/Ms)	**kd1 (1/s)**	**KD1 (M)**	**ka2 (1/Ms)**	**kd2 (1/s)**	**KD2 (M)**	**Rmax1 (RU)**	**Rmax2 (RU)**
LY-CoV016							
1.24E+06	**0.0172**	**1.38E-08**	1.79E+05	5.23E-05	2.91E-10	**59.66**	20.68
REGN10987							
1.28E+06	**0.0159**	**1.24E-08**	2.74E+05	1.02E-08	3.73E-14	**70.48**	28.6

a Measured dissociation is beyond the limit of detection. Reporting KD as an upper bound

b Rmax fixed to Rmax from Wuhan-Hu-1 fit (without constraint, Rmax from Omicron fit was
implausibly much lower than Rmax for corresponding Wuhan-Hu-1 binding)

c Values for the kinetic phase attributed to real binding (vs an artifact) are highlighted in bold

CHAPTER 12. ALTERED TMPRSS2 USAGE BY SARS-COV-2 OMICRON IMPACTS INFECTIVITY AND FUSOGENICITY.

In this study, I continue to describe the SARS-CoV-2 Omicron BA.1 variant which emerged in 2021. I describe the higher affinity for ACE2 demonstrated by Omicron compared with Delta. In the subsequent sections, I detail the marked changes in of Omicron compared to Delta in antigenicity, tissue specific replication rates, and cleavage. I discuss how alterations in TMPRSS2 by omicron explain the differences in replication show the Omicron spike inefficiently uses the cellular protease TMPRSS2.

Adapted from: Meng, B., Abdullahi, A., Ferreira, I.A.T.M., Goonawardane, N., Saito, A., Kimura, I., Yamasoba, D., Gerber, P.P., Fatihi, S., Rathore, S., Zepeda, S.K., Papa, G., Kemp, S.A., Ikeda, T., Toyoda, M., Tan, T.S., Kuramochi, J., Mitsunaga, S., Ueno, T., Shirakawa, K., Takaori-Kondo, A., Brevini, T., Mallery, D.L., Charles, O.J. (2022). "Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity." Nature 603(7902): 706-714.

12.1 Chapter Introduction

The Omicron variant of SARS-CoV-2, which was first detected in South Africa(Cele, 2021), carries over 30 mutations in its spike protein and has now spread internationally at a fast pace. More than 20 substitutions exist in the N-terminal domain and receptor-binding domain (RBD). The Omicron variant also contains six unique mutations in S2 (N764K, D796Y, N856K, Q954H, N969K and L981F) that were not previously detected in other variants of concern (VOCs). The spike proteins of the Alpha

and Delta variants were previously shown to confer more efficient cell–cell fusion kinetics compared with Wuhan-Hu-1 (refs. Mlkochova, 2021; Meng, 2021) due to mutations in the furin-cleavage site region that increase S1/S2 cleavage. Furthermore, Alpha and Delta showed more efficient syncytia formation, which was previously found to be associated with pathogenesis(Saito, 2021). Omicron has three mutations in the region of the furin-cleavage site (P681H, H655Y and N679K) and was initially predicted to be highly infectious(Peacock, 2021) and pathogenic. Indeed, Omicron has been associated with very rapid increases in case numbers and recent data demonstrate considerable reinfection and vaccine 'breakthrough', probably due to evasion of neutralizing antibody responses(Cele, 2021; Dejnirattisai, 2021). However, paradoxically, recent findings suggest reduced severity in Omicron infections compared to Delta(Wolter, 2022).

Here we examine the biological properties of Omicron, and focus on spike-mediated evasion of neutralizing antibodies, increased ACE2 binding affinity, a shift in tropism away from TMPRSS2-expressing cells and compromised ability to generate syncytia.

12.2 Omicron spike-binding affinity to ACE2

The Omicron variant has 15 amino acid mutations in the RBD (**Supplementary Figs. 12.1 and 12.2**). To understand the impact of these substitutions on receptor engagement, we determined the kinetics and affinity of monomeric human ACE2 binding to immobilized Omicron, Wuhan-Hu-1 and Delta RBDs using biolayer interferometry. We observed that the Omicron RBD has approximately threefold enhanced binding affinity for ACE2 relative to the Wuhan-Hu-1 and Delta RBDs

(**Extended Data Table 12.1 and Supplementary Fig. 12.2**), consistent with previous findings(Walls, 2020; Cameroni, 2021; Starr, 2020; McCallum, 2021). As previously observed for the Alpha RBD, which contains only the N501Y mutation(Collier, 2021), the modulation of binding is mediated by changes in ACE2 dissociation rates. Although the K417N mutation is known to decrease ACE2 engagement(Collier, 2021; McCallum) , a recently determined crystal structure of ACE2-bound Omicron RBD revealed that Q493R and Q498R introduce additional electrostatic interactions with ACE2 residues Glu35 and Asp38, respectively (McCallum, 2021), whereas S477N enables hydrogen-bonding with ACE2 Ser19 (**Supplementary Fig. 12.2**). Collectively, these mutations strengthen ACE2 binding, relative to the ancestral isolate. We extended this binding analysis to a cell-based model using cells that were transfected with full-length spike, followed by ACE2 antibody titration; we observed significantly higher ACE2 binding for Omicron spike compared with the spike proteins of the Wuhan-Hu-1 and Delta variants (**Supplementary Fig. 12.2**). This enhanced binding could be a factor in the enhanced transmissibility of Omicron relative to previous variants.

12.3 The sensitivity of Omicron to therapeutics

Omicron was predicted to have broad resistance to neutralizing antibodies based on mutations in the class I–IV antigenic regions in its RBD(Barnes, 2020). Current standard of care antiviral treatment for moderate to severe COVID-19 includes the use of the monoclonal antibody combination REGN10933 (casivirimab) and REGN10897 (imdevimab). The K417N, E484A, S477N and Q493R mutations preclude REGN10933 binding, whereas G446S clashes with REGN10987, consistent with the decreased binding to the Omicron RBD and S trimer and the loss of neutralization

previously described(Cameroni, 2021; McCallum, 2021). We next tested serial dilutions of component monoclonal antibodies, both individually and in combination, against Delta and Omicron live viruses in tissue culture (**Fig. 12.1a**). Although the Delta variant was effectively neutralized by casirivimab, imdevimab was only partially effective, consistent with previous data(Mlcochova, 2021). In combination, these monoclonal antibodies were highly potent against Delta. However, there was a complete loss of neutralizing activity against Omicron by these monoclonal antibodies alone or in combination (**Fig. 12.1a**). Given these results, we next tested the direct-acting antivirals remdesivir and the active metabolite of molnupiravir against live viruses. We observed similar antiviral activity against Delta and Omicron using serial titrations of both compounds (Supplementary **Fig. 12.3**).

12.4 Omicron and polyclonal antibodies

An important question is whether vaccine-elicited antibodies are able to neutralize Omicron. We synthesized codon-optimized spike expression plasmids for Omicron and Delta spike proteins and generated pseudotyped virus (PV) particles by co-transfecting the spike expression plasmids with a lentiviral gag-pol-expressing plasmid and a lentiviral transduction plasmid encoding the luciferase gene(Naldini, 1996; Mlcochova, 2020). We obtained longitudinal serum samples from 40 individuals who were vaccinated with either the BNT162b2 or ChAdOx-1 vaccine and performed serum titrations before mixing sera with our reporter PV particles. The participants had a median age of around 70 years and prospective serum samples were taken as follows: one month after dose two, six months after dose two and one month after dose three (**Extended Data Table 12.2**). We observed at least tenfold loss of neutralization against

Omicron after the second dose compared with Delta (**Fig. 12.1b, c and Extended Data Fig. 12.4**). Indeed, neutralization of Omicron was not detectable for the majority of individuals who had received two doses of ChAdOx-1. We also observed waning over time after the second dose for both vaccines (**Fig. 12.1b, c**). Both groups received a booster dose of BNT162b2 as a third dose, enabling us to compare the response to this booster dose. Considerable increases (greater than tenfold) in neutralization against both Omicron and Delta were observed for all of the variants after the third dose of vaccination, suggesting an increase in the breadth of responses as well as titre.

To confirm the loss of neutralizing activity against Omicron after the second dose, we next used a live virus experimental system to compare Delta and Omicron variants against sera taken four weeks after the second dose of BNT162b2, and obtained similar results to those obtained in the PV assay (**Extended Data Fig. 12.5a**). These live viruses were also used to assess the neutralization of the Omicron variant by sera derived from unvaccinated individuals who were previously infected with the early Wuhan-Hu-1 virus or the Delta variant. As expected, vaccine sera had significantly impaired activity against Omicron compared with Delta (**Extended Data Fig. 12.5a**). We also tested mRNA-1273 vaccine-elicited sera, which showed a similar reduction in neutralization to BNT162b2. However, coronavac sera showed little neutralization against Delta and 0 out of 9 participants had detectable neutralization against Omicron. Sera from Delta infections appeared to have lower cross-neutralization compared with those from the early pandemic period when Wuhan-Hu-1 D614G was dominant (**Extended Data Fig. 12.5b**).

12.5 Omicron replication and spike cleavage

We infected primary human nasal epithelial 3D cultures (hNECs) with an air–liquid interface (**Fig. 12.2a, b**). Infection with live SARS-CoV-2 Omicron or Delta was conducted at the apical surface with equal amounts of input virus and virus collected from the apical surface at 24 and 48 h, and virus was quantified using quantitative PCR (qPCR) analysis of the *E* gene and 50% tissue culture infectious dose ($TCID_{50}$). We observed similar replication kinetics for Omicron and Delta in the hNECs (**Fig. 12.2b**). We next infected Calu-3 lung cells (which are known to express endogenous *TMPRSS2*) with live isolates, and observed significantly greater viral replication for Delta than Omicron (**Fig. 12.2c**), manifesting as early as 24 h after infection. We also infected Caco-2 cells, a colon cancer cell line (known to express high levels of endogenous *TMPRSS2*)(Pommerenke, 2021) and, as for Calu-3, we found 1 log greater Delta RNA and higher infectious virus ($TCID_{50}$) in the supernatants compared with those observed for Omicron. Finally, we tested a cell line—HeLa, overexpressing *ACE2* and *TMPRSS2*—and obtained similar results to those of the Calu-3 and Caco-2 cells.

We next sought to examine the possibility that the impaired replication in some cells may be related to differential spike incorporation and/or cleavage. For example, Delta, which is known to have higher replication, is associated with a highly cleaved spike protein (Mlcochova, 2021). We tested purified virions from two independent Omicron isolate infections. We observed broadly similar amounts of Omicron and Delta spike incorporation into virions generated in VeroE6 ACE2/TMPRSS2 cells (**Fig. 12.2f**). However, there was reduced cleavage in whole Omicron virions compared with Delta as evidenced by the ratio of full-length spike to S1 and S2; the Omicron spike was

inefficiently cleaved in virions compared with Delta (**Fig. 12.2f**). However, more substantially reduced Omicron spike cleavage compared with Delta was observed in cell lysates (**Fig. 12.2g**).

Given the lower intracellular S1/S2 cleavage for Omicron versus Delta spike proteins, confocal microscopy was used to examine whether cleavage efficiency is related to differences in subcellular localization. We analysed the distribution of Spike in cells infected with Omicron and Delta SARS-CoV-2 live virus isolates (**Fig. 12.2h–j**). No clear differences were observed between the variants in distance from the nucleus or co-localization with a Golgi marker, indicating that the subcellular localization between Delta and Omicron is similar.

12.6 Omicron entry is impaired in TMPRSS2^{+} cells

SARS-CoV-2 spike is a major determinant of viral infectivity and mediates cell entry through interaction with ACE2 (refs. Hoffmann, 2020; Walls, 2020) and either TMPRSS2 (ref. Hoffmann, 2020) at the plasma membrane or cathepsins in the endosomes. The plasma membrane route of entry, and indeed transmissibility and disease severity in animal models, is critically dependent on the polybasic cleavage site (PBCS) between S1 and S2 (refs. Saito, 2021; Peacock, 2021) and cleavage of spike before virion release from producer cells; this contrasts with the endosomal entry route, which does not require spike cleavage in producer cells(Hoffmann, 2020). Plasma membrane fusion also enables the virus to avoid restriction factors in the endosomes(Peacock, 2021).

We tested spike-mediated viral entry of the wild-type (WT) Wuhan-Hu-1 D614G, Delta and Omicron spike proteins (**Fig. 12.3a**) using the PV system. We first probed PV

virions for spike protein and, after western blot analysis, noted a reduced incorporation of Omicron spike into virions compared with the incorporation of Delta spike (**Extended Data Fig. 12.6**). We also noted that the Omicron spike was less cleaved than the Delta spike, as observed with the live virus (**Extended Data Fig. 12.6**). Notably, cleavage of Omicron spike in cells was also lower compared with Delta and WT spike, again similar to that in the live virus.

Using normalized amounts of Delta and Omicron spike PV, we infected primary 3D lower-airway organoids and gallbladder organoids(Youk, 202; Sachs, 2019) (**Fig. 12.3a and Extended Data Fig. 12.6**), in addition to Calu-3 lung cells. Mirroring the lower replication of Omicron relative to Delta in live virus assays, we observed impaired entry efficiency for Omicron spike in both organoid systems and the Calu-3 cells in comparison to Delta and WT Wuhan-Hu-1 D614G. By contrast, in H1299 lung cancer epithelial cells, HeLa-ACE2 cells (overexpressing *ACE2*), and HEK293T cells overexpressing *ACE2*and deleted for *TMPRSS2* (293T-A2ΔT2), we did not observe large differences in the entry efficiency of Delta and Omicron (**Fig. 12.3a**).

To further examine our PV entry and infection findings, we studied the expression of *ACE2* and *TMPRSS2* across our target cells using qPCR analysis of RNA extracts from cell lysates (**Fig. 12.3b**). We found greater levels of *TMPRSS2* mRNA in cells in which Omicron PV entry was impaired relative to Delta—for example, Calu-3 cells and organoids had higher levels of *TMPRSS2* mRNA compared with H1299, HeLa-ACE2 and 293T-A2ΔT2 cells. *ACE2* levels were variable as expected, and did not appear to be correlated with differences in infection between Omicron and Delta PV. We therefore

hypothesized that the Omicron PV was disadvantaged relative to Delta in *TMPRSS2*-expressing cells.

To experimentally demonstrate differential usage of TMPRSS2 as a cofactor for virus entry by Omicron, we directly compared the efficiency of viral entry between 293T-A2ΔT2 and 293T-A2T2 cells. Enhanced infectivity was observed for both the WT and Delta when *TMPRSS2* was overexpressed, suggesting that TMPRSS2 is also required for their optimal virus entry. By contrast, no difference was observed for Omicron PV, indicating that Omicron is inefficient in using TMPRSS2 for its entry (**Fig. 12.3c**). We hypothesized that the modest degree of TMPRSS2-dependent enhancement (3.4×) for Delta and a complete lack of enhancement for Omicron over a single infection round might relate to the overexpression of *ACE2* in this system. To examine this possibility, we tested standard HEK293T cells expressing low endogenous levels of *ACE2* and those transduced to overexpress *TMPRSS2* (293T-T2). Indeed, we observed a much greater increase in infection for Delta (34×) under TMPRSS2 overexpression conditions and a smaller increase (6×) for Omicron in the presence of overexpressed *TMPRSS2* (**Fig. 12.3d**). Together, these data support a reduced contribution of TMPRSS2 cofactor usage by Omicron spike, with modulation of the effect size by *ACE2* expression levels. Given previous data on the relationship between S1/S2 cleavage in producer cells and TMPRSS2 dependence(Papa, 2021; Ou, 2021), the most likely explanation for altered TMPRSS2 usage is the suboptimal S1/S2 cleavage in Omicron leading to inefficient S2′ processing and fusion peptide exposure.

A change in the use of TMPRSS2 for entry would be predicted to alter the entry pathway of the virus (**Fig. 12.3e**). To test the hypothesis that the entry route preference

by Omicron spike is changed, we used inhibitors of proteases specific to either the endocytic pathway (E64d blockade of cathepsins) or the plasma membrane pathway (camostat blockade of TMPRSS2). We reasoned that, if Omicron has become more dependent on the endocytic cell entry route, it should be more sensitive to cathepsin inhibition. We infected A549 cells overexpressing *ACE2* and *TMPRSS2* (A549-A2T2) cells in the presence of serial dilutions of E64d or camostat. Indeed, E64d had a greater effect on cell entry by Omicron PVs versus Delta PVs and, conversely, camostat had a greater effect on the entry of Delta PVs (**Fig. 12.3f**). We next performed this experiment with live virus isolates, using an indicator lung cell line (A549-A2T2) with a similar, albeit less marked, result as shown for PV (**Fig. 12.3g**). These drug data in both live virus and PV systems further support the shift in tropism away from *TMPRSS2*-expressing cells.

Having established that TMPRSS2 modulates entry mediated by Delta to a greater extent than for Omicron BA.1 spike, we sought to understand the distribution of *TMPRSS2* and *ACE2* expression in human respiratory cells. We used single-nucleus RNA-sequencing (RNA-seq) data from five locations in the human lungs(Madissoon, 2021). The comparison revealed higher expression of *TMPRSS2* in alveolar AT1 and AT2 pneumocytes and, in general, lower expression in the trachea (upper airway; **Extended Data Fig. 12.7**). *ACE2* expression was generally lower overall, but appeared to be higher in AT1 and AT2 cells compared with in other cell types, and also in club cells(Salahudeen, 2020). We experimentally extended these findings by performing qPCR analysis of human lung tissue samples(Salahudeen, 2020), showing higher mRNA expression of *TMPRSS2* in the lung parenchyma (alveolar) tissue compared with

in the upper-airway bronchial tissues and a trend towards *ACE2* being higher in the upper-airway tissue (**Fig. 12.3h**).

12.7 Impaired cell–cell fusion by Omicron S

The ability of viral membrane glycoproteins to induce cell–cell fusion and syncytium formation is well established(Buchrieser, 2021; Cattin, Ortola, 2021; Li, 2003), providing an additional route for SARS-CoV-2 dissemination that may facilitate evasion of neutralizing antibodies. The role of syncytium formation in viral replication and the pathogenesis of severe COVID-19 has been reported and may be a druggable process to treat COVID-19 pathology(Braga, 2021; Bussani, 2020). SARS-CoV-2-mediated cell–cell fusion requires the PBCS and spike cleavage at S1/S2, and the process is known to be accelerated by the presence of TMPRSS2 (ref. Buchrieser, 2021).

Mutations at P681 in the PBCS have been observed in multiple SARS-CoV-2 lineages, most notably in the B.1.1.7/Alpha variant(Meng, 2021) and the Delta/Kappa variants from the B.1.617 lineage(Mlcochova; 2021; Ferreira, 2021). We previously showed that spike proteins of these variants, all bearing Pro681 mutations, had significantly higher fusogenic potential compared with the D614G Wuhan-Hu-1 spike (Mlcochova, 2021; Meng, 2021). Omicron bears P681H, in addition to the N679K and H655Y mutations (**Supplementary Fig. 12.6a**).

Given the requirement of TMPRSS2 for optimal cell–cell fusion, we hypothesized that Omicron may be impaired in mediating this process. We used a split GFP system(Cabantous, 2005) to monitor cell–cell fusion in real time (**Fig. 12.4a**). As a control to show that spike cleavage is needed for fusion in our assay system, we titrated

CMK (a furin inhibitor) into donor cell medium before transfection with spike-expressing plasmids, observing dose-dependent inhibition of fusion (**Extended Data Fig. 12.8**). As a further control to demonstrate the need for ACE2 engagement by expressed spike, we tested the ability of convalescent serum containing neutralizing antibodies to block cell–cell fusion. Indeed, serum blocked syncytia formation in a dose-dependent manner (**Extended Data Fig. 12.8**).

We performed flow cytometry to verify that spike was expressed at the cell surface (**Fig. 12.4b**). We proceeded to transfect spike-bearing plasmids into HEK293T cells expressing GFP1-10 and Vero E6 cells stably expressing GFP11, such that the GFP signal could be detected after cell–cell fusion and measured over time. (**Fig. 12.4c, d**). We observed increased fusion for Delta spike compared with D614G Wuhan-Hu-1 spike (**Fig. 12.4d**). However, the Omicron spike resulted in very poor fusion (**Fig. 12.4d**).

We predicted that poor fusion would impair cell–cell spread of Omicron, and analysed infection focus size using live virus infection of H1299 cells, in which we have shown similar entry efficiency for Omicron and Delta spike PV (**Fig. 12.3**). Infection foci in spreading virus infections occur because of localized (cell-to-cell) infection, probably facilitated by syncytium formation. The infection was followed by overlay of semi-solid medium (carboxymethylcellulose) to inhibit cell-free infection, thereby favouring cell–cell infection. Infected cells were stained with an anti-S antibodies, and local spread was characterized by the staining of foci visible to the naked eye (**Fig. 12.4e**). Omicron infection resulted in slightly higher numbers of foci (**Fig. 12.4e–g**), but each infection focus was substantially smaller in size relative to those formed by Delta infection

(**Fig. 12.4e**). For a more detailed view of the foci, we imaged at a higher magnification using a microscope objective, and used cells that were imaged 2 h after infection with ancestral virus—when relatively little viral transmission is expected to be completed—as an approximate measure of the size of a single cell (**Extended Data Fig. 12.8**). We compared this to cells that were infected for 18 h by Omicron or Delta isolates. The size of foci at 2 h after infection with ancestral virus was 1.2×10^3 μm^2 (95% confidence interval (CI) = 0.9–1.5×10^3). The size of Omicron foci was 1.5×10^3 μm^2 (95% CI = 1.4–1.5×10^3) and, by contrast, the size of Delta foci was 3.4×10^3 μm^2 (95% CI = 3.1–3.7×10^3), 2.3-fold higher than Omicron (**Extended Data Fig. 12.8**). Together, these results show that localized cell–cell infection of Omicron is reduced relative to Delta, consistent with attenuated cell–cell fusion.

12.8 Discussion

Here we examined the biological properties of Omicron from the perspectives of spike-mediated immune evasion, ACE2-binding interactions and cellular entry pathways that likely dictate tissue tropism. First, we have shown that the Omicron spike confers substantial evasion of vaccine-elicited neutralizing antibodies that is greater for ChAdOx-1 versus BNT162b2 vaccine sera. In longitudinally sampled participants, second-dose waning was mitigated by a third dose of an mRNA vaccination that also increased and broadened neutralization of Omicron in the short term. This observation is critical for informing third-dose vaccination efforts worldwide. Notably, ChAdOx-1 is widely used in low-income settings in which third doses of mRNA vaccines are not widely available, and Omicron may contribute to higher infection rates in such settings unless third doses can be implemented. Of further concern for low-income countries

was the absence of any neutralizing activity to Omicron in sera after two doses of coronavac in all nine participants, in addition to poor Delta neutralization.

In terms of the implications for treatment options for moderate to severe clinical disease, we show high-level loss of in vitro activity for the widely used monoclonal antibody combination therapy REGN2 against Omicron, but no significant loss in the activity of the polymerase inhibitors remdesivir and molnupiravir against the live virus. The monoclonal antibody sotrovimab has also been reported to retain significant in vitro activity against Omicron(Cameroni, 2021; Cathcart,2021).

Despite the presence of three mutations that are individually predicted to favour spike S1/S2 cleavage, the observed cleavage of spike in cells infected with live virus was lower compared with Delta and WT Wuhan-Hu-1 D614G. Similarly, live Omicron virions had a lower proportion of cleaved spike compared with Delta. Cleavage deficiency was also evident in PV particles. We previously noted that the proportion of cleaved spike was lower within cells as compared to virions(Mlcochova, 2021), suggesting that the incorporation of spike into virions produced in the endoplasmic reticulum–Golgi intermediate compartment may favour cleaved spike over full-length spike, or be cleaved as virions are secreted through exocytic pathways(V'Kovski, 2021).

Live Omicron virus showed a significant replication defect compared with Delta in cell culture systems in which TMPRSS2 was present. Omicron spike was also associated with decreased entry relative to both WT and Delta spike proteins in target lower-airway organoids, gallbladder organoids or Calu-3 lung cells expressing TMPRSS2, but displayed no difference in cells with low *TMPRSS2* expression, for example H1299, HeLa-ACE2 and 293T-A2 cells. We proceeded to directly compare

entry in *ACE2*-overexpressing cells in which *TMPRSS2* was either knocked down or overexpressed, and showed that Omicron entry was not affected by variations in *TMPRSS2* expression, in contrast to WT and Delta spike PV entry. Importantly, when *ACE2* was not overexpressed, we observed a large increase in Delta PV entry with TMPRSS2 availability, and a much smaller increase in entry for Omicron. These data suggest that TMPRSS2 cofactor usage may be affected by ACE2 levels. The hypothesis that Omicron spike has altered its preference to the cathepsin-dependent endosomal route of entry (TMPRSS2 independent) was borne out by inhibitor experiments against cathepsin and TMPRSS2 in both PV and live virus isolates.

TMPRSS2 is a member of the type II transmembrane serine protease family. This family comprises 17 members with diverse physiological functions. These proteins require activation and remain associated with membranes, and TMPRSS2 in particular can undergo autocatalytic activation(Bugge, 2009), although its role in human physiology is unclear. TMPRSS2 has been shown to be a cofactor for SARS-CoV-1 and MERS as well as SARS-CoV-2 (refs. Hoffmann, 2020; Park, 2016). Importantly, TMPRSS2 dependence is correlated with efficient cleavage at the PBCS between S1/S2 (refs. Papa, 2021; Ou, 2021)), with the removal of the furin-cleavage site in SARS-CoV-2 reducing TMPRSS2 dependence. By contrast, in experiments in which a furin-cleavage site was introduced into SARS-CoV-1, the virus demonstrated greater TMPRSS2 dependence.

Omicron presents an interesting parallel to SARS-CoV-1, which is not cleaved in the producer cell due to lack of the PBCS but, rather, on the target cell surface by TMPRSS2 or by cathepsins in endosomes. SARS-CoV-1 is able to infect cells though

TMPRSS2 (and therefore syncytia can be induced)(Li, 2003) or by endocytosis(Hoffmann, 2020,Oum 2021). The SARS-CoV-2 Omicron variant appears to also favour endocytosis through reduced cleavage efficiency of spike impairing TMPRSS2-mediated entry. Indeed, the accumulated mutations in omicron include multiple additional basic amino acids leading to an increase in the overall charge of the S protein. This change may make S more sensitive to low-pH-induced conformational changes, and be an adaptation that facilitates the use of the low-pH endosomal entry route, and/or be an adaptation to entry in the lower pH environment in the upper airway. Furthermore, our analysis of single-cell RNA-seq datasets and direct qPCR measurements on human respiratory tissue samples suggest that lung cells have higher *TMPRSS2* mRNA compared with cells found in the upper airway. Indeed recent data indicate lower virus burden in deep lung tissue and reduced pathogenicity in Omicron versus Delta infections using Syrian hamster models(Shuai, 2021).

As expected from suboptimal cleavage of spike in producer cells, and the known link between efficient spike cleavage and ability to use TMPRSS2 for cell entry/cell–cell fusion, we observed that the Omicron spike protein has low fusogenic potential. This phenomenon could translate to impaired cell–cell spread and, indeed, we observed smaller plaque sizes for Omicron compared with Delta in a system in which cell-free infection was limited by semi-solid medium. The smaller focus size for Omicron could be attributable to a failure to induce syncytia, without compromised entry efficiency given that a similar number of foci was observed. The semi-solid medium prevented cell-free infection and our measurement should therefore primarily reflect infection spread over a limited time period through syncytia formation. However, the smaller foci could have

differing implications in vivo. First, there may be decreased virus production due to less cell death, therefore favouring Omicron over Delta. Alternatively, syncytia along with cell survival and larger infection foci may increase spread, thereby favouring Delta. The difference in outcome from syncytia is likely to be cell and tissue dependent and we caution against extrapolating these in vitro infection data to viral loads in the respiratory tract.

Our data showing tropism differences for Omicron in organoid systems and human nasal epithelial cultures are limited by the fact that they are in vitro systems, albeit using primary human tissue. Note also that levels of TMPRSS2 may affect ACE2, particularly as TMPRSS2 has been implicated in ACE2 cleavage(Heurich, 2014) and our effect sizes were affected by ACE2 expression. Finally, E64d may inhibit viral protease in addition to host cellular cathepsins, and this may have affected the strength of the observed phenotype.

In summary, Omicron has gained immune evasion properties while compromising cell entry in TMPRSS2-expressing cells such as those in alveoli, as well as compromising syncytia formation—a combination of traits consistent with reduced pathogenicity in vivo. Crucially, experience with Omicron has demonstrated that predictions regarding replication and tropism based on sequence alone can be misleading, and a detailed molecular understanding of the phenotypic impact of complex sets of mutations is vital as variants continue to emerge.

12.9 Methods

Serum and tissue samples and ethical approval

Ethical approval for study of vaccine-elicited antibodies in sera from vaccinated individuals was obtained from the East of England–Cambridge Central Research Ethics Committee Cambridge (REC ref: 17/EE/0025). The participants provided informed consent. The samples of human upper and lower airways were obtained from the lungs of a multiorgan donor whose lungs were deemed to be unsuitable for clinical transplantation and after their next of kin consented to their use in research. The studies using human donor lung tissue were approved by National Research Ethics Committee (NREC) 16/NE/0230.

All protocols involving samples from human participants recruited at Kyoto University, Kuramochi Clinic Interpark and Universidad San Francisco de Quito were reviewed and approved by the Institutional Review Boards of Kyoto University (approval ID: G0697), Kuramochi Clinic Interpark (approval ID: G2021-004) and Universidad San Francisco de Quito (approval ID: CEISH P2020-022IN) and the Ecuadorian Ministry of Health (approval IDs: MSP-CGDES-2020-0121-O and MSP-CGDES-061-2020). The export of sera from Ecuador to Japan was approved by ARCSA ID: ARCSA-ARCSA-CGTC-DTRSNSOYA-2021-1626-M. All human participants provided written informed consent. All protocols for the use of human specimens were reviewed and approved by the Institutional Review Boards of The Institute of Medical Science, The University of Tokyo (approval ID: 2021-1-0416), Kumamoto University (approval IDs: 2066 and 2074) and University of Miyazaki (approval ID: O-1021)

Omicron spike model showing surface mutations with annotation

Spike genomes for the original Wuhan strain, and Omicron VOC were obtained from GISAID EpiCoV database accessed on 30th November 2021. A consensus Spike genome was created from all complete and high-coverage genomes, excluding all sequences with >5% Ns using Geneious Prime v.2022. The consensus genome was translated to poly-proteins and the spike gene was aligned to the original Wuhan strain using mafft (v.7.490)(Katoh, 2013) with the parameters --globalpair --maxiterate 1000 flags.

Alternatively, 3D structural models of the spike homotrimer protein complex were generated using Alphafold (v.2.1.1)(Jumper, 2021). In its validation at the 14th edition of the Critical Assessment of protein Structure Prediction (CASP14), the predictions generated were demonstrated to be comparative to experimental structures. When a close homologue structure is available to Alphafold, the predictions that it generates for those positions are within typical experimental error. The required databases were downloaded on 2 December 2021. The program was run with the parameters --max_template_date=2021-12-01 --model_preset=monomer --db_preset=full_dbs --is_prokaryote_list=false. Protein structures were visualized in ChimeraX (v.1.3)(Goddard, 2018). As predicted structures for the whole spike protein include poorly resolved chains at the terminal ends, these residues were identified by overlaying structures on Protein Data Bank (PDB) entry 6ZP2, then selected and removed from PDB files using the delete atoms/bonds action. Two further monomers were overlayed on 6ZP2 to generate a homotrimer structure. Mutated residues were then coloured in red and labelled manually with respect to the Wuhan strain.

Spike model comparing Delta and Omicron

The SARS-CoV-2 Delta (B.1.617.2) and Omicron (B.1.529) models in Extended **Data Fig. 12.1a** were generated using the cryo-EM structure PDB 7A94 as a reference. The experimental structure has been resolved in open conformational state with one RBD in the 'up' conformation bound to ACE2. To generate the complete structure, the missing loop regions in the structure (71–75, 625–632, 677–688, 828–852 and 941–943) were modelled using the Modeller plug-in(Cele, 2021). The hACE2 structure in the template also had six residues missing patches that were added by using PDB 1R42 as a reference(Hoffman, 2020). The top model was selected on the basis of the zDOPE score and the least root-mean-square deviation (r.m.s.d.) from the reference structure. To create variant models, mutations were induced using the Dunbrack 2010 backbone-dependent rotamer library in Chimera4. Further processing of steps involving minimization and peptide bond building (after deletions for B.1.617.2 and B.1.529 spike) were performed using the Gromacs. The resultant Delta and Omicron modelled complexes have 3,988 and 3,994 residues, respectively, with spike trimer consisting of three chains (14–1,146 residues). The structural superimposition of the Delta and Omicron spike model with the cryo-EM structure displayed an overall r.m.s.d. of 0.66 Å and 0.87 Å, respectively. All of the molecular images were rendered using ChimeraX (Goddard, 2018).

Biolayer interferometry analysis

Assays were performed on an Octet Red (ForteBio) instrument at 30 °C with shaking at 1,000 rpm. Streptavidin (SA) biosensors were hydrated in water for 10 min before a 1 min incubation in undiluted 10× kinetics buffer. SARS-CoV-2 Wuhan, Delta or

Omicron RBDs were loaded at 2.5 μg ml^{-1} in 10× kinetics buffer for 100–150 s before baseline equilibration for 120 s in 10× kinetics buffer. Human ACE2 association with Wuhan, Delta and Omicron RBDs was assessed in 10× kinetics buffer at various concentrations in a threefold dilution series from 333.3 to 4.1 nM and carried out for 300 s before dissociation for 300 s. The data were normalized to the baseline and fitting was performed using a 1:1 binding model and the ForteBio data analysis software. Kinetics values (K_D, k_{on}, k_{off}) were determined with a global fit applied to all data.

The SARS-CoV-2-RBD-Avi constructs were synthesized by GenScript into pcDNA3.1- with an N-terminal mu-phosphatase signal peptide and a C-terminal octa-histidine tag, flexible linker and avi tag (GHHHHHHHHGGSSGLNDIFEAQKIEWHE). The boundaries of the construct are N-328RFPN331 and 528KKST531-C(Walls, 2020; McCallum, 2021). The human angiotensin-converting enzyme ectodomain (hACE2) consists of residues 1–614 fused to a C-terminal octa-histidine tag(Starr, 2020). Proteins were produced in Expi293F Cells (Thermo Fisher Scientific) grown in suspension using Expi293 Expression Medium (Thermo Fisher Scientific) at 37 °C in a humidified 8% CO_2 incubator rotating at 130 rpm. Cells grown to a density of 3 million cells per ml were transfected using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific) and cultivated for 4 days. Proteins were purified from clarified supernatants using a nickel HisTrap HP affinity column (Cytiva) and washed with ten column volumes of 20 mM imidazole, 25 mM sodium phosphate pH 8.0 and 300 mM NaCl before elution on a gradient to 500 mM imidazole. Proteins were buffer-exchanged into 20 mM sodium phosphate pH 8 and 100 mM NaCl and concentrated using centrifugal filters (Amicon Ultra) before being flash-frozen.

Flow cytometry analysis of spike protein surface expression

ExpiCHO cells were seeded into 50 ml Mini Bioreactor Centrifuge Tube (Corning) at 6×10^6 cells per ml in 5 ml ExpiCHO Expression medium (Life technology). Plasmids encoding SARS-CoV-2 Wuhan, Alpha (B.1.1.7), Delta (B.1.617.2) or Omicron (B.1.1.529) spike protein (5 mg) were diluted in OptiPRO SFM (Life Technology) and mixed with Expifectamine CHO Reagent (Life Technology). After incubation for 1 min at room temperature, transfection mixes were added to the cell suspensions. Next, the cells were incubated at 37 °C under 8% CO_2 with an orbital shaking speed of 120 rpm for the next 48 h.

Transiently transfected cells were collected and washed with PBS, 1% BSA and 2 mM EDTA. Cells were counted, dispensed into round-bottom 96-well plates (Corning) and incubated with human IgG Fc-conjugated ACE2 serial dilutions (concentration range: 30,000–0.17 ng ml^{-1}) for 45 min at 4 °C. After two washes, Alexa Fluor647 Goat Anti-Human IgG secondary antibodies (1.5 mg ml^{-1}) (Jackson Immunoresearch) was added to the cells, which were then incubated for 30 min at 4 °C in the dark. The cells were washed twice and resuspended in wash buffer for data acquisition at ZE5 cytometer (Bio-Rad). To assess spike protein expression level, an aliquot of each transfectant cell line was stained with 10 mg ml^{-1} of S2P6 antibody(Cathcart, 2021) for 45 min at 4 °C.

Analysis of single-nucleus and single-cell RNA-seq datasets

For the tissue data(Madissoon, 2021), only single-nucleus RNA-seq were selected. Scanpy v.1.7.1 was used to process the data. *ACE2* and *TMPRSS2* expression was plotted using the violin plot function in scanpy.

Cells and cell culture

Caco-2 cells transduced with BVDV NPro (Caco-NPro) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 % FCS, non-essential amino acid (NEAA), glutamine and 100 U ml^{-1} penicillin and 100 µg ml^{-1}streptomycin at 37 °C with 5 % CO_2. HOS-ACE2/TMPRSS2 cells (HOS cells stably expressing human ACE2 and TMPRESS2)[6,44] were maintained in Dulbecco's modified Eagle's medium (high glucose) (Wako, 044-29765) containing 10% foetal bovine serum (FBS) and 1% penicillin–streptomycin. VeroE6/TMPRSS2 cells (an African green monkey (*Chlorocebus sabaeus*) kidney cell line; JCRB1819)(Maatsuyama, 2020) were maintained in Dulbecco's modified Eagle's medium (low glucose) (Wako, 041-29775) containing 10% FBS, G418 (1 mg ml^{-1}; Nacalai Tesque, G8168-10ML) and 1% penicillin–streptomycin. Calu-3 (a human lung epithelial cell line; a gift from P. Lehner, or ATCC HTB-55) were maintained in Eagle's minimum essential medium (Sigma-Aldrich, M4655-500ML) containing 10% FBS and 1% penicillin–streptomycin.Vero E6 (ATCC-CRL1586) and Vero-ACE2/TMPRSS2 (a gift from E. Thomson) were maintained in Dulbecco's modified Eagle's medium (low glucose) (Wako, 041-29775) containing 10% FBS, G418 (1 mg ml^{-1}; Nacalai Tesque, G8168-10ML) and 1% penicillin–streptomycin. HeLa-ACE2 (a gift from J. Voss), A549-ACE2/TMPRSS2 (a gift from M. Palmarini) and HeLa-ACE2/TMPRSS2 were maintained in DMEM containing 10% FBS and 1% penicillin–streptomycin. HEK293T (CRL-3216) cells, and their derivative cell lines, including 293T-ACE2/TMPRSS2, 293T-ACE2ΔTMPRSS2, 293T-TMPRSS2 and 293T-GFP11, have been described previously[22] and were provided by L. James. All the HEK293T cell lines as well as

Vero-GFP1-10 (also a gift from L. James) were maintained in DMEM with 10% FBS and 1% penicillin–streptomycin.

H1299 cells (a gift from S. Cook) were maintained in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% FCS and 1% penicillin–streptomycin. All of the cells were regularly tested and are mycoplasma free. Airway epithelial organoids and gallbladder organoids were obtained and maintained as previously described (Youk, 2020; Brevini, 2021). For airway organoids, human distal lung parenchymal tissues were obtained from adult donors with no background lung pathologies from Papworth Hospital Research Tissue Bank (T02233). Airway organoids were cultured in 48-well plate were passaged every 2 weeks. For passaging, Matrigel was melted by incubation with dispase at 37 °C for 30 min, followed by single-cell dissociation by TrypLE (Gibco) at 37 °C for 5 min. Then, 5,000 cells were resuspended in fresh GFR-Matrigel, followed by plating in 48-well plates. Primary cholangiocyte organoids were isolated from primary biliary tissue (intrahepatic ducts, common bile duct and gallbladder) and cultured using established methodology as previously reported(Brevini, 2021).

Virus isolates

The lineages B .1.1.617.2 (Delta, GISAID: EPI_ISL_1731019) and B .1.1.529 (ref. Dejnirattisai, 2022) (Omicron UK isolate(Dejnirattisai, 2022), G. Screaton) were received as part of the work conducted by G2P-UK National Virology Consortium (Barclay PI). Virus was propagated on Vero-ACE2-TMPRSS2 (VAT) cells for 3 days

until cytopathic effect (CPE) was observed. These viruses were sequence-confirmed and were used for western blotting and Caco-3 gut experiments.

A second Omicron isolate from South Africa was obtained (GISAID: EPI_ISL_7886688), and used specifically as a second isolate for western blot analysis of live virions and for focus-size experiments. To isolate virus, ACE2-expressing H1299 cells were seeded at 4.5×10^5 cells in a six-well plate well and incubated for 18–20 h. After one DPBS wash, the sub-confluent cell monolayer was inoculated with 500 µl universal transport medium diluted 1:1 with growth medium filtered through a 0.45 µm filter. Cells were incubated for 1 h. Wells were then filled with 3 ml complete growth medium. After 4 days of infection (completion of passage 1 (P1)), cells were trypsinized, centrifuged at 300 rcf for 3 min and resuspended in 4 ml growth medium. All of the infected cells were then added to Vero E6 cells that had been seeded at 2×10^5 cells per ml, 20 ml total, 18–20 h earlier in a T75 flask for cell-to-cell infection. The coculture of ACE2-expressing H1299-E3 (clone 3) and Vero E6 cells was incubated for 1 h and the flask was then filled with 20 ml of complete growth medium and incubated for 4 days. The viral supernatant (passage 2 (P2) stock) was used for experiments. P2 stock was sequenced and confirmed to be Omicron BA.1. The sequence was deposited in GISAID under accession number EPI_ISL_7886688.

A third Omicron isolate (BA.1 lineage, strain TY38-873; GISAID: EPI_ISL_7418017) with same mutational profile as other Omicron variant isolates and provided by the National Institute for Infectious Diseases, Japan, was used for hNEC, Calu-3 and HeLa-ACE2-TMPRSS2 infection experiments. Virus isolation was performed as previously described(Matsuyama, 2020). In brief, the saliva was inoculated into

VeroE6/TMPRSS2 cells and CPE was observed 4 days after inoculation. The supernatant was then collected and stored at −80 °C as an original virus (GISAID: EPI_ISL_7418017). After one more passage in VeroE6/TMPRSS2 cells, the following virus was additionally obtained from National Institute of Infectious Diseases, Japan for hNEC experiments: Delta isolate (B.1.617.2 lineage, strain TKYTK1734; GISAID: EPI_ISL_2378732)(Saito,2021) Virus preparation and titration of virus isolates was performed as previously described **Saito, 2021; Motozono, 2021**. To prepare the working virus stock, 100 µl of the seed virus was inoculated into VeroE6/TMPRSS2 cells (5×10^6 cells in a T-75 flask). In brief, 1 h after infection, the culture medium was replaced with fresh culture medium. At 3 days after infection, the culture medium was collected and centrifuged, and the supernatants were collected as the working virus stock. The titre of the prepared working virus was measured as the $TCID_{50}$ in VeroE6/TMPRSS2 cells. Serially diluted virus stocks were inoculated into the cells and incubated at 37 °C for 4 days. The cells were observed under microscopy to judge the CPE appearance. The value of $TCID_{50}$ per ml was calculated using the Reed–Muench method.

Serum collection for live-virus neutralization experiments

Vaccine sera were collected from 14 vaccinated individuals 4 weeks after the second vaccination with BNT162b2 (Pfizer-BioNTech) (average age: 46, range: 38–55, 21% male). The sera obtained from 7 vaccinated individuals 2–3 weeks after the second vaccination with ChAdOx1 (Oxford-AstraZeneca) (average age: 46, range: 35–54, 71% male) were purchased from BioIVT.

Convalescent sera were collected from ten vaccine-naive individuals who had been infected with Delta variant (B.1.617.2 or AY.29) (average age: 46, range: 22–63, 70% male). To determine the SARS-CoV-2 variants that individuals were infected with, saliva samples were collected from patients with COVID-19 during onset and RNA was extracted using the QIAamp viral RNA mini kit (Qiagen, Qiagen, 52906) according to the manufacturer's protocol. The sample was analysed using whole-genome sequencing based on a modified ARTIC Network protocol (Itokawa, 2020), and the near-full-length SARS-CoV-2 genome sequences were obtained. Alternatively, we also performed the capture-hybridization method. The RNA isolated from saliva was treated with DNase I (Takara, EN0521) and the sequencing library was prepared using the Twist library preparation kit (Twist Bioscience, 101058). The capture-hybridization was conducted using xGen COVID-19 capture panel and xGen hybridization and wash kit (Integrated DNA Technologies, 1080578) according to the manufacturer's instruction. Illumina sequencing was performed using the MiSeq reagent kit v2 (300 cycles) and a MiSeq sequencer (Illumina). For the data analysis, trimmomatic-0.39 (ref. Bolger, 2014) was used to remove the adapters and low-quality reads from the raw sequencing data. The trimmed paired-end reads were aligned to the human genome hg38 using bowtie2 (v.2.3.4.3)(Langmead, 2012) and unmapped reads were mapped to the original SARS-CoV-2 genome (strain Wuhan-Hu-1, GenBank accession no. NC_045512.2) using Bwa-mem2 (https://github.com/bwa-mem2/bwa-mem2). The PCR duplicates were removed by gencore (v.0.16.0)(Chen, 2019) and a consensus sequence was obtained using IGV (v.2.10.2)(Thorvaldsdottir, 2013). The mutations detected and viral lineage were determined by using CoVsurver (https://corona.bii.a-star.edu.sg) and Pangolin COVID-

19 lineage assigner (https://pangolin.cog-uk.io/). The twelve convalescent sera during early pandemic (until April 2020) (average age: 71, range: 52–92, 8% male) were purchased from RayBiotech. Sera were inactivated at 56 °C for 30 min and stored at −80 °C until use.

hNEC infection

Primary human nasal epithelial cells (EP02, MP0010) were purchased from Epithelix and maintained according to the manufacturer's procedure. The infection experiment using primary human nasal epithelial cells was performed as follows. In brief, the working Omicron (EPI_ISL_7418017) and Delta (EPI_ISL_2378732) viruses were diluted with Opti-MEM (Thermo Fisher Scientific, 11058021). The diluted viruses (1,000 $TCID_{50}$ in 100 µl) were inoculated onto the apical side of the culture and incubated at 37 °C for 1 h. The inoculated viruses were removed and washed twice with Opti-MEM. To collect the viruses on the apical side of the culture, 100 µl Opti-MEM was applied onto the apical side of the culture and incubated at 37 °C for 10 min. The Opti-MEM applied was collected and analysed using qPCR with reverse transcription (RT–qPCR) to quantify the viral RNA copy number (see below).

Caco-2 gut cell infection

Caco-2-Npro cells were prepared in 48-well plates and infected in biological triplicates with either SARS-CoV-2 Delta (EPI_ISL_1731019) or Omicron (M21021166) variant at a m.o.i. of 1 or 0.1 $TCID_{50}$ per cell in 250 µl DMEM containing 2% FBS for 2 h at 37 °C with 5% CO_2. Unabsorbed inoculum was then removed, cells were washed and

the medium was replaced with DMEM containing 2% FBS. Virus infections were maintained up to 72 h. Cell culture supernatants were collected at 0, 24 and 48 h after infection, and viral RNAs were extracted and quantified using RT–qPCR using specific primer probe against the *E* gene. An absolute quantification against a standard curve generated was carried out using in vitro-transcribed RNA from fragment 11 clone provided by V. Thiel. Data were collected using the ViiA 7 Real-Time PCR System (Applied Biosystems). Moreover, tenfold serial dilutions of collected virus supernatants were prepared in DMEM culture medium for measuring the infectious dose by $TCID_{50}$. Virus titres were recorded at 72 h after infection for CPE appearance and expressed as $TCID_{50}$ per ml values using the Reed–Muench method .

Calu-3 lung cell and HeLa-ACE2-TMPRSS2 live virus infection

One day before infection, Calu-3 cells (20,000 cells) and HeLa-ACE2/TMPRSS2 cells (10,000 cells) were seeded onto a 96-well plate. SARS-CoV-2 (100 $TCID_{50}$ for HeLa-ACE2/TMPRSS2 cells; and 2,000 $TCID_{50}$ for Calu-3 cells) Omicron (EPI_ISL_7418017) or Delta (EPI_ISL_1731019) was inoculated and incubated at 37 °C for 1 h. The infected cells were washed, and 180 µl of culture medium was added. The culture supernatant (10 µl) was collected at the indicated time points and used for real-time RT–PCR to quantify the viral RNA copy number(Saito, 2021; Motozono, 2021). In brief, 5 µl of culture supernatant was mixed with 5 µl of 2× RNA lysis buffer (2% Triton X-100, 50 mM KCl, 100 mM Tris-HCl (pH 7.4), 40% glycerol, 0.8 U μl^{-1} recombinant RNase inhibitor (Takara, 2313B)) and incubated at room temperature for 10 min. RNase-free water (90 µl) was added, and the diluted sample (2.5 µl) was used as the template for RT–qPCR performed according to the manufacturer's protocol using the

One Step TB Green PrimeScript PLUS RT-PCR kit (Takara, RR096A) and the following primers: forward *N*, 5′-AGC CTC TTC TCG TTC CTC ATC AC-3′; and Reverse *N*, 5′-CCG CCA TTG CCA GCC ATT C-3′. The viral RNA copy number was standardized using the SARS-CoV-2 direct detection RT-qPCR kit (Takara, RC300A). Fluorescent signals were acquired using the QuantStudio 3 Real-Time PCR system (Thermo Fisher Scientific), a CFX Connect Real-Time PCR Detection system (Bio-Rad), an Eco Real-Time PCR System (Illumina) or a 7500 Real Time PCR System (Applied Biosystems).

Western blot analysis of infected cells and virions

VeroE6-ACE2/TMPRSS2 cells were infected with Delta (EPI_ISL_1731019), Omicron isolate 1 (M21021166) and 2 (EPI_ISL_7886688) at an m.o.i. of 1. At 24 and 48 h of infection, cells and culture medium were collected. The culture media were centrifuged, and the supernatants were collected. To pellet down SARS-CoV-2 in the supernatants, an equal volume of clarified supernatants was mixed with 20% PEG6000 in PBS and centrifuged at 12,000*g* for 30 min at 4 °C (ref. Papa, 2021), followed by pellet resuspension in 1× SDS sample buffer.

For cell lysates, the collected cells were washed and lysed in lysis buffer (Cell Signalling) and the lysates were diluted with 4 × sample buffer (Bio-Rad) and boiled for 10 min before analysed using western blotting. For protein detection, the following antibodies were used: rabbit anti-SARS-CoV-2 S monoclonal antibodies (PA1-41165, Thermo Fisher Scientific), mouse anti-SARS-CoV-2 S1 antibodies (MAB105403, R&D systems), rabbit anti-SARS-CoV-2 N monoclonal antibodies (HL344, GeneTex, GTX635679), rabbit anti-GAPDH polyclonal antibodies (10494-1-AP, Proteintech), horseradish peroxidase (HRP)-conjugated anti-rabbit and anti-mouse IgG polyclonal

antibodies (Cell Signalling). Chemiluminescence was detected using the ChemiDoc Touch Imaging System (Bio-Rad). The cleavage ratio of S1 or S2 to FL in virions was determined by densitometry using ImageJ (NIH)

Pseudotype virus preparation for testing against vaccine-elicited antibodies and cell entry

Plasmids encoding the spike protein of SARS-CoV-2 D614 with a C terminal 19 amino acid deletion with D614G were used. Omicron and Delta spikes were generated by gene synthesis. Viral vectors were prepared by transfection of HEK293T cells using the Fugene HD transfection reagent (Promega). HEK293T cells were transfected with a mixture of 11 ul of Fugene HD, 1 µg of pCDNAΔ19 spike, 1 ug of p8.91 HIV-1 gag-pol expression vector and 1.5 µg of pCSFLW (expressing the firefly luciferase reporter gene with the HIV-1 packaging signal). Viral supernatant was collected at 48 and 72 h after transfection, filtered through 0.45 um filter and stored at −80 °C as previously described. Infectivity was measured by luciferase detection in target cells.

Neutralization titre analyses

Neutralization by vaccine-elicited antibodies after two doses of the BNT162b2 and Chad-Ox-1 vaccine, in addition to a third dose with BNT162b2, was determined by PV infections in the presence of serial dilutions of sera. The ID_{50} values within groups were summarized as a geometric mean titre and statistical comparisons between groups were performed using Mann–Whitney *U*-tests or Wilcoxon signed-rank tests. Statistical analyses were performed using Stata v.13 and Prism v.9.

Preparation of monoclonal antibodies

Casirivimab and Imdevimab were prepared as previously described53. To construct the plasmids expressing anti-SARS-CoV-2 monoclonal antibodies (casirivimab and imdevimab), the sequences of the variable regions of casirivimab and imdevimab were obtained from KEGG Drug Database (https://www.genome.jp/kegg/drug/) and were artificially synthesized (Fasmac). The obtained coding sequences of the variable regions of the heavy and light chains were cloned into the pCAGGS vector containing the sequences of the human immunoglobulin 1 and κ constant region (provided by H. Arase).

To prepare these monoclonal antibodies, the pCAGGS vectors containing the sequences encoding the immunoglobulin heavy and light chains were cotransfected into HEK293T cells using PEI Max (Polysciences, 24765-1). At 48 h after transfection, the cell culture supernatants were collected, and the antibodies were purified using the NAb protein A plus spin kit (Thermo Fisher Scientific, 89948) according to the manufacturer's protocol.

Neutralization assay for monoclonal antibodies with live virus

One day before infection, VeroE6/TMPRSS2 (10,000 cells) were seeded into a 96-well plate. The monoclonal antibodies (casirivimab, imdevimab or casirivimab/imdevimab) and the heat-inactivated human sera were serially diluted with DMEM supplemented with 10% FCS and 1% penicillin–streptomycin. The diluted antibodies and sera were incubated with SARS-CoV-2 (120 $TCID_{50}$) Omicron (EPI_ISL_7418017) or Delta (EPI_ISL_2378732) isolates at 37 °C for 1 h. The viruses without antibodies or sera were included as controls. The mixture (containing the virus

at 100 $TCID_{50}$) was inoculated onto a monolayer of VeroE6/TMPRSS2 cells and incubated at 37 °C for 1 h. The cells were next washed with DMEM and cultured in DMEM supplemented with 10% FCS and 1% penicillin–streptomycin. At 24 h after infection, the culture supernatants were collected and viral RNA was quantified using RT–qPCR. The assay of each antibody or serum was performed in triplicate or quadruplicate, and the 50% neutralization titre was calculated using Prism 9 (GraphPad).

Antiviral drug assay with live virus

One day before infection, HOS-ACE2-TMPRSS2 cells (10,000 cells) were seeded into a 96-well plate. The cells were infected with SARS-CoV-2 (100 $TCID_{50}$) Omicron (EPI_ISL_7418017) or Delta (EPI_ISL_2378732) isolates at 37 °C for 1 h. The cells were next washed with DMEM and cultured in DMEM supplemented with 10% FCS and 1% penicillin–streptomycin and the serially diluted remdesivir (Selleck, S8932) or β-d-N4-hydroxycytidine (a derivative of molnupiravir; Cell Signalling Technology, 81178S). At 24 h after infection, the culture supernatants were collected and viral RNA was quantified using RT–qPCR . The assay of each compound was performed in quadruplicate, and the 50% neutralization titre was calculated using Prism 9 (GraphPad).

Cytotoxicity assay

The CPE of the remdesivir and β-d-N4-hydroxycytidine was tested using the cell counting kit-8 (Dojindo, CK04-11) according to the manufacturer's instructions. One day before the assay, HOS-ACE2-TMPRSS2 cells (10,000 cells) were seeded into a 96-well

plate. The cells were cultured with the serially diluted compound for 24 h. The cell counting kit-8 solution (10 µl) was added to each well, and the cells were incubated at 37 °C for 90 min. Absorbance was measured at 450 nm using the GloMax Explorer Microplate Reader (Promega).

PV SG-PERT and infectivity

The Delta and Omicron spike-PVs were prepared as described in the 'Neutralization titre analyses' section and standardized using a SYBR Green-based product-enhanced PCR assay (SG-PERT) as described previously.54. In brief, tenfold dilutions of virus supernatant were lysed at a 1:1 ratio in a 2× lysis solution (made up of 40% glycerol (v/v) 0.25% Triton X-100 (v/v) 100 mM KCl, RNase inhibitor 0.8 U ml^{-1}, TrisHCL 100 mM, buffered to pH 7.4) for 10 min at room temperature. Sample lysates (12 µl) were added to 13 µl of SYBR Green master mix (containing 0.5 µM of MS2-RNA Fwd and Rev primers, 3.5 pmol ml^{-1} of MS2-RNA and 0.125 U $µl^{-1}$ of Ribolock RNase inhibitor) and cycled in a QuantStudio system. Relative amounts of reverse transcriptase activity were determined as the rate of transcription of bacteriophage MS2 RNA, with absolute RT activity calculated by comparing the relative amounts of RT to an RT standard of known activity.

PVs were used to transduce different target cell lines and 3D organoid cultures. After 48 or 72 h after transduction, the cells were lysed with direct addition of Bright Glo (Promega). The raw readings were then normalized to the SG-PERT and plotted using GraphPad Prism.

Entry inhibitors in A549-ACE2-TMPRSS2 cells with PV and live virus

A549-ACE2/TMPRSS2 (A549-A2T2) cells were treated with either E64d (Tocris) or camostat (Sigma-Aldrich) for 2 h at each drug concentration before the addition of a comparable amount of input viruses pseudotyped with Delta or Omicron (approximately 100,000 RLU). The cells were then left for 48 h before addition of luciferase substrate (Promega) and read on a Glomax plate reader (Promega).

Quantification of viral replication in A549-ACE2/TMPRSS2-based luminescent reporter cells was performed as previously described (ref. Gerber, 2022). In brief, A549-A2T2 reporter cells (clone E8) over-expressing *Renilla* luciferase (Rluc) and SARS-CoV-2 Papain-like protease-activatable circularly permuted firefly luciferase (FFluc) were seeded in flat-bottomed 96-well plates. The next day, the cells were treated with the indicated doses of camostat or E64d for 2 h, then infected with the Delta (EPI_ISL_1731019) or Omicron (M21021166) variant of SARS-CoV-2 at an m.o.i. of 0.01. After 24 h, the cells were lysed in Dual-Glo Luciferase Buffer (Promega) diluted 1:1 with PBS and 1% NP-40. Lysates were then transferred to opaque 96-well plates, and viral replication was quantified as the ratio of FFluc/Rluc luminescence measured using the Dual-Glo kit (Promega) according to the manufacturer's instructions. Percentage inhibition was calculated relative to the maximum luminescence signal for each condition, then analysed using the Sigmoidal, 4PL, X is log(concentration) function in GraphPad Prism.

Immunofluorescence assays

HeLa-ACE2 cells were grown on 19 mm glass coverslips in 24-well plates and infected with Delta and Omicron for 24 h. Cells were fixed in 4% paraformaldehyde

(Applichem), and permeabilized in PBS-T (0.1% (v/v) Triton X-100 in PBS) for 5 min. Cells were blocked in PBS-T containing 5% (w/v) BSA for 30 min and labelled with primary antibodies against SARS-CoV-2 Spike (GTX632604) or anti-GM130[EP892Y]-cis-Golgi (ab52649) in PBS-T and 1% (w/v) BSA for 1 h at room temperature. After washing, the cells were stained with fluorochrome-conjugated secondary antibodies in PBS-T, 1% (w/v) BSA for 1 h at room temperature in the dark. Nuclei were counterstained with DAPI and phalloidin 647 (A22287, Molecular Probes). Confocal images were acquired on a Leica TCS SP8 microscope. Post-acquisition analysis was performed using the Leica LAS X or Fiji (v.1.49) software. For distance analysis, the two-dimensional coordinates of the centroids of spike proteins were calculated using the Analyze Particles module of Fiji (ImageJ). The distance of each particle to the edge of the nucleus, visualized using DAPI staining, was looked up using a Euclidean distance map computed with the Distance Transform module of Fiji and exported as a list of distance measurements using the Analyze Particle function. Colocalization of Spike with the Golgi was estimated through Pearson's correlation coefficients (*R*) using the PSC colocalization plug-in (ImageJ, NIH). *R* values between −1 (perfect negative correlation) and +1 (perfect positive correlation), with 0 corresponding to no correlation, were assigned[56]. Colocalization analyses were performed on >20 cells.

Cell–cell fusion assay

Cell–cell fusion assays were carried out as follows. In brief, HEK293T GFP11 and Vero-GFP1-10 cells were seeded at 80% confluence at a 1:1 ratio in 48-well plates the day before. Cells were co-transfected with 0.5 µg of spike expression plasmids using Fugene 6 according to the manufacturer's instructions (Promega). Cell–cell fusion

was measured using an Incucyte and determined as the proportion of green area to total phase area. Graphs were generated using Prism 8. To measure cell surface spike expression, HEK293 cells were transfected with *S* expression plasmids and stained with rabbit anti-SARS-CoV-2 S S1/S2 polyclonal antibodies (Thermo Fisher Scientific, PA5-112048, 1:100). Normal rabbit IgG (SouthernBiotech, 0111-01, 1:100) was used as a negative control, and APC-conjugated goat anti-rabbit IgG polyclonal antibodies (Jackson ImmunoResearch, 111-136-144, 1:50) were used as secondary antibodies. The surface expression level of S proteins was analysed using FACS Canto II (BD Biosciences) and FlowJo v.10.7.1 (BD Biosciences). The gating strategy for flow cytometry is shown in **Supplementary Fig. 12.8**.

Analysis of the size of infection foci

Vero E6 cells (ATCC CRL-1586, obtained from Cellonex) were propagated in complete growth medium consisting of DMEM with 10% fetal bovine serum (Hyclone) containing 10 mM HEPES, 1 mM sodium pyruvate, 2 mM L-glutamine and 0.1 mM nonessential amino acids (Sigma-Aldrich). Vero E6 cells were passaged every 3–4 days. H1299 cell lines were propagated in growth medium consisting of complete RPMI 1640 medium with 10% fetal bovine serum containing 10 mM HEPES, 1 mM sodium pyruvate, 2 mM L-glutamine and 0.1 mM non-essential amino acids. H1299 cells were passaged every second day. The H1299-E3 (H1299-ACE2, clone E3) cell line was derived from H1299 (CRL-5803).

Virus stock using a South Africa-derived (EPI_ISL_7886688) isolate, and Delta variant isolate, was produced as described previously(Cele,2021) and used at approximately 50–100 focus-forming units per microwell. H1299-E3 cells were plated in

a 96-well plate (Corning) at 30,000 cells per well 1 day before infection. Cells were infected with 100 µl of the virus working stock for 1 h, then 100 µl of a 1× RPMI 1640 (Sigma-Aldrich, R6504), 1.5% carboxymethylcellulose (Sigma-Aldrich, C4888) overlay was added without removing the inoculum. Cells were fixed 2 h or 18 h after infection using 4% PFA (Sigma-Aldrich) for 20 min. Foci were stained with a rabbit anti-spike monoclonal antibody (BS-R2B12, GenScript A02058) at 0.5 µg ml^{-1} in a permeabilization buffer containing 0.1% saponin (Sigma-Aldrich), 0.1% BSA (Sigma-Aldrich) and 0.05% Tween-20 (Sigma-Aldrich) in PBS. Plates were incubated with primary antibodies overnight at 4 °C, then washed with wash buffer containing 0.05% Tween-20 in PBS. Secondary goat anti-rabbit HRP-conjugated antibodies (Abcam, ab205718) was added at 1 µg ml^{-1} and incubated for 2 h at room temperature with shaking. TrueBlue peroxidase substrate (SeraCare, 5510-0030) was then added at 50 µl per well and incubated for 20 min at room temperature. Plates were imaged using the ImmunoSpot Ultra-V S6-02-6140 Analyzer ELISPOT instrument with BioSpot Professional built-in image analysis (CTL). For microscopy, wells were imaged using a Metamorph-controlled Nikon TiE motorized microscope with a ×10 objective. Automated image analysis was performed using a custom script in MATLAB 2019b. Statistics and graphing were performed using Prism v.8.0.

Data Availability

The sequences of the virus isolates used are available at GISAID: B.1.1.617.2 (Delta, GISAID: EPI_ISL_1731019), B.1.1.529 (Omicron, BA.1 isolate from the UK; ENA project no PRJEB50520, accession number ERS10754850), B.1.1.529 (Omicron, BA.1 isolate from South Africa, GISAID: EPI_ISL_7886688) and (BA.1 isolate from Japan,

GISAID: EPI_ISL_7418017). The structural models used in this study for Delta and Omicron variants are available at GitHub (https://github.com/CSB-Thukral-Lab/Spike_structural_models_Delta_and_Omicron and https://github.com/ojcharles/viral_alphafold). The scRNA-seq data analysis is available at GitHub (https://github.com/elo073/Lung5loc_scRNAseq_omcr_entrygenes). Source data are provided with this paper.

Code Availability

The code for scRNA-seq data analysis is available at GitHub (https://github.com/elo073/Lung5loc_scRNAseq_omcr_entrygenes).

12.10 Figures and Tables

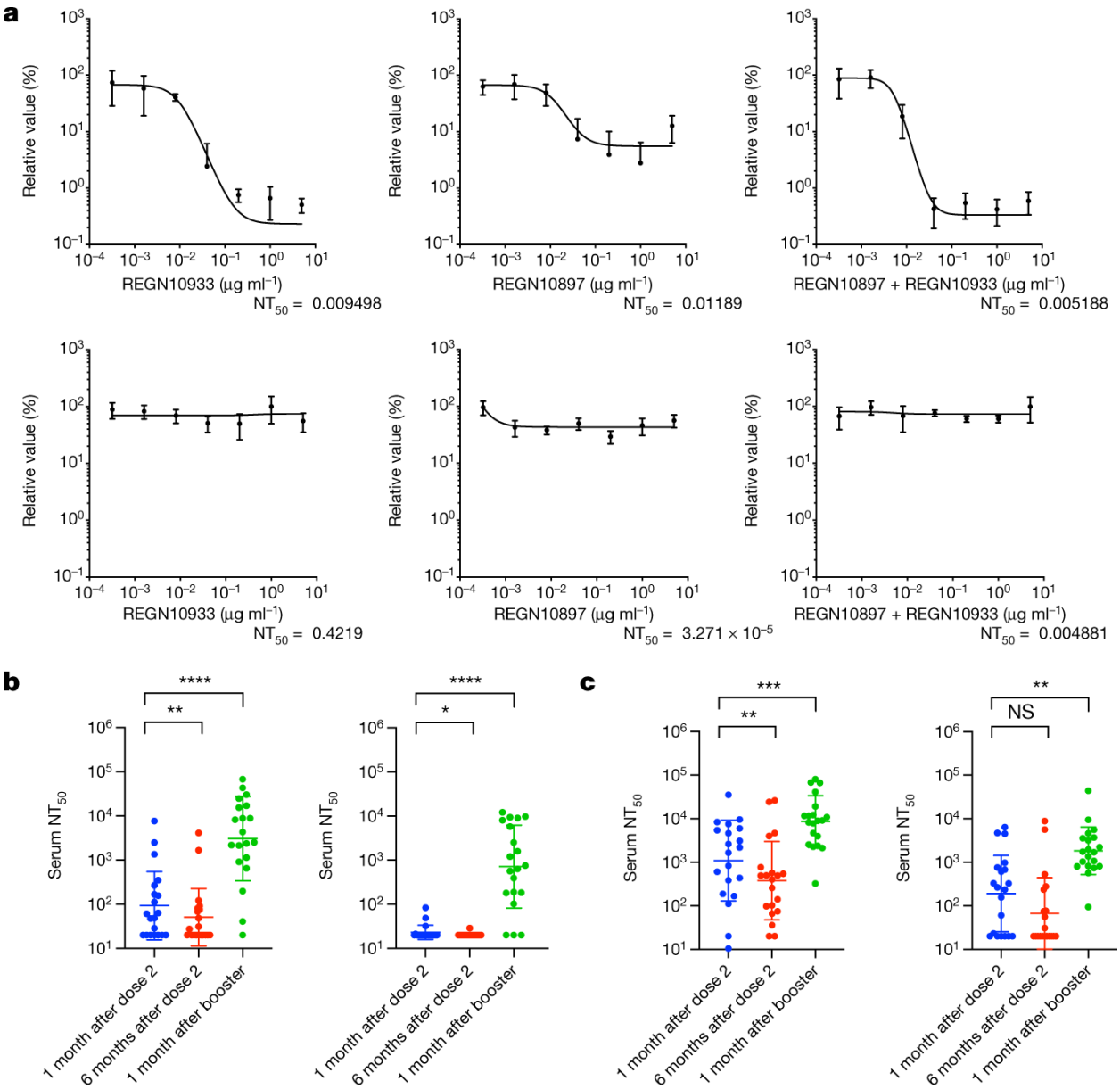


Figure 12.1: SARS-CoV-2 Omicron and Delta Variant live virus replication in 3D tissue culture systems and 2D cell lines.

a, Titration of the monoclonal antibodies REGN10933 and REGN10897 individually and a combination of both against replication-competent Delta (top) and Omicron (bottom) viruses. Data are mean ± s.e.m. $n = 4$ of technical replicates at each dilution for each virus. Curve fitting for dose response was performed in GraphPad Prism. NT_{50}, 50% neutralization titre. **b**, **c**, Neutralization of Delta (left) and Omicron (right) spike PV by sera from vaccinated individuals (ChAdOx-1 (**b**; $n = 20$) or BNT12b2 (**c**; $n = 20$)) over

three time points after dose two (ChAdOx-1 or BNT162b2) and dose three (BNT162b2 only). Data are the geometric mean titre ± s.d., representative of two independent experiments each with two technical replicates. Statistical analysis was performed using Wilcoxon matched-pairs signed-rank tests; $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$; NS, not significant.

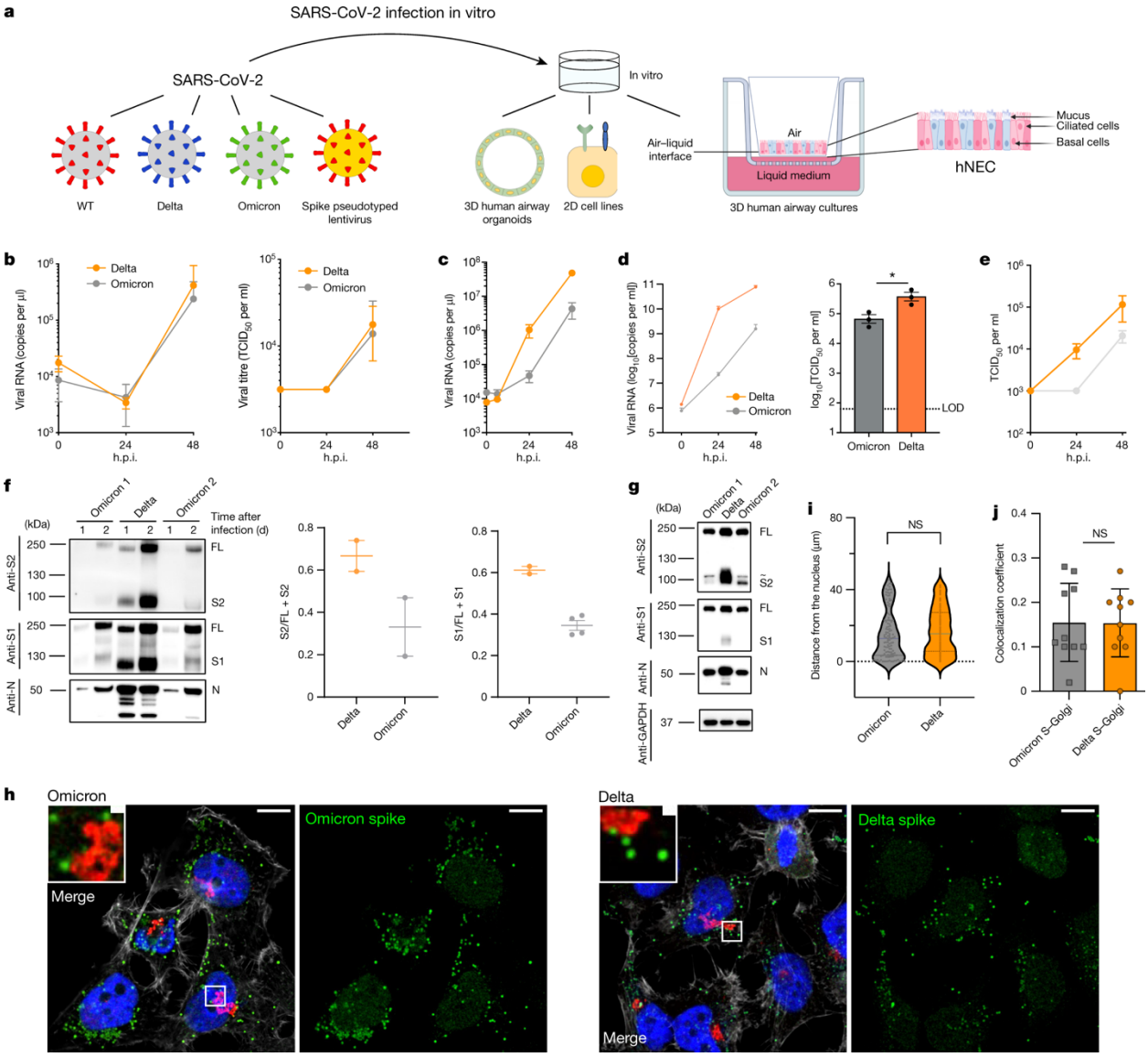


Figure 12.2: SARS-CoV-2 Omicron and Delta Variant live virus replication in 3D tissue culture systems and 2D cell lines.

a, Overview of the viruses and culture systems used. The schematic was created using BioRender.com. **b**–**e**, The spread of infection by replication-competent Omicron versus Delta variants over time in hNECs (**b**), and Calu-3 (**c**), Caco-2-Npro (**d**) and HeLa-A2T2 (**e**) cells. Viral RNA and/or infectious virus in the supernatant ($TCID_{50}$) were measured. Data are mean ± s.d. of the technical replicates shown, representative of two independent experiments. Statistical analysis was performed using two-sided unpaired Student's *t*-tests; $^{*}P < 0.05$. h.p.i., hours post-infection. **f**, Western blot analysis of two

live Omicron isolates and one Delta virus isolate in Vero-ACE2/TMPRSS2 cells probed with antibodies against S2, S1 and N, with quantification of the ratio of S2 and S1 to total spike. Data are mean ± s.e.m. of 2–4 biological replicates. **g**, Vero-ACE2/TMPRSS2 producer cell lysates infected with live isolates probed with antibodies against S2, S1 and N. GAPDH was used as the loading control. A non-specific band above the S2 band is indicated (~). Data are representative of two independent experiments. **h**–**j**, The subcellular localization of spike in cells infected with SARS-CoV-2 Delta versus Omicron. The subcellular distribution of Omicron and Delta spike proteins in HeLa-ACE2 cells infected with live virus isolates. **h**, Cells on coverslips were infected for 24 h, fixed and stained with anti-spike, anti-GM130-cis-Golgi, phalloidin 647 and 4′,6′-diamidino-2-phenylindole dihydrochloride (DAPI), and imaged on a Leica TCS SP8 confocal microscope. **i**, The distance of spike proteins from the nucleus at 24 h after infection. Data points ($n = 144$ for each virus) are shown along with the median ± interquartile range. **j**, Quantification of spike–Golgi colocalization in infected cells. Data are mean ± sd. $n = 10$ for each virus. Values were determined using Pearson's correlation coefficient. Scale bars, 10 μm (main images) and 1 μm (insets). Statistical analysis was performed using two-sided unpaired Student's *t*-tests; NS, not significant.

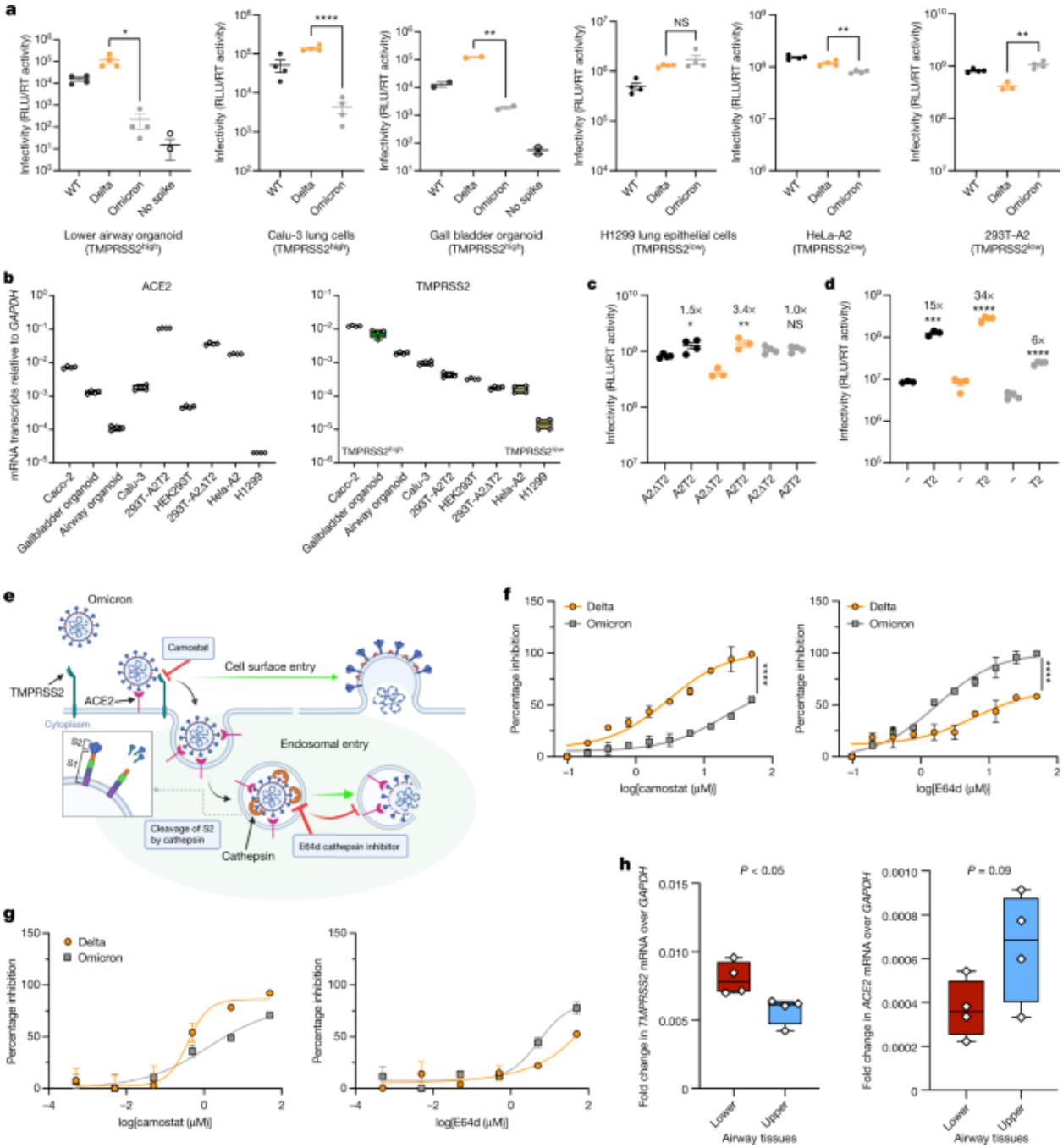


Figure 12.3: SARS-CoV-2 Omicron variant spike enters cells less efficiently by TMPRSS2-dependent plasma membrane fusion.

a, PV entry in airway organoids, Calu-3 lung cells, gall bladded organoids, H1299 lung cells, HeLa *ACE2*-overexpressing cells and HEK293 *ACE2*-overexpressing cells. Black, WT Wuhan-Hu-1 D614G; orange, Delta/B.1.617.2; grey, Omicron/BA.1. Data are mean ± s.e.m. of $n = 2–4$ technical replicates. RLU, relative light unit; RT, reverse transcriptase. Statistical analysis was performed using two-sided unpaired Student's *t*-tests; $*P < 0.05$, $**P < 0.01$, $****P < 0.0001$. Data are representative of three independent experiments. **b**, *ACE2* and *TMPRSS2*mRNA transcripts in the indicated cell types and

organoids as measured using qPCR. Samples were run as $n = 4$ technical replicates. The centre line shows the median, the box limits show the interquartile range and the whiskers denote the range. Data are representative of two independent experiments. **c**, Entry of PV expressing spike in HEK293T cells transduced to overexpress *ACE2* and either depleted for *TMPRSS2* (A2ΔT2) or overexpressing *TMPRSS2* (A2T2). **d**, Entry of PV expressing spike in HEK293T cells with endogenous (−) or overexpressed *TMPRSS2* (T2). For **c** and **d**, data are the mean ± s.e.m. of $n = 4$ technical replicates, representative of three independent experiments. Statistical analysis was performed using two-sided unpaired Student's *t*-tests; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. **e**, An illustration of two cell entry pathways that are known to be used by SARS-CoV-2. The schematic was created using BioRender.com. **f**, Titration of inhibitors in A549-ACE2/TMPRSS2 cells using PV expressing Delta (orange) or Omicron (grey) spike in the presence of the indicated doses of Camostat or E64d, then analysed after 48 h. The percentage inhibition was calculated relative to the maximum luminescence signal for each condition. For each variant and dilution, data are mean ± s.e.m. of experiments conducted in duplicate. Statistical analysis was performed using two-sided unpaired *t*-tests; $^{****}P < 0.0001$. **g**, Titration of inhibitors in A549-ACE2/TMPRSS2-based luminescent reporter cells using live virus. Cells were infected at a multiplicity of infection (m.o.i.) of 0.01 with Delta or Omicron variants in the presence of the indicated doses of Camostat or E64d, and then analysed after 24 h. The percentage inhibition was calculated relative to the maximum luminescence signal for each condition. For each variant and dilution, data are mean ± s.e.m. of experiments conducted in triplicate. Data in **f** and **g** are representative of two independent experiments. **h**, *ACE2* and *TMPRSS2* mRNA expression determined using qPCR in human lung tissue (four pieces of tissue each from upper (main bronchus) and lower (lung parenchyma) airways). Each data point is the mean of $n = 2$ technical replicates. The centre line shows the median of biological replicates ($n = 4$ for each lung region), the box limits show the interquartile range and the whiskers denote the range. Statistical analysis was performed using two-sided unpaired *t*-tests. No adjustments were made for multiple comparisons.

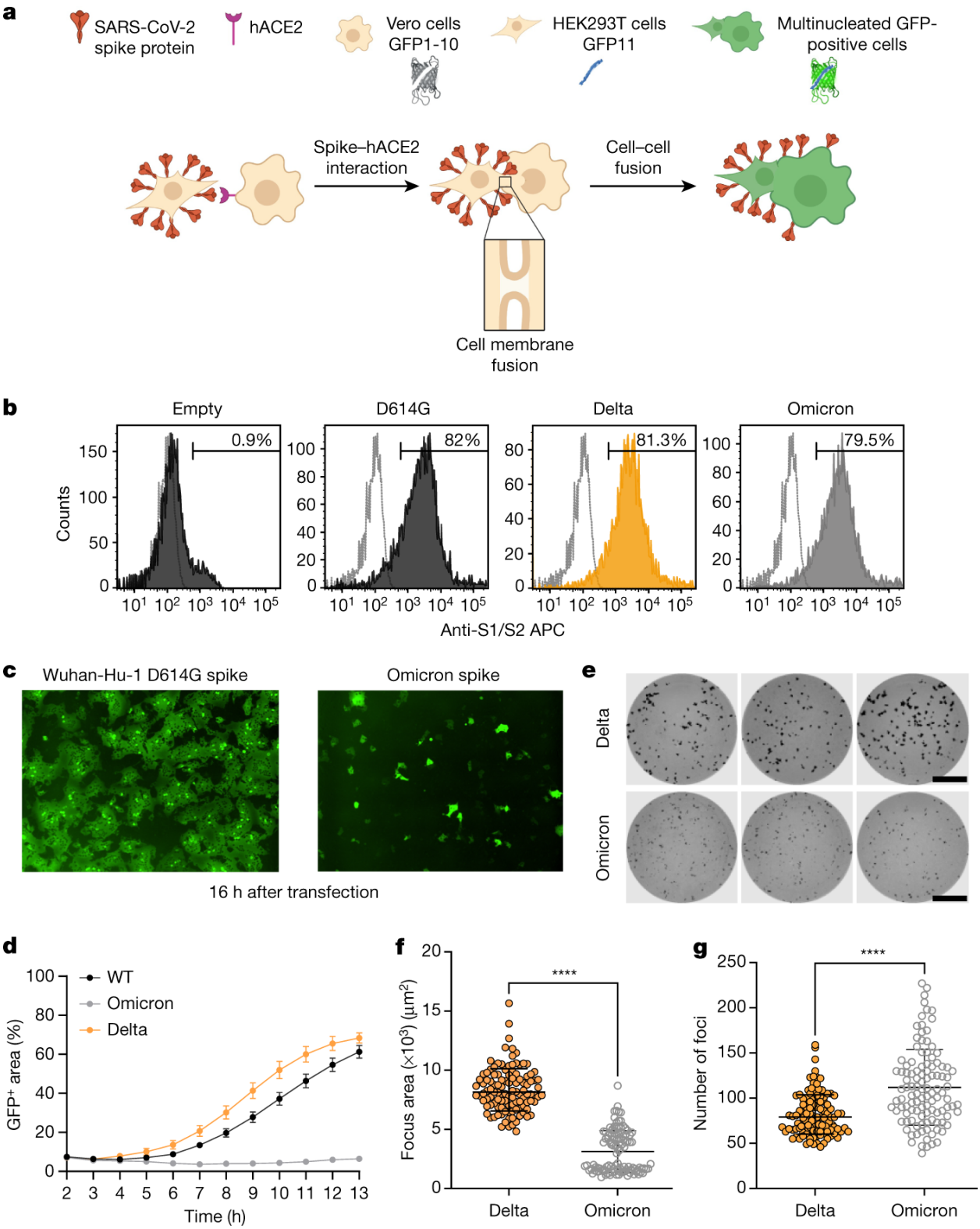
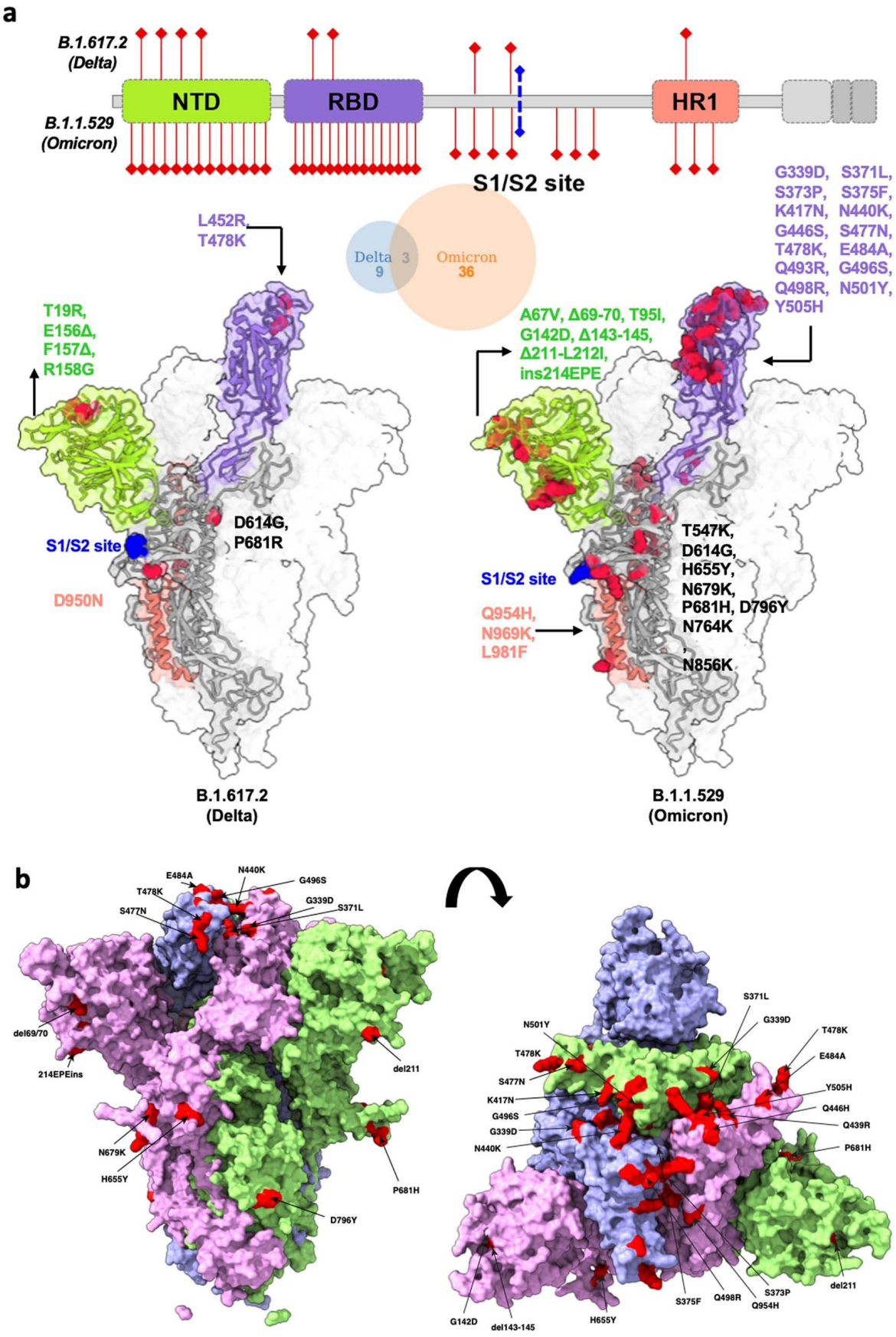


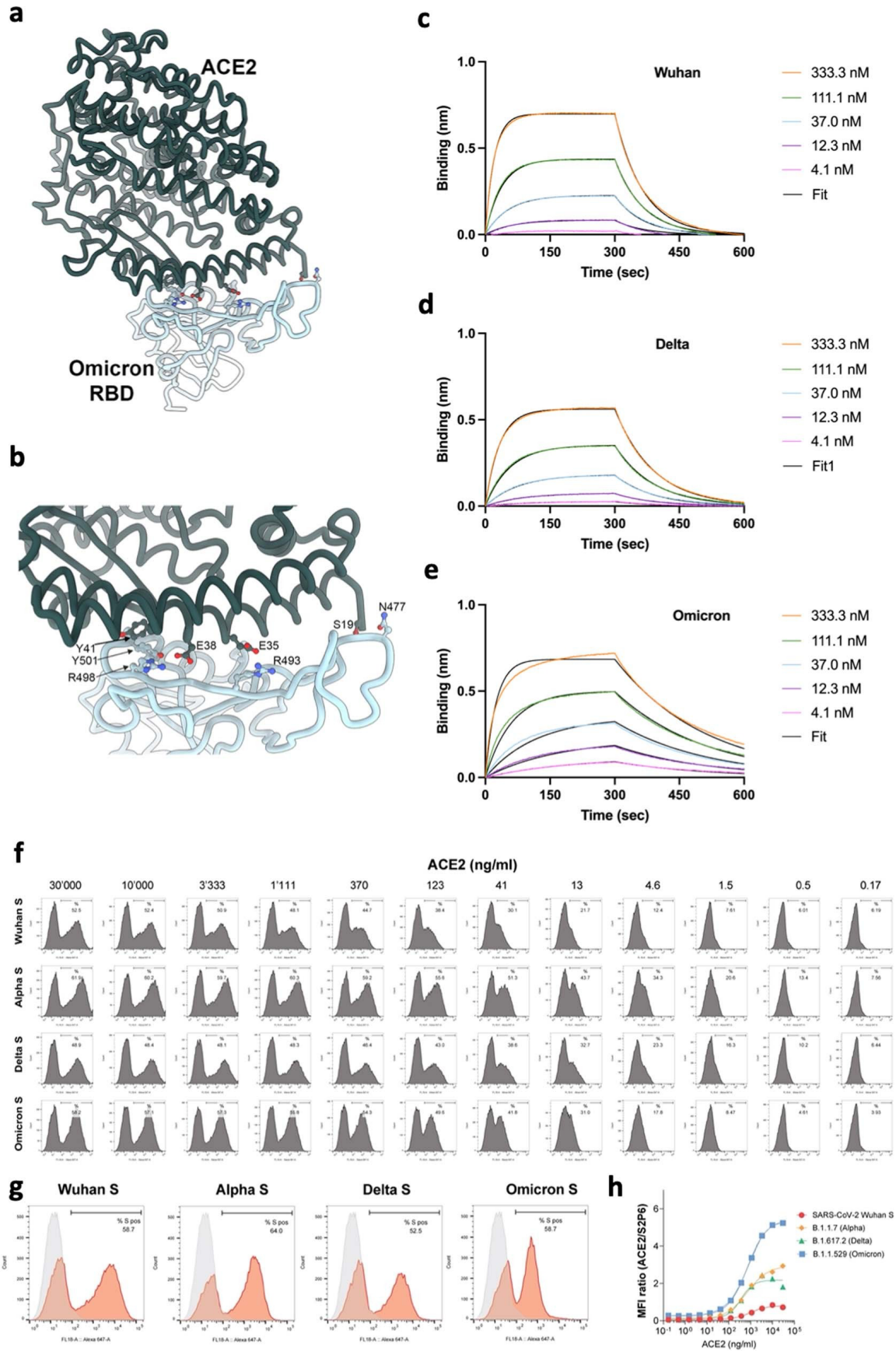
Figure 12.4: SARS-CoV-2 Omicron variant spike shows impaired cell–cell fusion activity and smaller infection foci generated by live virus.

a, Schematic of the cell–cell fusion assay. The schematic was created using BioRender.com. b, Spike expression at the cell surface as determined by flow cytometry, showing the distribution of fluorescence intensity. The percentage of spike-positive cells is indicated. c, Reconstructed images of GFP^+ syncytia at 16 h. d, Quantification of cell–cell fusion kinetics showing the percentage of the GFP^+ area to the total cell area over time. WT, Wuhan-Hu-1 D614G. Data are mean ± s.e.m. from eight fields of views at each time point. Data are representative of at least two independent experiments. e, Three representative images of wells from 96-well plates with infection foci formed by Delta (top row) or Omicron (bottom row) live virus. Scale bars, 2 mm. f, Focus area as determined by ELISpot image analysis for Omicron (n = 111 wells) and Delta (n = 112 wells). Data are the geometric mean ± 95% CI. g, Focus number per well for the same experiments as in f. Data are the geometric mean ± 95% CI for focus number per well for Omicron and Delta infections. For f and g, statistical analysis was performed using two-sided Wilcoxon rank-sum tests; ****$P < 0.0001$.


a
B.1.617.2
(Delta)
B.1.1.529
(Omicron)
NTD
RBD
HR1
S1/S2 site
Delta 3
9
Omicron
36
L452R,
T478K
T19R,
E156Δ,
F157Δ,
R158G
D614G,
P681R
S1/S2 site
D950N
B.1.617.2
(Delta)
G339D, S371L,
S373P, S375F,
K417N, N440K,
G446S, S477N,
T478K, E484A,
Q493R, G496S,
Q498R, N501Y,
Y505H
A67V, Δ69-70, T95I,
G142D, Δ143-145,
Δ211-L212I,
ins214EPE
T547K,
D614G,
H655Y,
N679K,
P681H, D796Y
N764K
,
N856K
S1/S2 site
Q954H,
N969K,
L981F
B.1.1.529
(Omicron)
b
E484A
N440K
G496S
T478K
G339D
S371L
S477N
del69/70
214EPEins
del211
N679K
H655Y
P681H
D796Y
S371L
G339D
N501Y
T478K
T478K
E484A
S477N
Y505H
K417N
Q446H
G496S
Q439R
G339D
P681H
N440K
S375F
Q498R
S373P
Q954H
del211
G142D
del143-145
H655Y

Extended Data Figure 12-1: Structural model of SARS-CoV-2 Delta and Omicron spike variants.

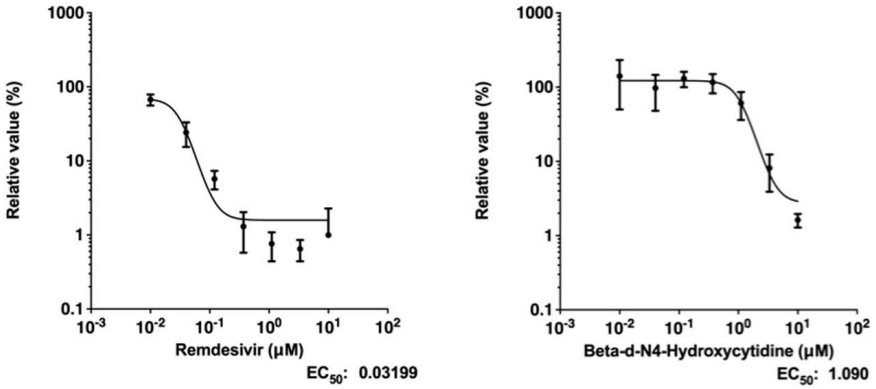
a. Variant structures for Delta and Omicron of spike protein displaying mutational sites in red colour, generated using the cryo-EM structure PDB# 7A94 as a reference. The schematic highlights differences in domain-wise mutations across protein length. **b**. Side and top down surface representation of the Omicron spike protein. Spike homotrimer structures were created predicted *in silico* by the Alphafold2 software package. Individual mutations making up the Omicron spike are highlighted in red on each of the three homotrimers.


a
ACE2
Omicron
RBD
b
Y41
Y501
R498
E38
E35
R493
S19
N477
c
Wuhan
Binding (nm)
Time (sec)
333.3 nM
111.1 nM
37.0 nM
12.3 nM
4.1 nM
Fit
d
Delta
Fit1
e
Omicron
f
ACE2 (ng/ml)
30'000
10'000
3'333
1'111
370
123
41
13
4.6
1.5
0.5
0.17
Wuhan S
Alpha S
Delta S
Omicron S
g
Wuhan S
Alpha S
Delta S
Omicron S
% S pos
58.7
64.0
52.5
58.7
h
MFI ratio (ACE2/S2P6)
ACE2 (ng/ml)
SARS-CoV-2 Wuhan S
B.1.1.7 (Alpha)
B.1.617.2 (Delta)
B.1.1.529 (Omicron)

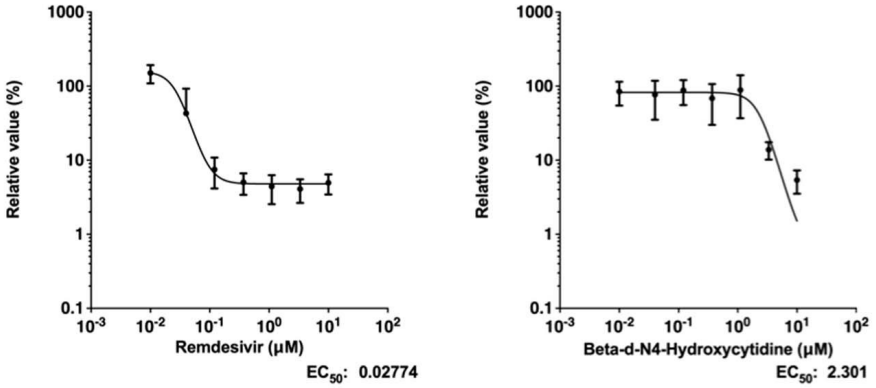
Extended Data Figure 12-2: Omicron spike binding to ACE2.

a. Rendering of the contacts formed between the Omicron RBD (light blue) and human ACE2 (dark green) using PDB 7TN0 described in15 **b**. Zoomed-in view showing select amino acid side chains participating in remodelling of interactions between the Omicron RBD and ACE2. **c**–**e**. Biolayer interferometry binding analysis of human ACE2 to immobilized SARS-CoV-2 Wuhan-Hu-1 **c**. Delta **d**. or Omicron **e**. RBDs. Black lines correspond to a global fit of the data using a 1:1 binding model. **f**. Flow cytometry analysis of ACE2 binding to SARS-CoV-2 Wuhan, B.1.1.7 (Alpha), B.1.617.2 (Delta) or Omicron (B.1.1.529) spike proteins transiently expressed in ExpiCHO cells. **g**. Expression level of each spike as measured via binding by S2-subunit targeting S2P6 mAb. Grey histograms represent background florescence of cells stained with the secondary antibody only. **h**. Normalized ACE2 binding results based on spike protein expression levels. Values on the Y axis indicate the ratio between the mean fluorescence intensity (MFI) of positive cells stained for ACE2 binding and the MFI of the S2P6 positive cells. Data points in h. are mean of technical triplicates and data shown are representative two independent experiments.
Source data

a

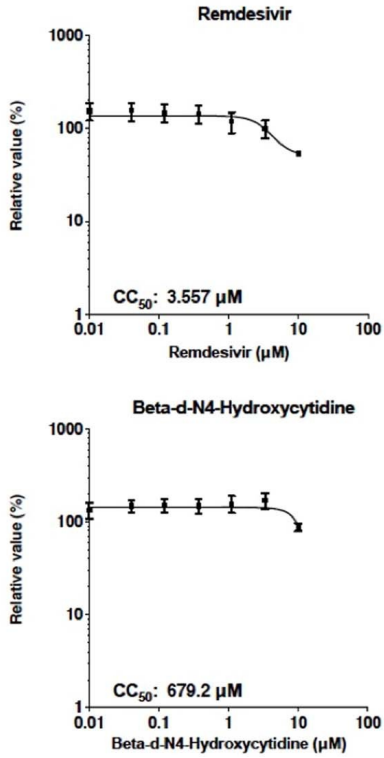


DELTA



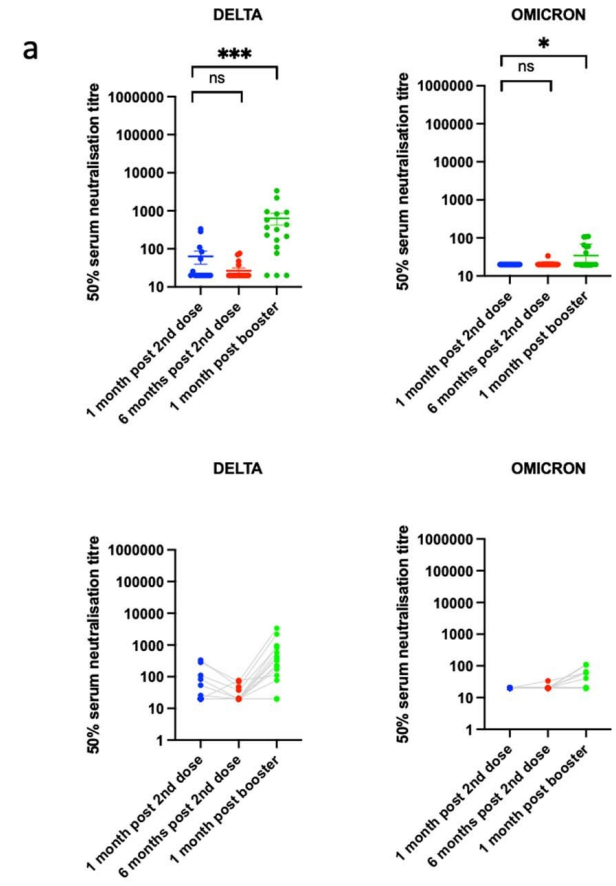
OMICRON

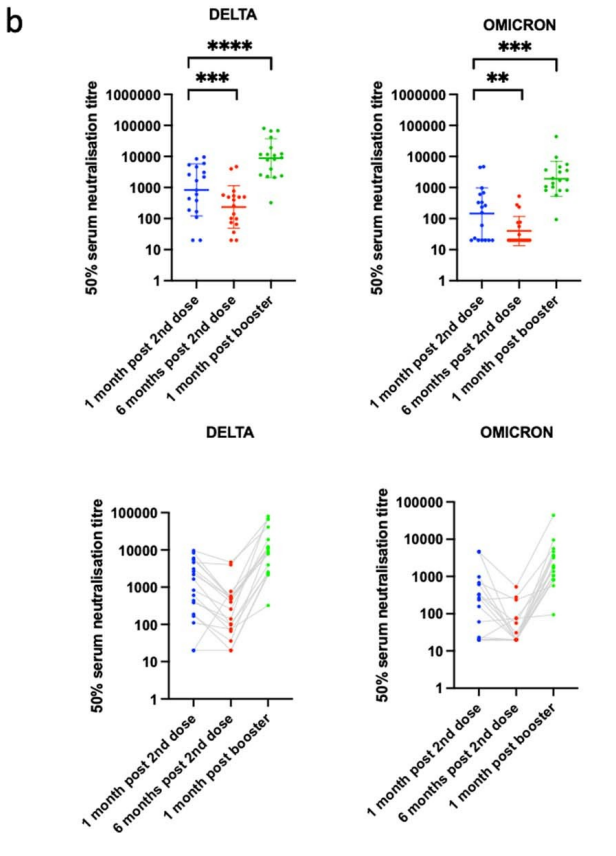
b



Extended Data Figure 12-3: Sensitivity of replication competent SARS-CoV-2 Omicron and Delta variants to clinically approved direct acting antiviral molecules remdesivir and the active metabolite of molnupiravir.

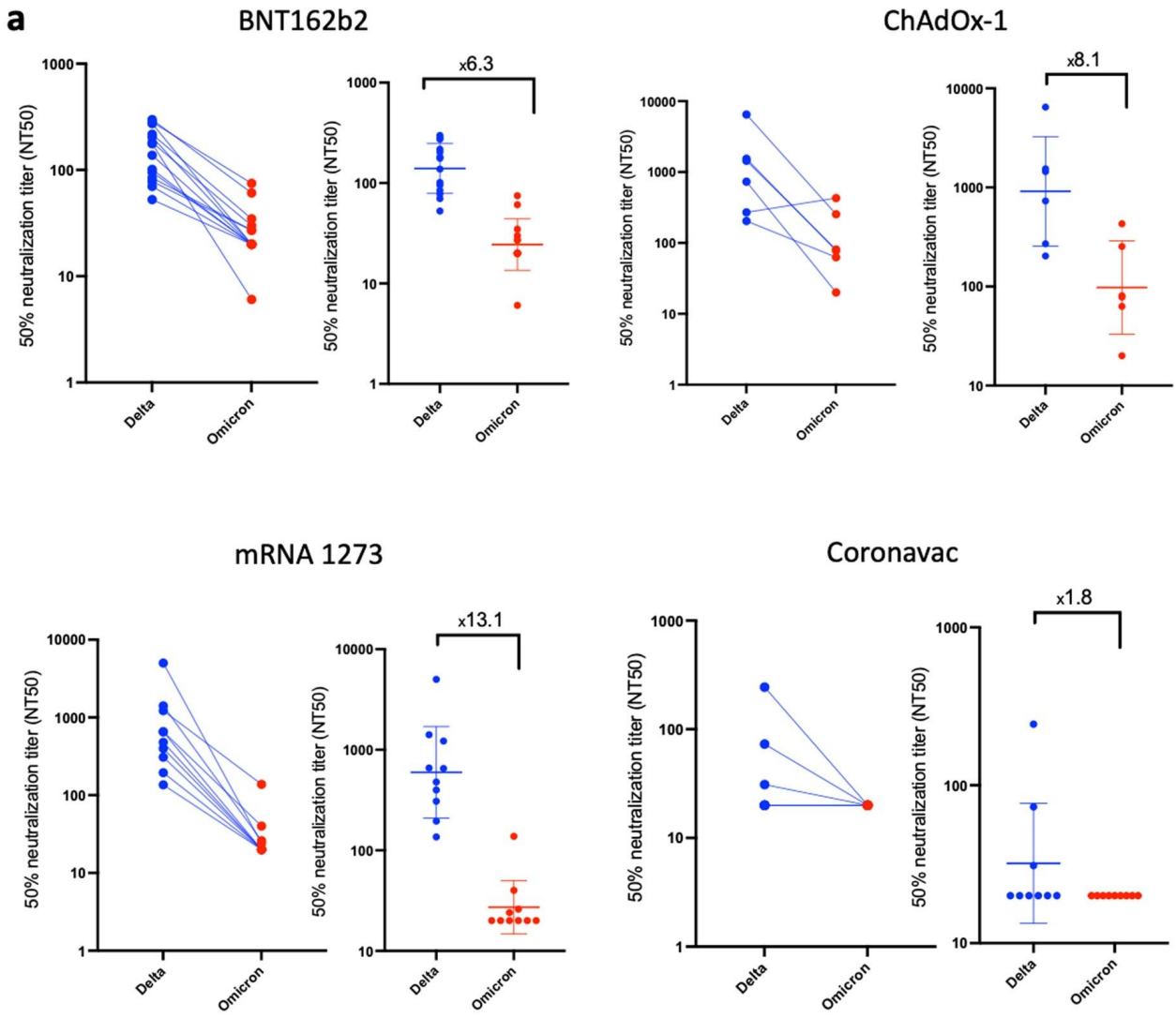
HOS cells overexpressing ACE2 and TMPRSS2 were used and a viral input of 1,000 $TCID_{50}$ was used ($TCID_{50}$ measured using VeroE6/TMPRSS2 cells). **a**. Dose-response curves. Infection as measured by viral RNA copies relative to the no drug control (100%) is plotted on the y axis with serial drug dilution on the x axis. EC50 is indicated for each panel and calculated in GraphPad Prism. Data points are mean of technical quadruplicates with +/- SEM shown. Curve fitting for dose response was done in GraphPad prism. **b**. toxicity assay showing relative cell viability at a range of drug doses. Data points are mean of technical quadruplicates with +/- SEM shown. All data are representative of two independent experiments.


a
DELTA
OMICRON
***
ns
*
ns
50% serum neutralisation titre
1000000
100000
10000
1000
100
10
1
1 month post 2nd dose
6 months post 2nd dose
1 month post booster


b
DELTA
OMICRON
****
***
***
**
50% serum neutralisation titre
1000000
100000
10000
1000
100
10
1
1 month post 2nd dose
6 months post 2nd dose
1 month post booster

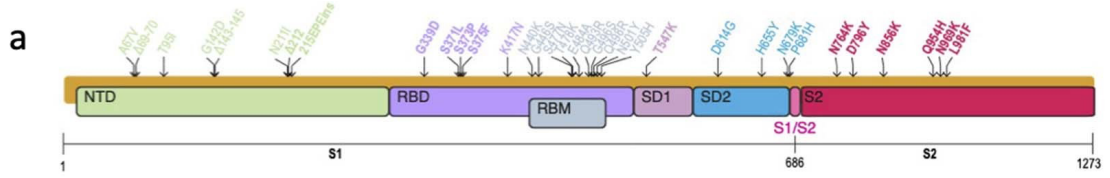
Extended Data Figure 12-4: Neutralization of spike pseudotyped virus by sera from vaccinated individuals over three time points following dose two (ChAdOx-1 or BNT162b2) and dose three (BNT162b2 only).

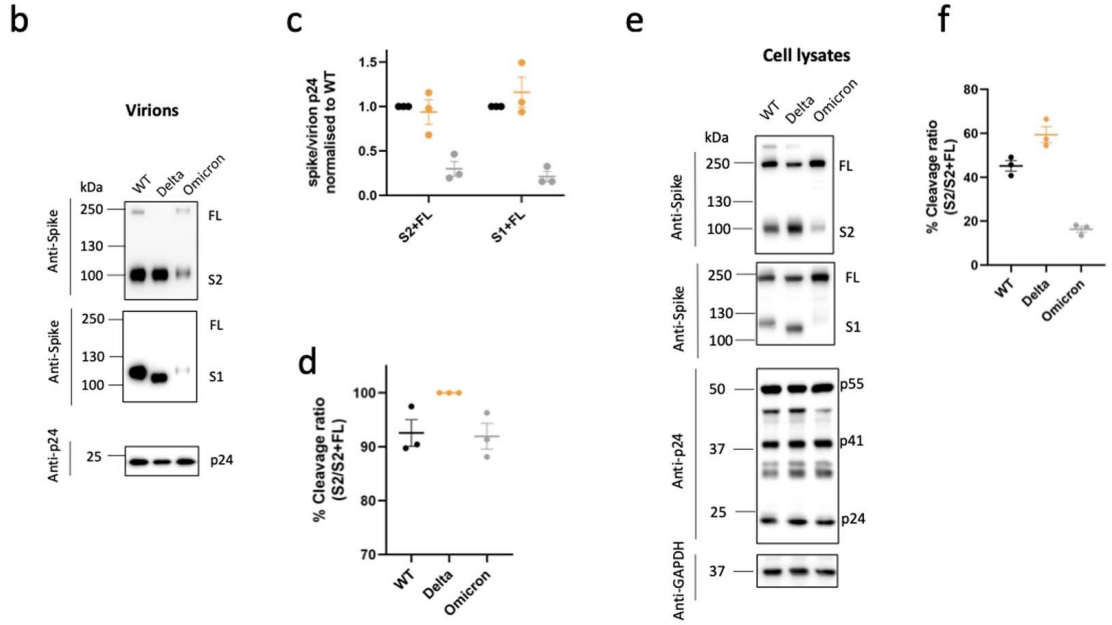
a. n = 17 ChAdOx-1 or **b**. n = 18 BNT12b2. GMT (geometric mean titre) shown as a line with s.d as bars are presented. Data representative of two independent experiments each with two technical replicates. **p < 0.01, *** p < 0.001, ****p < 0.0001 Wilcoxon matched-pairs signed rank test, ns not significant, two-sided. Data are for individuals who tested negative for anti-N IgG at each time point.

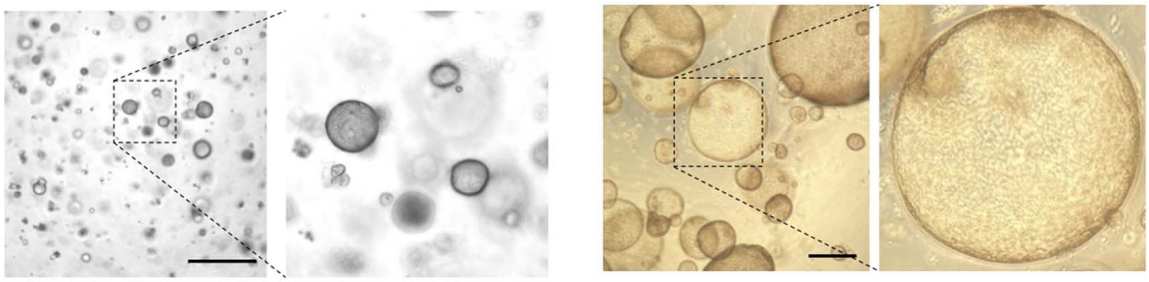

a
BNT162b2
ChAdOx-1
mRNA 1273
Coronavac
50% neutralization titer (NT50)
Delta
Omicron
x6.3
x8.1
x13.1
x1.8
b
Early pandemic 2020 convalescent Sera
Convalescent Sera from Delta variant infections
x13.6
x59.2

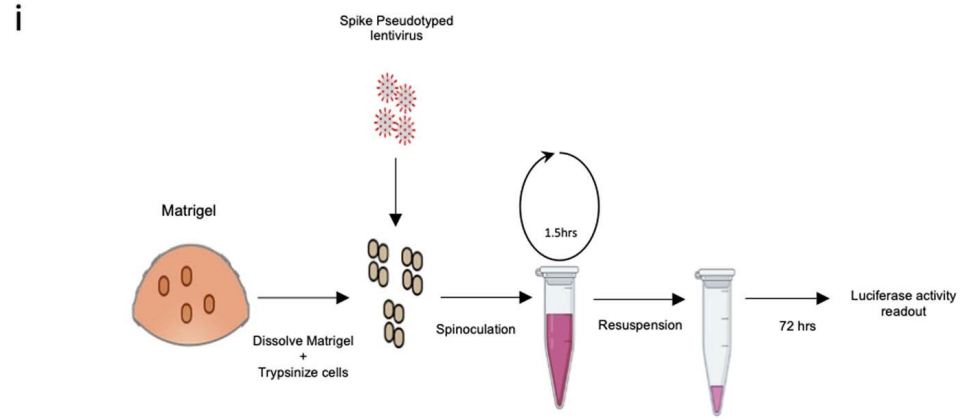
Extended Data Figure 12-5: Live virus neutralization by vaccine-elicited and convalescent sera.

a. Neutralization of replication competent Delta and Omicron viruses by sera from vaccinated individuals following dose two of ChAdOx-1 (n = 5), BNT162b2 (n = 10), mRNA 1273(n = 10), and Coronavac (n = 9) vaccines. **b**. neutralization of live viruses by sera derived from recovered individuals following infection in early 2020 or with confirmed Delta infection. GMT (geometric mean titre) of three technical replicates are shown as a line with s.d as bars are presented. Data representative of two independent experiments. Fold change in NT50 is indicated.
Source data


a
NTD
RBD
RBM
SD1
SD2
S2
S1/S2
S1
S2
1
686
1273

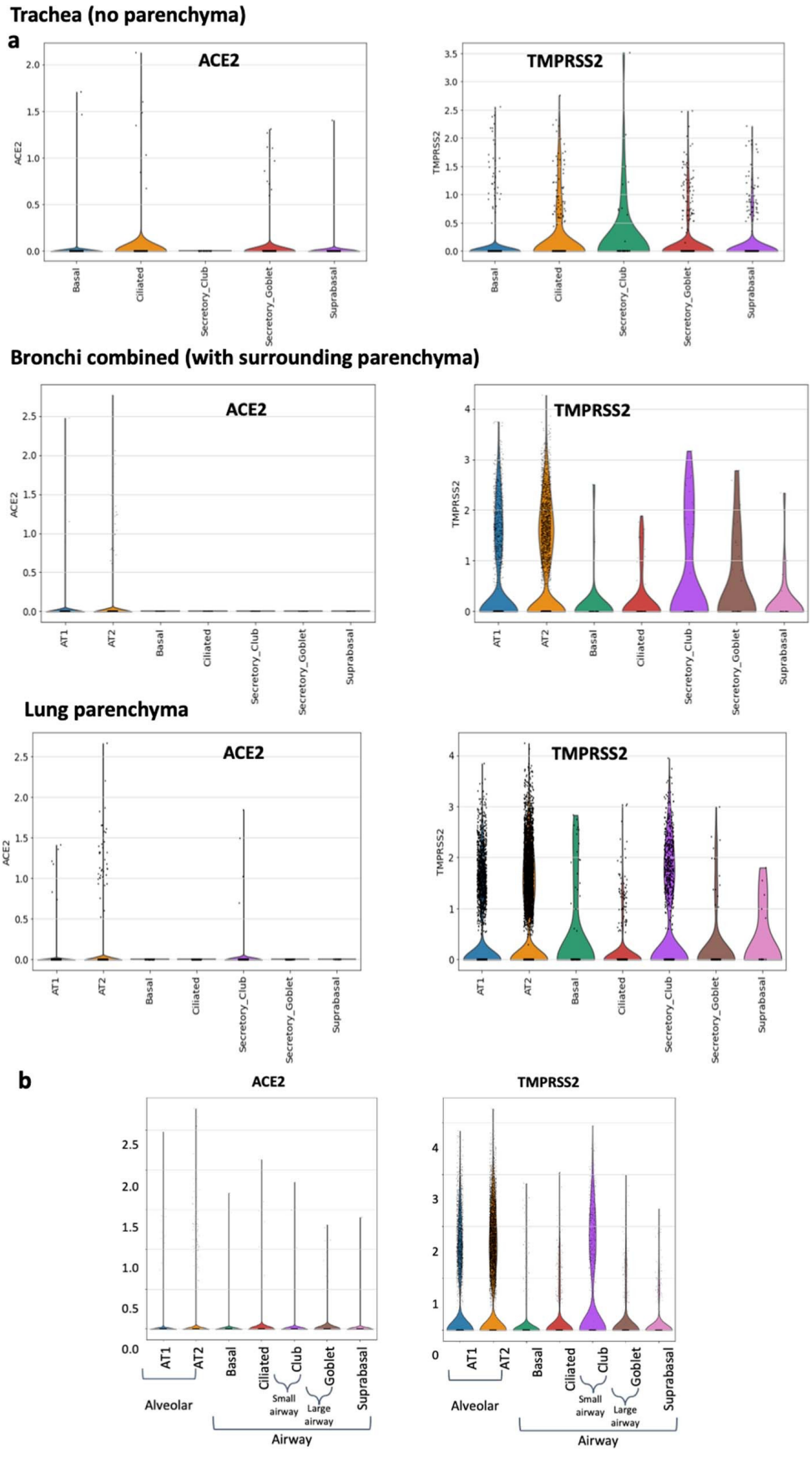

b
Virions
kDa
WT
Delta
Omicron
Anti-Spike
250
130
100
FL
S2
Anti-Spike
FL
S1
Anti-p24
25
p24
c
spike/virion p24 normalised to WT
1.5
1.0
0.5
0.0
S2+FL
S1+FL
d
% Cleavage ratio (S2/S2+FL)
100
90
80
70
WT
Delta
Omicron
e
Cell lysates
WT
Delta
Omicron
kDa
Anti-Spike
250
130
100
FL
S2
Anti-Spike
FL
S1
Anti-p24
50
37
25
p55
p41
p24
Anti-GAPDH
37
f
% Cleavage ratio (S2/S2+FL)
80
60
40
20
0
WT
Delta
Omicron


g
h


i
Spike Pseudotyped lentivirus
Matrigel
Dissolve Matrigel
+
Trypsinize cells
Spinoculation
1.5hrs
Resuspension
72 hrs
Luciferase activity readout

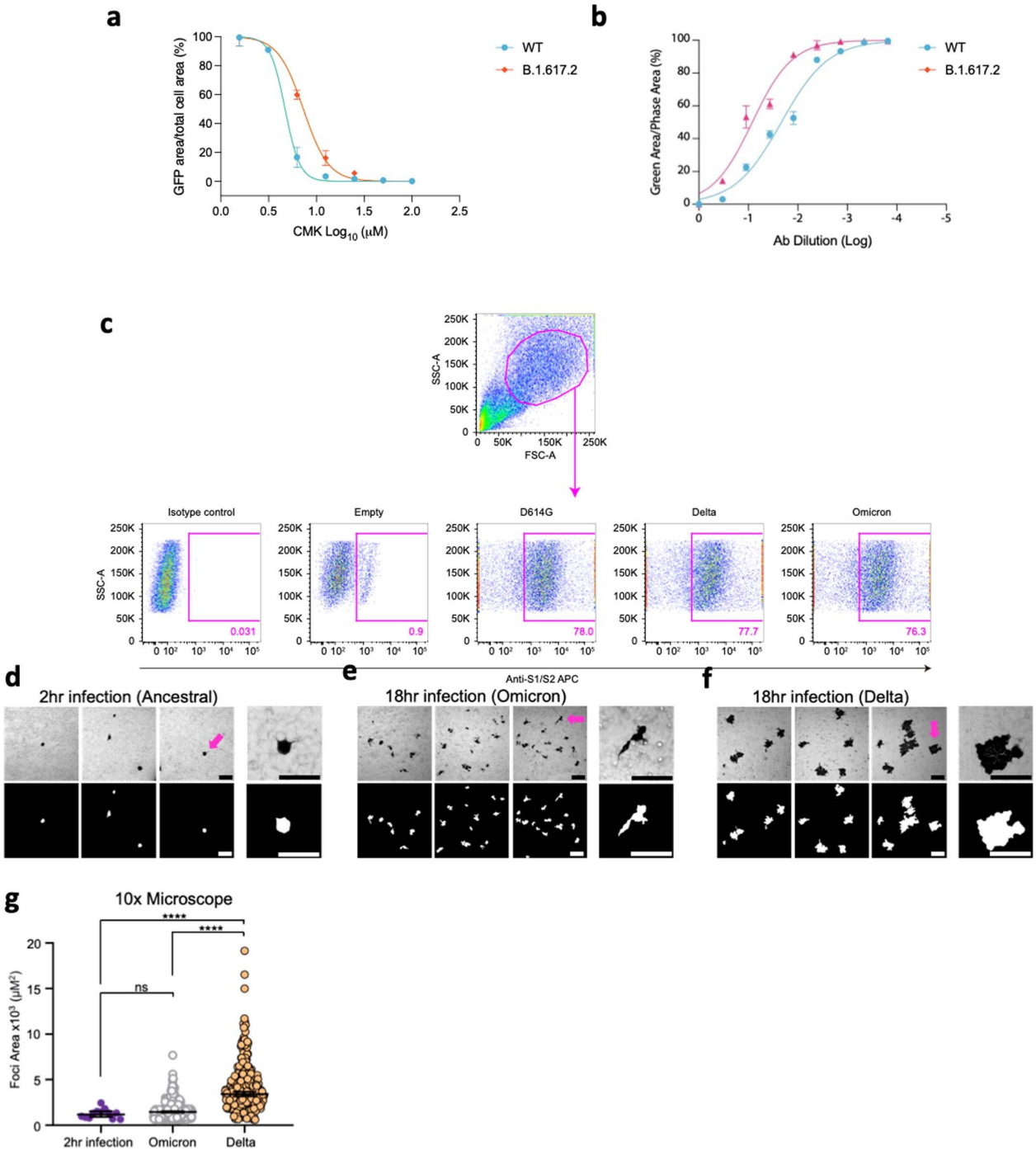
Extended Data Figure 12-6: SARS-CoV-2 Omicron Variant spike pseudotyped lentiviruses.

a. Graphical representation of Omicron spike mutations present in expression plasmid used to generate lentiviral pseudotyped virus (PV). Mutations coloured according to location in spike; bold mutations are novel to this lineage and have not been identified in previous VOCs. **b**. western blots of pseudotyped virus (PV) virions from 293T producer cells following transfection with plasmids expressing lentiviral vectors and SARS-CoV-2 S plasmids. (WT- Wuhan-1 with D614G), probed with antibodies for HIV-1 p24 and SARS-Cov-2 S2 (top) and S1 (bottom). **c**, **d**. quantification of western blots showing **c**. ratio of spike:p24 in virions, **d**. ratio of S2:total spike. Mean ratio +/- sd is shown for 3 biological replicates **e**. Western blot of cell lysates used to produce virions with **f**. quantification of S2 to total spike ratio. Mean ratio +/- sd is shown for 3 biological replicates **g**. brightfield images of lower airway organoids **h**. brightfield images of of cholangiocyte organoids (Scale bars 200 μm). **i**. infection schematic for entry assays in organoids created with BioRender.com.


Trachea (no parenchyma)
a
ACE2
TMPRSS2
Basal
Ciliated
Secretory_Club
Secretory_Goblet
Suprabasal
Bronchi combined (with surrounding parenchyma)
AT1
AT2
Lung parenchyma
b
Alveolar
Airway
Small airway
Large airway

Extended Data Figure 12-7: ACE2 and TMPRSS2 levels by single cell RNAseq.

a. Log-normalized expression of ACE2 and TMPRSS2 genes in single-nucleus RNAseq data for indicated tissue types and. **b**. Log-normalized aggregrated expression of ACE2 and TMPRSS2 genes in single-nucleus RNAseq data from human alveolar (AT1 - alveolar type 1, AT2 - alveolar type 2 pneumocytes), and airway epithelial cells (basal, suprabasal, goblet and ciliated). Data are derived from in Madissoon et al.



Extended Data Figure 12-8: SARS-CoV-2 cell-cell fusion and infection focus formation.

a. effect of furin inhibition using CMK on cell-cell fusion for WT (Wuhan-1 D614G) and Delta variant. CMK drug dilution indicated on x axis and fusion on y axis. Quantification of cell-cell fusion shows percentage of green area to total cell area. Mean of technical replicates is plotted with error bars representing SEM. Data are representative of at least two independent experiments. **b**. impact of neutralising

antibodies from UK wave one 2020 sera on cell-cell fusion for WT (Wuhan-1 D614G) and Delta variant. Serial dilutions of sera added to acceptor cells before co-cultivation with donor cells. Data points are mean of technical replicates is plotted with error bars representing SEM. **c**. Gating strategy for cell surface staining of SARS-CoV-2 spike proteins from 293T cells transfected with plasmids expressing spike, using rabbit anti-SARS-CoV-2 spike S1/2 polyclonal antibody at 1:100. **d**–**g**. Omicron shows attenuated localized spread compared to Delta. **d**. Microscope images of foci (pink arrow) 2 h post-infection with ancestral SARS-CoV-2. Bar is 200 microns. **e**. Images of Omicron foci 18 h post-infection. **f**. Images of Delta foci 18 h post-infection **g**. Geometric mean and 95% confidence intervals of focus area of microscope images of 2 h post-ancestral infection (n = 12 foci, purple), 18 h post-Omicron infection (n = 363, grey), and 18 h post-Delta infection (n = 275, orange). p-values are not significant (ns) or **** < 0.0001 as determined by two sided Wilcoxon rank sum test.

Extended Data Table 12.1: Kinetic Analysis of human ACE2 binding to SARS-CoV-2 VOC by Biolayer Interferometry.

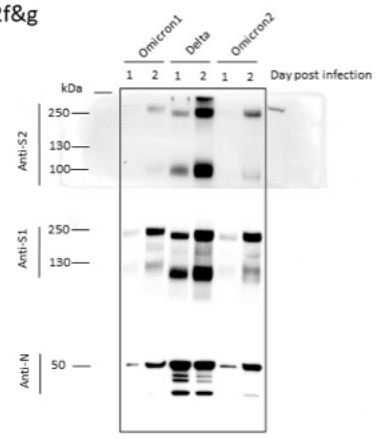
	Wuhan	Delta	Omicron
KD (nM)	127	190	44
k_{on} (M^{-1} s^{-1})	1.1×10^{5}	5.9×10^{4}	1.1×10^{5}
k_{off} (s^{-1})	1.4×10^{-2}	1.1×10^{-2}	4.7×10^{-3}

Values reported represent the global fit to the data shown in Extended Data Fig. 2.

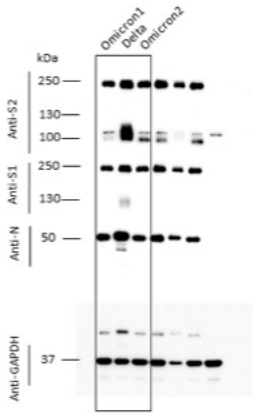
Extended Data Table 12.2: Demographic characteristics of participants in longitudinal vaccine-elicited sera neutralization study

Characteristic	Total	Pfizer (BNT162b2)	AstraZeneca (ChAdOx1)
Total number, n (%)	40	20 (50)	20 (50)
Age, median years (IQR)	70.5 (62.5, 74.0)	71.0 (55.8, 75.8)	69.5 (63.0, 73.8)
≥80, n (%)	4 (10)	4 (20)	0 (0)
<80, n (%)	36 (90)	16 (80)	20 (100)
Gender, n (%)			
Male	15 (37.5)	8 (40)	7 (35)
Female	25 (62.5)	12 (60)	13 (65)
Previous COVID-19			
Yes	5 (12.5)	2 (10)	3 (15)
No	35 (87.5)	18 (90)	17 (85)
Type of booster			
Moderna	3 (7.5)	0	3 (15)
Pfizer	37 (92.5)	20 (100)	17 (85)
Duration between booster dose and sampling, median days (IQR)	34 (32, 36)	34.0 (32.0, 36.0)	33.5 (30.3, 35.8)

SI for Figure 2f&g

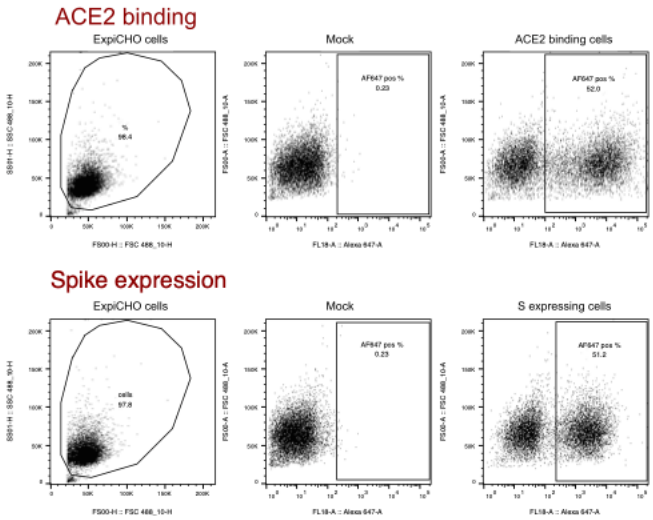


Cut the blot into two halves around 75 KDa
Probed S2 first then stripped for S1
Probed N

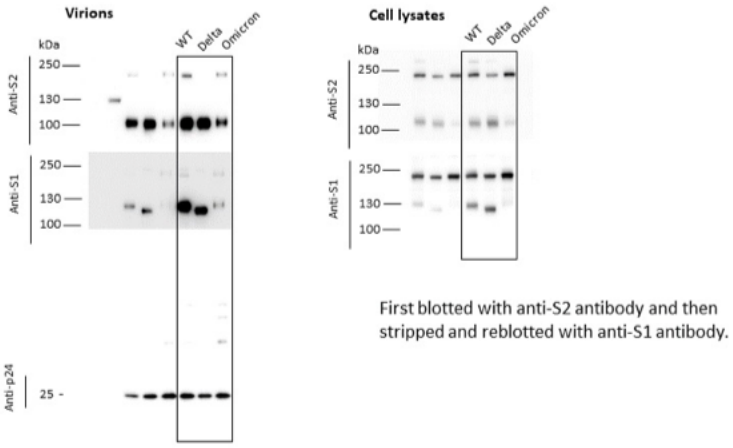


Cut the blot into two halves around 75 KDa
Probed S2 first then stripped for S1
Probed N first then GAPDH

SI for ED Fig 2: gating strategy for ACE2 binding and Spike expression analyses



SI for ED Fig 6



First blotted with anti-S2 antibody and then stripped and reblotted with anti-S1 antibody.

Cut blot into two halves around 75 kDa;
First blotted with anti-S2 antibody and then stripped and reblotted with anti-S1 antibody;
The other half blotted with anti-p24 antibody

CHAPTER 13. BROADLY NEUTRALIZING ANTIBODIES OVERCOME SARS-COV-2 OMICRON ANTIGENIC SHIFT.

In the following story I continue to discuss our efforts towards pandemic response, specifically in response to the emergence of the SARS-CoV-2 Omicron . Omicron encoded 37 amino acid substitutions in the spike protein, 15 of which are in the receptor-binding domain (RBD). In the subsequent sections I detail how the Omicron RBD binds to human ACE2 with enhanced affinity, relative to the Wuhan-Hu-1 RBD, and binds to mouse ACE2. I also discuss changes in antigenicity. Specifically, reductions in neutralizing activity were observed against Omicron compared to the ancestral pseudovirus in plasma from convalescent individuals and from individuals who had been vaccinated against SARS-CoV-2 twice. In addition, I characterize the loss of neutralizing potency for most monoclonal antibodies that are directed against the receptor-binding motif, with only 3 out of 29 monoclonal antibodies retaining unaltered potency. The magnitude of Omicron-mediated immune evasion marked a major antigenic shift in SARS-CoV-2.

Adapted from: Cameroni, E., Bowen, J.E., Rosen, L.E., Saliba, C., **Zepeda, S.K.**, Culap, K., Pinto, D., VanBlargan, L.A., De Marco, A., di Iulio ,J., Zatta, F., Kaiser, H., Noack, J., Farhat, N., Czudnochowski, N., Havenar-Daughton, C., Sprouse, K.R., Dillen, J.R., Powell, A.E., Chen, A., Maher, C., Yin, L., Sun, D., Soriaga, L., Bassi, J., Silacci-Fregni, C., Gustafsson, C., Franko, N.M., Logue, J., Iqbal, N.T., Mazzitelli, I., Geffner, J., Grifantini, R., Chu, H., Gori, A., Riva, A., Giannini, O., Ceschi, A., Ferrari, P., Cippà, P.E., Franzetti-Pellanda, A., Garzoni, C., Halfmann, P.J., Kawaoka, Y., Hebner, C., Purcell, L.A., Piccoli, L., Pizzuto, M.S., Walls, A.C., Diamond M.S.,

Telenti, A., Virgin, H.W., Lanzavecchia, A., Snell, G., Veesler, D., Corti D. (2021). "Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift." Nature. https://doi.org/10.1038/s41586-021-04386-2

13.1 Chapter Introduction

The evolution of RNA viruses can result in immune escape and modulation of binding to host receptors through the accumulation of mutations (Telenti, 2021). Previously emerged SARS-CoV-2 variants of concern (VOCs) have developed resistance to neutralizing antibodies, including some clinical antibodies that are used as therapeutics (Wang, 2021; Chen, 2021; Corti, 2021). The B.1.351 (Beta) VOC showed the greatest magnitude of immune evasion from serum neutralizing antibodies (Wang, 2021; Chen, 2021), whereas B.1.617.2 (Delta) quickly outcompeted all other circulating isolates through the acquisition of mutations that enhanced transmission and pathogenicity(McCallum, 2021, Mlcoshova, 2021; Sheikh, 2021) and eroded the neutralizing activity of antibody responses (McCallum, 2021).

The Omicron (B.1.1.529) variant was first detected in November 2021, was immediately declared to be a VOC by the World Health Organization (WHO) and quickly rose in frequency worldwide. The Omicron variant is substantially mutated compared to any previously described SARS-CoV-2 isolates, including 37 substitutions of residues in the spike protein in the predominant haplotype (Fig. 1a, Extended Data Figs. 1–4). Fifteen of the Omicron mutations are clustered in the RBD, which is the main target of neutralizing antibodies after infection or vaccination (Piccoli, 2020, Gr) (Piccoli, 2020; Greaney, 2021), suggesting that Omicron might escape infection- and vaccine-elicited

antibodies and therapeutic monoclonal antibodies. Nine of these mutations map to the receptor-binding motif (RBM), which is the RBD subdomain that directly interacts with the host receptor, ACE2 (Walls, 2020).

Preliminary reports indicate that the neutralizing activity of plasma from individuals who had received the Pfizer–BioNTech BNT162b2 vaccine is reduced against the Omicron variant (Cele, 2021; Wilhelm, 2021), documenting a substantial—albeit not complete—escape from mRNA-vaccine-elicited neutralizing antibodies. Another report also shows that vaccine effectiveness against symptomatic disease induced by the Omicron variant is significantly lower than for the Delta variant (Andrews, 2021). The potential for booster doses to ameliorate this decline in neutralization is being investigated. In addition, the neutralizing activity of several therapeutic monoclonal antibodies appears to be decreased or abolished against Omicron (Wilhelm, 2021, Cao, 2021).

To understand the consequences of the high number of mutations found in the Omicron spike protein, we used a pseudovirus assay to study receptor use and neutralization mediated by monoclonal and polyclonal antibodies, as well as surface plasmon resonance (SPR) to measure binding of the RBD to human and mouse ACE2 receptors.

13.2 Omicron RBD binds ACE2 with increased affinity

At present, 23 out of the 37 amino acid mutations in the Omicron spike protein have been individually observed previously in SARS-CoV-2 variants of interest, VOCs or other sarbecoviruses, whereas the remaining 14 substitutions have not to our knowledge been described before (Extended Data Fig. 5a). Analysis of the GISAID database indicates that there are rarely more than 10–15 Omicron spike mutations

present in a given non-Omicron haplotype or Pango lineage (Extended Data Fig. 5b–d). Although we have not formally assessed the possibility of recombination events, persistent replication in immunocompromised individuals or inter-species ping-pong transmission(Telenti, 2021) are possible scenarios for the rapid accumulation of mutations that could have been selected on the basis of viral fitness and immune evasion.

Several of the Omicron RBD mutations are found at positions that are key contact sites with human ACE2, such as K417N, Q493R and G496S (Leist, 2020). Except for N501Y, which increases ACE2-binding affinity by sixfold (Dinnon, 2020; Collier, 2021) , all other substitutions were shown by deep mutational scanning either to reduce binding or to have no effect on human ACE2 affinity when present individually (Starr, 2020), resulting in an overall predicted decrease of binding affinity (Supplementary Table 1). However, we found that the Omicron RBD has a 2.4-fold increased binding affinity to human ACE2 (Fig. 1b, c, Extended Data Fig. 6a), suggesting epistasis of the full constellation of RBD mutations. It remains to be determined whether and how the spike mutations in Omicron may influence the dynamics of RBD opening, which may also affect the engagement of the RBD with ACE2.

The presence of the N501Y mutation has previously been reported to enable some SARS-CoV-2 VOCs to infect mice (Shuai, 2021). As Omicron contains the N501Y mutation, along with 14 other RBD mutations, we investigated whether the Omicron RBD binds mouse ACE2 using SPR (Fig. 1b, Extended Data Fig. 6). The Omicron RBD binds mouse ACE2 with a 1:1 binding affinity of 470 nM (Fig. 1b), whereas weak binding of the Beta RBD and very weak binding of the Alpha RBD to mouse ACE2 was

observed (Fig. 1b, Extended Data Fig. 6b), consistent with previous reports (Shuai, 2021; Pan, 2021). Conversely, our assay did not detect any binding of the Wuhan-Hu-1, Delta or K417N RBDs to mouse ACE2. The enhanced binding of the Omicron RBD to mouse ACE2 is likely to be explained by theQ493R substitution, which is similar to the Q493K mutation isolated in mouse-adapted SARS-CoV-2 (Leist, 2020). Our binding data correlate with our observation of Omicron spike protein-mediated but not Wuhan-Hu-1/G614 spike protein-mediated entry of VSV pseudoviruses into mouseACE2-expressing cells (Fig. 1d), as recently reported (Hoffman, 2021). Collectively, these findings highlight the plasticity of the SARS-CoV-2 RBM, which in the case of the Omicron VOC acquired enhanced binding to human and mouse ACE2 orthologues, relative to other SARS-CoV-2 isolates. The influence of these findings on viral load and replication kinetics in humans and animal models remains to be evaluated, owing to the interplay of additional factors besides receptor binding. Preliminary data suggest that Omicron appears to be attenuated in some laboratory mouse strains (M.S.D., personal communication) and that it replicates less efficiently in human lung tissue as compared to Delta (Nicholls, 2021).

13.3 Extent of Omicron escape from plasma antibodies

To investigate the magnitude of immune evasion that is mediated by the 37 mutations present in the Omicron spike protein, we used Wuhan-Hu-1 and Omicron spike VSV pseudoviruses and compared plasma neutralizing activity in different cohorts of convalescent individuals (that is, individuals who had recovered from COVID-19) or individuals who had been vaccinated with six major COVID-19 vaccines (mRNA-1273, BNT162b2, AZD1222, Ad26.COV2.S, Sputnik V and BBIBP-CorV) (Fig. 2, Supplementary Figs. 1–3, Extended Data Table 1).

Convalescent individuals and individuals who had been vaccinated with Ad26.COV2.S (single dose), Sputnik V or BBIBP-CorV had no detectable neutralizing activity against Omicron, with only one convalescent subject and three BBIBP-CorV vaccinees exhibiting ID_{50} values above 10 (Fig. 2a). Individuals who were immunized with mRNA-1273, BNT162b2 and AZD1222 had more potent neutralizing activity against Wuhan-Hu-1 and retained detectable neutralization against Omicron, with decreases of 39-, 37- and 21-fold, respectively (Fig. 2a). The dampening of neutralizing activity against Omicron was comparable to that observed against SARS-CoV, a virus that differs from Wuhan-Hu-1 by 52 residues in the RBD. Reductions of neutralization potency were less pronounced in vaccinated individuals who had been previously infected (fivefold; Fig. 2b) and in individuals undergoing dialysis (fourfold; Fig. 2c) who were boosted with a third mRNA vaccine dose. In the same cohort of individuals undergoing dialysis, the levels of antibodies that neutralize the vaccine-matched Wuhan-Hu-1 strain were found to be low (less than 1/100) or undetectable in 44% of individuals after the second dose of mRNA vaccine (Bassi, 2021).

Collectively, these findings provide evidence of a substantial reduction in plasma neutralizing activity against Omicron as compared to the ancestral virus, with neutralizing activity probably falling below the protective threshold in several cases28. Our data further indicate that multiple exposures to the ancestral virus through infection or vaccination result in the production of antibodies that can neutralize divergent viruses, such as Omicron or even SARS-CoV, as a consequence of affinity maturation or epitope masking by immune-dominant RBM antibodies (Khoury, 2021; Stamatotos, 2021; Bergstrom, 2017).

13.4 Broadly neutralizing antibodies inhibit Omicron

Neutralizing monoclonal antibodies with demonstrated in vivo efficacy in the prevention or treatment of COVID-19 (Baum, 2021; Hansen, 2020; Shi, 2020; Gottlieb, 2021; Jones, 2021; Dong, 2021; Zost, 2020) can be divided into two groups on the basis of whether they do or do not block the binding of the spike protein to ACE2. Of the eight currently authorized or approved monoclonal antibodies, seven (LY-CoV555, LY-CoV016, REGN10987, REGN10933, COV2-2130, COV2-2196 and CT-P59; all synthesized on the basis of publicly available sequences) block the binding of spike protein to ACE2 and are often used as two-antibody cocktails (Corti, 2021). They bind to epitopes that overlap with the RBM (Fig. 3a), which is structurally and evolutionary plastic (Thomson, 2021), as shown by the accumulation of mutations throughout the pandemic and the genetic diversity of this subdomain among ACE2-using sarbecoviruses (Starr, 2021). Combining two such ACE2-blocking monoclonal antibodies can provide greater resistance to variant viruses that carry RBM mutations (Baum, 2020). The second class of monoclonal antibodies, represented by sotrovimab, do not block ACE2 binding but neutralize SARS-CoV-2 by targeting non-RBM epitopes that are shared across many sarbecoviruses, including SARS-CoV(Starr, 2021; Lempp, 2021).

We compared the in vitro neutralizing activity of these therapeutic monoclonal antibodies side-by-side against Wuhan-Hu-1 and Omicron spike proteins using VSV pseudoviruses (Fig. 3). Although sotrovimab had a threefold-reduced potency against Omicron and Omicron-R346K variant VSV pseudoviruses, all of the other (RBM-specific) monoclonal antibodies completely lost their neutralizing activity—with the exception of the combination of COV2-2130 and COV2-2196, for which we determined

an approximately 100-fold-reduced potency (Fig. 3b, c). Moreover, sotrovimab exhibited a reduction of less than twofold in neutralizing activity against authentic Omicron SARS-CoV-2 as compared to the WA1/2020 D614G virus (Fig. 3c, Extended Data Fig. 7), consistent with recent reports on S309, the parent of sotrovimab (VanBlargan, 2022; Aggarwal, 2021). The threefold and less-than-twofold decrease in the neutralizing activity of sotrovimab against pseudoviruses and authentic virus, respectively, is within the currently defined threshold of 'no change' as defined by the US Food and Drug Administration (FDA; the FDA fact sheet for sotrovimab denotes no change as a reduction of less than fivefold in susceptibility(FDA, 2021)). Overall, our findings agree with two preliminary reports (Wilhelm, 2021; Cao, 2021) and, together with serological data, support the conclusion that the Omicron VOC has undergone antigenic shift.

We next tested a larger panel of 36 neutralizing NTD- or RBD-specific monoclonal antibodies for which the epitopes have been characterized structurally or assigned to a given antigenic site through competition studies(Greaney, 2021; Walls, 2020; Leist, 2020; Starr, 2020;Tortorici, 2020; McCallum, 2021) (**Fig. 13.4a, Extended Data Table 13.2**, **Extended Data Fig. 13.8**). The four NTD-specific antibodies completely lost activity against Omicron, consistent with the presence of mutations and deletions in the NTD antigenic supersite (Collier, 2021; McCallum, 2021; McCallum 2021). Three out of the twenty-two monoclonal antibodies that target the RBD antigenic site I (RBM) retained potent neutralizing activity against Omicron, including S2K146, which binds the RBD of SARS-CoV-2, SARS-CoV and other sarbecoviruses through ACE2 molecular mimicry (Park, 2022). Of the nine monoclonal antibodies that are specific for the conserved RBD site II (refs. Starr, 2021; Piccoli, 2020), only S2X259 (Totorici,

2021) retained activity against Omicron, whereas neutralization was decreased by more than tenfold or abolished for the remaining antibodies. Finally, the S2H97 monoclonal antibody retained neutralizing activity against Omicron through recognition of the highly conserved cryptic site V (ref. Starr, 2021). The panel of 44 monoclonal antibodies tested in this study includes members of each of the four classes of neutralizing monoclonal antibodies, defined by their cognate RBD-binding sites (sites I, II, IV and V) (Piccoli, 2020). Our findings show that a member, or members, of each of the four classes can retain Omicron neutralization: S2K146, S2X324 and S2N28 targeting site I, S2X259 targeting site II, sotrovimab targeting site IV and S2H97 targeting site V (**Fig. 13.4b**). Several of these monoclonal antibodies cross-react with and neutralize sarbecoviruses beyond the SARS-CoV-2 clade 1b (Park, 2022; Greaney, Walls, 2020), indicating that targeting of conserved epitopes can lead to neutralization breadth and resilience to antigenic shift associated with viral evolution.

13.5 Discussion

The high number of substitutions present in the Omicron spike protein marks a pronounced shift in antigenicity and is associated with immune evasion of considerable magnitude for SARS-CoV-2. Antigenic shift of the influenza virus is defined as genetic reassortment of the RNA genome segments, but the mechanism for the abrupt appearance of a large number of mutations in SARS-CoV-2 Omicron spike protein remains to be determined. Although recombination events are a hallmark of coronaviruses (Fischer, 2021), we and others (Kupferschmidt, 2021) propose that the Omicron shift may result from extensive viral replication in immunodeficient hosts (Fischer, 2021; Corey, 2021), although we cannot rule out the possibility of a

contribution of inter-species ping-pong transmission (Telenti, 2021) between humans and rodents, as previously described for minks (Oude, 2021).

Consistent with the variable decrease in plasma neutralizing antibody titres, we found that only 6 out of a panel of 44 neutralizing monoclonal antibodies retained potent neutralizing activity against Omicron. The monoclonal antibodies that retain neutralization recognize RBD antigenic sites that are conserved in Omicron and other sarbecoviruses. Notably, three of these antibodies bind to the RBM, including one that is a molecular mimic of the ACE2 receptor (S2K146) (Park, 2022). Collectively, these data may guide future efforts to develop SARS-CoV-2 vaccines and therapies to counteract antigenic shift and future sarbecovirus zoonotic spillovers.

13.6 Methods

Cell lines

Cell lines used in this study were obtained from ATCC (HEK293T and Vero E6), Thermo Fisher Scientific (Expi-CHO-S cells, FreeStyle 293-F cells and Expi293F cells), Takara (Lenti-X 293T cells) or generated in-house (Vero E6-TMPRSS2) (Lempp, 2021). Vero-TMPRSS2 (Zang, 2020)cells were cultured at 37 °C in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 10 mM HEPES pH 7.3 and 100 U ml^{-1} of penicillin–streptomycin and supplemented with 5 µg ml^{-1} of blasticidin. None of the cell lines used was authenticated. Cell lines were routinely tested for mycoplasma contamination.

Omicron prevalence analysis

The viral sequences and the corresponding metadata were obtained from the GISAID EpiCoV project (https://www.gisaid.org/). Analysis was performed on sequences

submitted to GISAID up to 20 December 2021. Spike protein sequences were obtained from the genomic sequences with the exonerate52 2 2.4.0–haf93ef1_3 (https://quay.io/repository/biocontainers/exonerate?tab=tags) using protein to DNA alignment with parameters -m protein2dna –refine full –minintron 999999 –percent 20 and using accession YP_009724390.1 as a reference. Multiple sequence alignment of all human spike proteins was performed with mafft53 7.475–h516909a_0 (https://quay.io/repository/biocontainers/mafft?tab=tags) with parameters –auto –reorder –mapout –keeplength –addfragments using the same reference as above. The –mapout parameter was used to retrieve insertions. Spike sequences that contained more than 10% ambiguous amino acids or that were less than 80% of the canonical protein length were discarded. Figures were generated with R 4.0.2 (https://cran.r-project.org/) using the ggplot2 3.3.2 and sf 0.9-7 packages. To identify each mutation prevalence, missingness (or ambiguous amino acids) was taken into account in both nominator and denominator.

Monoclonal antibodies

Sotrovimab and VIR-7832 (VIR-7832 (Yamin, 2021) is derived from sotrovimab, Fc further engineered to carry GAALIE) were produced at WuXi Biologics. Antibody VH and VL sequences for the monoclonal antibodies COV2-2130 (Protein Data Bank (PDB) ID 7L7E), COV2-2196 (PDB ID 7L7E, 7L7D), REGN10933 (PDB ID 6XDG), REGN10987 (PDB ID 6XDG) and ADI-58125 (PCT application WO2021207597, seq. IDs 22301 and 22311) were subcloned into heavy chain (human IgG1) and the corresponding light chain (human Ig κ-chain, Ig λ-chain) expression vectors respectively and produced in transiently transfected ExpiCHO-S cells (Thermo Fisher Scientific,

A29133) at 37 °C and 8% CO_2. Cells were transfected using ExpiFectamine. Transfected cells were supplemented 1 day after transfection with ExpiCHO Feed and ExpiFectamine CHO Enhancer. Cell culture supernatant was collected eight days after transfection and filtered through a 0.2-μm filter. Recombinant antibodies were affinity purified on an ÄKTA Xpress FPLC device using 5 ml HiTrap MabSelect PrismA columns followed by buffer exchange to histidine buffer (20 mM histidine, 8% sucrose, pH 6) using HiPrep 26/10 desalting columns. Antibody VH and VL sequences for LY-CoV555, LY-CoV016 and CT-P59 were obtained from PDB IDs 7KMG, 7C01 and 7CM4, respectively, and monoclonal antibodies were produced as recombinant IgG1 by ATUM. The remaining monoclonal antibodies were discovered at VIR and have been produced as recombinant IgG1 in ExpiCHO-S cells as described above. The identity of the produced monoclonal antibodies was confirmed by liquid chromatography–mass spectrometry (LC–MS) analysis.

IgG mass quantification by LC–MS intact protein mass analysis

Fc N-linked glycan from monoclonal antibodies was removed by PNGase F after overnight non-denaturing reaction at room temperature. Deglycosylated protein (4 μg) was injected to the LC–MS system to acquire intact MS signal. Thermo MS (Q Exactive Plus Orbitrap) was used to acquire intact protein mass under denaturing condition with *m*/*z*window from 1,000 to 6,000. BioPharma Finder 3.2 software was used to deconvolute the raw *m*/*z* data to protein average mass. The theoretical mass for each monoclonal antibody was calculated with GPMAW 10.10 software. Post-translational modifications such as N-terminal pyroglutamate cyclization, C-terminal lysine cleavage and formation of 16–18 disulfide bonds were added into the calculation.

Sample donors

Samples were obtained from SARS-CoV-2 convalescent and vaccinated individuals under study protocols approved by the local institutional review boards (Canton Ticino Ethics Committee, Switzerland, Comitato Etico Milano Area 1). All donors provided written informed consent for the use of blood and blood derivatives (such as peripheral blood mononuclear cells, sera or plasma) for research. Plasma samples from convalescent individuals and individuals who had been vaccinated with Ad26.COV2.S, mRNA-1273 or BNT162b2 were obtained from the HAARVI study, approved by the University of Washington Human Subjects Division Institutional Review Board (STUDY00000959). Samples from individuals who had been vaccinated with AZD1222 were obtained from INGM, Ospedale Maggio Policlinico of Milan and approved by the local review board Study Polimmune. Samples from individuals who had been vaccinated with Sputnik V were obtained from healthcare workers at the hospital de Clínicas 'José de San Martín', Buenos Aires, Argentina. Samples from individuals who had been vaccinated with Sinopharm were enrolled from Aga Khan University under the institutional review board of the UWARN (United World Antivirus Research Network) study.

Pseudovirus neutralization assays

Generation of VSV pseudovirus used in Vero E6 cells

The plasmid encoding the Omicron SARS-CoV-2 spike variant was generated by overlap PCR mutagenesis of the wild-type plasmid, pcDNA3.1(+)-spike-D19 (Giroglou, 2004). Replication defective VSV pseudovirus expressing SARS-CoV-2 spike proteins

corresponding to the ancestral Wuhan-Hu-1 virus and the Omicron VOC were generated as previously described (McCallum, 2021) with some modifications. Lenti-X 293T cells (Takara) were seeded in 15-cm^2 dishes at a density of 10×10^6cells per dish and the following day were transfected with 25 µg of spike expression plasmid with TransIT-Lenti (Mirus, 6600) according to the manufacturer's instructions. One day after transfection, cells were infected with VSV-luc (VSV-G) with a multiplicity of infection (MOI) of 3 for 1 h, rinsed three times with PBS containing Ca^{2+} and Mg^{2+}, then incubated for an additional 24 h in complete medium at 37 °C. The cell supernatant was clarified by centrifugation, aliquoted, and frozen at −80 °C.

Generation of VSV pseudovirus used in Vero E6-TMPRSS2 cells

Comparison of Omicron SARS-CoV-2 spike VSV to SARS-CoV-2 G614 spike (YP 009724390.1) VSV and Beta spike VSV used pseudotyped particles prepared as described previously (McCallum, 2021; Walls, 2021). In brief, HEK293T cells in DMEM supplemented with 10% FBS, 1% penicillin–streptomycin seeded in 10-cm dishes were transfected with the plasmid encoding the corresponding spike glycoprotein using Lipofectamine 2000 (Life Technologies) following the manufacturer's instructions. One day after transfection, cells were infected with VSV(G*ΔG-luciferase) (Kaname, 2010) and after 2 h were washed five times with DMEM before adding medium supplemented with anti-VSV-G antibody (I1-mouse hybridoma supernatant, CRL-2700, ATCC). Virus pseudotypes were collected 18–24 h after inoculation, clarified by centrifugation at 2,500*g* for 5 min, filtered through a 0.45-µm cut-off membrane, concentrated 10 times with a 30-kDa cut-off membrane, aliquoted and stored at −80 °C.

VSV pseudovirus neutralization

Assay performed using Vero E6 cells

Vero E6 cells were grown in DMEM supplemented with 10% FBS and seeded into clear bottom white 96 well plates (PerkinElmer, 6005688) at a density of 20,000 cells per well. The next day, monoclonal antibodies or plasma were serially diluted in pre-warmed complete medium, mixed with pseudoviruses and incubated for 1 h at 37 °C in round bottom polypropylene plates. Medium from cells was aspirated and 50 µl of virus–monoclonal antibody–plasma complexes was added to cells, which were then incubated for 1 h at 37 °C. An additional 100 µl of prewarmed complete medium was then added on top of complexes and cells were incubated for an additional 16–24 h. Conditions were tested in duplicate wells on each plate and eight wells per plate contained untreated infected cells (defining the 0% of neutralization, 'MAX RLU' value) and infected cells in the presence of S309 and S2X259 at 20 µg ml^{-1} each (defining the 100% of neutralization, 'MIN RLU' value). Virus–monoclonal antibody–plasma-containing medium was then aspirated from cells and 100 µl of a 1:2 dilution of SteadyLite Plus (PerkinElmer, 6066759) in PBS with Ca^{2+} and Mg^{2+} was added to cells. Plates were incubated for 15 min at room temperature and then were analysed on the Synergy-H1 (Biotek). The average relative light units (RLUs) of untreated infected wells (MAX RLU_{ave}) was subtracted by the average of MIN RLU (MIN RLU_{ave}) and used to normalize percentage of neutralization of individual RLU values of experimental data according to the following formula: $(1 - (RLU_x - \text{MIN } RLU_{ave})/(\text{MAX } RLU_{ave} - \text{MIN } RLU_{ave})) \times 100$. Data were analysed and visualized with Prism (v.9.1.0). IC_{50} (monoclonal antibodies) and ID_{50}(plasma) values were calculated from the

interpolated value from the log(inhibitor) versus response, using variable slope (four parameters) nonlinear regression with an upper constraint of ≤100, and a lower constrain equal to 0. Each neutralization experiment was conducted on two independent experiments—that is, biological replicates—in which each biological replicate contains a technical duplicate. IC_{50} values across biological replicates are presented as arithmetic mean ± s.d. The loss or gain of neutralization potency across spike variants was calculated by dividing the variant IC_{50}/ID_{50} by the parental IC_{50}/ID_{50} within each biological replicate, and then visualized as arithmetic mean ± s.d.

Assay performed using Vero E6-TMPRSS2 cells

VeroE6-TMPRSS2 cells were cultured in DMEM with 10% FBS (Hyclone), 1% penicillin–streptomycin and 8 μg ml^{-1}puromycin (to ensure retention of TMPRSS2) with 5% CO_2 in a 37 °C incubator (Thermo Fisher Scientific). Cells were trypsinized using 0.05% trypsin and plated to be at 90% confluence the following day. In an empty half-area 96-well plate, a 1:3 serial dilution of serum was made in DMEM and diluted pseudovirus was then added and incubated at room temperature for 30–60 min before addition of the serum–virus mixture to the cells at 37 °C. Two hours later, 40 μl of a DMEM solution containing 20% FBS and 2% penicillin–streptomycin (Thermo Fisher Scientific, 10,000 units per ml of penicillin and 10,000 μg ml^{-1} of streptomycin when undiluted) was added to each well. After 17–20 h, 40 μl per well of One-Glo-EX substrate (Promega) was added to the cells and they were incubated in the dark for 5–10 min before reading on a BioTek plate reader. Measurements were done at least in duplicate using distinct batches of pseudoviruses and one representative experiment is

shown. RLUs were plotted and normalized in Prism (GraphPad). Nonlinear regression of log(inhibitor) versus normalized response was used to determine IC_{50} values from curve fits. Normality was tested using the D'Agostino-Pearson test and in the absence of a normal distribution, Kruskal–Wallis tests were used to compare two groups to determine whether differences reached statistical significance. Fold changes were determined by comparing individual IC_{50} and then averaging the individual fold changes for reporting.

Focus reduction neutralization test

The WA1/2020 strain with a D614G substitution was described previously (Plante, 2021). The B.1.1.529 isolate (hCoV-19/USA/WI-WSLH-221686/2021) was obtained from a nasal swab and passaged on Vero-TMPRSS2 cells as described (Imai, 2020). The B.1.1.529 isolate was sequenced (GISAID: EPI_ISL_7263803) to confirm the stability of substitutions. All virus experiments were performed in an approved biosafety level 3 (BSL-3) facility.

Serial dilutions of sotrovimab were incubated with 10^2 focus-forming units of SARS-CoV-2 (WA1/2020 D614G or B.1.1.529) for 1 h at 37 °C. Antibody–virus complexes were added to Vero-TMPRSS2 cell monolayers in 96-well plates and incubated at 37 °C for 1 h. Subsequently, cells were overlaid with 1% (w/v) methylcellulose in MEM. Plates were collected 30 h later (WA1/2020 D614G on Vero-TMPRSS2 cells) or 70 h later (B.1.1.529 on Vero-TMPRSS2 cells) by removal of overlays and fixation with 4% PFA in PBS for 20 min at room temperature. Plates with WA1/2020 D614G were washed and sequentially incubated with an oligoclonal pool of SARS2-2, SARS2-11, SARS2-16, SARS2-31, SARS2-38, SARS2-57 and SARS2-71

(Liu, 2021) anti-S antibodies. Plates with B.1.1.529 were additionally incubated with a pool of monoclonal antibodies that cross-react with SARS-CoV-1 and bind a CR3022-competing epitope on the RBD (VanBlargan, 2021). All plates were subsequently stained with HRP-conjugated goat anti-mouse IgG (Sigma, A8924) in PBS supplemented with 0.1% saponin and 0.1% bovine serum albumin. SARS-CoV-2-infected cell foci were visualized using TrueBlue peroxidase substrate (KPL) and quantitated on an ImmunoSpot microanalyser (Cellular Technologies). Antibody dose response curves were analysed using nonlinear regression analysis with a variable slope (GraphPad Software), and the IC_{50} was calculated.

VSV pseudovirus entry assays using mouse ACE2

HEK293T (293T) cells (ATCC CRL-11268) were cultured in 10% FBS, 1% penicillin–streptomycin in DMEM at 37 °C in a humidified 8% CO_2 incubator. Transient transfection of mouse ACE2 in 293T cells was done 18–24 h before infection using Lipofectamine 2000 (Life Technologies) and an HDM plasmid containing full-length mouse ACE2 (GenBank: Q8R010, synthesized by GenScript) in Opti-MEM. After a 5-h incubation at 37 °C in a humidified 8% CO_2 incubator, DMEM with 10% FBS was added and cells were incubated at 37 °C in a humidified 8% CO_2 incubator for 18–24 h. Immediately before infection, 293T cells with transient expression of mouse ACE2 were washed with DMEM 1×, then plated with pseudovirus at a 1:75 dilution in DMEM. Infection in DMEM was done with cells between 60% and 80% confluence for 2.5 h before adding FBS and penicillin–streptomycin to final concentrations of 10% and 1%, respectively. After 18–24 h of infection, One-Glo-EX (Promega) was added to the cells and they were incubated in the dark for 5 min before reading on a Synergy H1 Hybrid

Multi-Mode plate reader (Biotek). Cell entry levels of pseudovirus generated on different days (biological replicates) were plotted in GraphPad Prism as individual points, and average cell entry across biological replicates was calculated as the geometric mean.

Production of recombinant RBD proteins

SARS-CoV-2 RBD proteins for SPR binding assays (residues 328–531 of the spike protein from GenBank NC_045512.2 with N-terminal signal peptide and C-terminal thrombin cleavage site-TwinStrep-8×His-tag) were expressed in Expi293F (Thermo Fisher Scientific) cells at 37 °C and 8% CO_2. Transfections were performed using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Cell culture supernatants were collected two to four days after transfection and supplemented with 10× PBS to a final concentration of 2.5× PBS (342.5 mM NaCl, 6.75 mM KCl and 29.75 mM phosphates). SARS-CoV-2 RBDs were purified using cobalt-based immobilized metal affinity chromatography followed by buffer exchange into PBS using a HiPrep 26/10 desalting column (Cytiva) or, for the second batch of Omicron RBD used for SPR, a Superdex 200 Increase 10/300 GL column (Cytiva).

The SARS-CoV-2 Wuhan-Hu-1 and Delta (B.1.617.2) RBD-Avi constructs were synthesized by GenScript into pcDNA3.1- with an N-terminal mu-phosphatase signal peptide and a C-terminal octa-histidine tag, flexible linker and avi tag (GHHHHHHHHGGSSGLNDIFEAQKIEWHE).The boundaries of the construct are N-$_{328}$RFPN$_{331}$ and $_{528}$KKST$_{531}$-C (refs. McCallum, 2021; Walls, 2020). Proteins were produced in Expi293F cells (Thermo Fisher Scientific) grown in suspension using Expi293 Expression Medium (Thermo Fisher Scientific) at 37 °C in a humidified 8% CO_2 incubator rotating at 130 rpm. Cells grown to a density of 3 million cells per ml were

transfected using the the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific) and cultivated for 3–5 days. Proteins were purified from clarified supernatants using a nickel HisTrap HP affinity column (Cytiva) and washed with 10 column volumes of 20 mM imidazole, 25 mM sodium phosphate pH 8.0 and 300 mM NaCl before elution on a gradient to 500 mM imidazole. Proteins were biotinylated overnight using the BirA Biotin-Protein Ligase Kit (Avidity) and purified again using theHisTrapHP affinity column. After a wash and elution as before, proteins were buffer-exchanged into 20 mM sodium phosphate pH 8 and 100 mM NaCl, and concentrated using centrifugal filters (Amicon Ultra) before being flash-frozen.

Recombinant production of ACE2 orthologues

Recombinant human ACE2 (residues 19–615 from Uniprot Q9BYF1 with a C-terminal AviTag-10×His-GGG-tag, and N-terminal signal peptide) was produced by ATUM. Protein was purified via Ni Sepharose resin followed by isolation of the monomeric hACE2 by size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva) pre-equilibrated with PBS. The mouse (*Mus musculus*) ACE2 ectodomain construct (GenBank: Q8R0I0) was synthesized by GenScript and placed into a pCMV plasmid. The domain boundaries for the ectodomain are residues 19–615. The native signal tag was identified using SignalP-5.0 (residues 1–18) and replaced with an N-terminal mu-phosphatase signal peptide. This construct was then fused to a sequence encoding a thrombin cleavage site and a human Fc fragment or an 8×His tag at the C terminus. ACE2-Fc and ACE2-His constructs were produced in Expi293 cells (Thermo Fisher Scientific, A14527) in Gibco Expi293 Expression Medium at 37 °C in a humidified 8% CO_2incubator rotating at 130 rpm. The cultures were

transfected using PEI-25K (Polyscience) with cells grown to a density of 3 million cells per ml and cultivated for 4–5 days. Proteins were purified from clarified supernatants using a 1-ml HiTrap Protein A HP affinity column (Cytiva) or a 1-ml HisTrap HP affinity column (Cytiva), concentrated and flash-frozen in 1× PBS, pH 7.4 (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl).

ACE2-binding measurements using SPR

Measurements were performed using a Biacore T200 instrument, in triplicate for monomeric human and mouse ACE2 and duplicate for dimeric mouse ACE2. A CM5 chip covalently immobilized with StrepTactin XT (IBA LifeSciences) was used for surface capture of TwinStrepTag-containing RBDs (Wuhan-Hu-1, Alpha, Beta, Omicron, K417N) and a Cytiva Biotin CAPture Kit was used for surface capture of biotinylated RBDs (Delta and Wuhan-Hu-1 used for fold-change comparison to Delta). Two different batches of Omicron RBD were used for the experiments. Running buffer was HBS-EP+ pH 7.4 (Cytiva) and measurements were performed at 25 °C. Experiments were performed with a threefold dilution series of human ACE2 (300, 100, 33, 11 nM) or mouse ACE2 (900, 300, 100, 33 nM) and were run as single-cycle kinetics. Monomeric ACE2-binding data were double-reference-subtracted and fit to a 1:1 binding model using Biacore Evaluation software. High concentrations of dimeric mouse ACE2 exhibited significant binding to the CAP sensor chip reference flow cell.

Statistical analysis

Neutralization measurements were performed in duplicate and RLUs were converted to per cent neutralization and plotted with a nonlinear regression model to determine

IC_{50}/ID_{50} values using GraphPad Prism software (v.9.0.0). Comparisons between two groups of paired two-sided data were made with Wilcoxon rank test.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

13.7 Figures and Tables

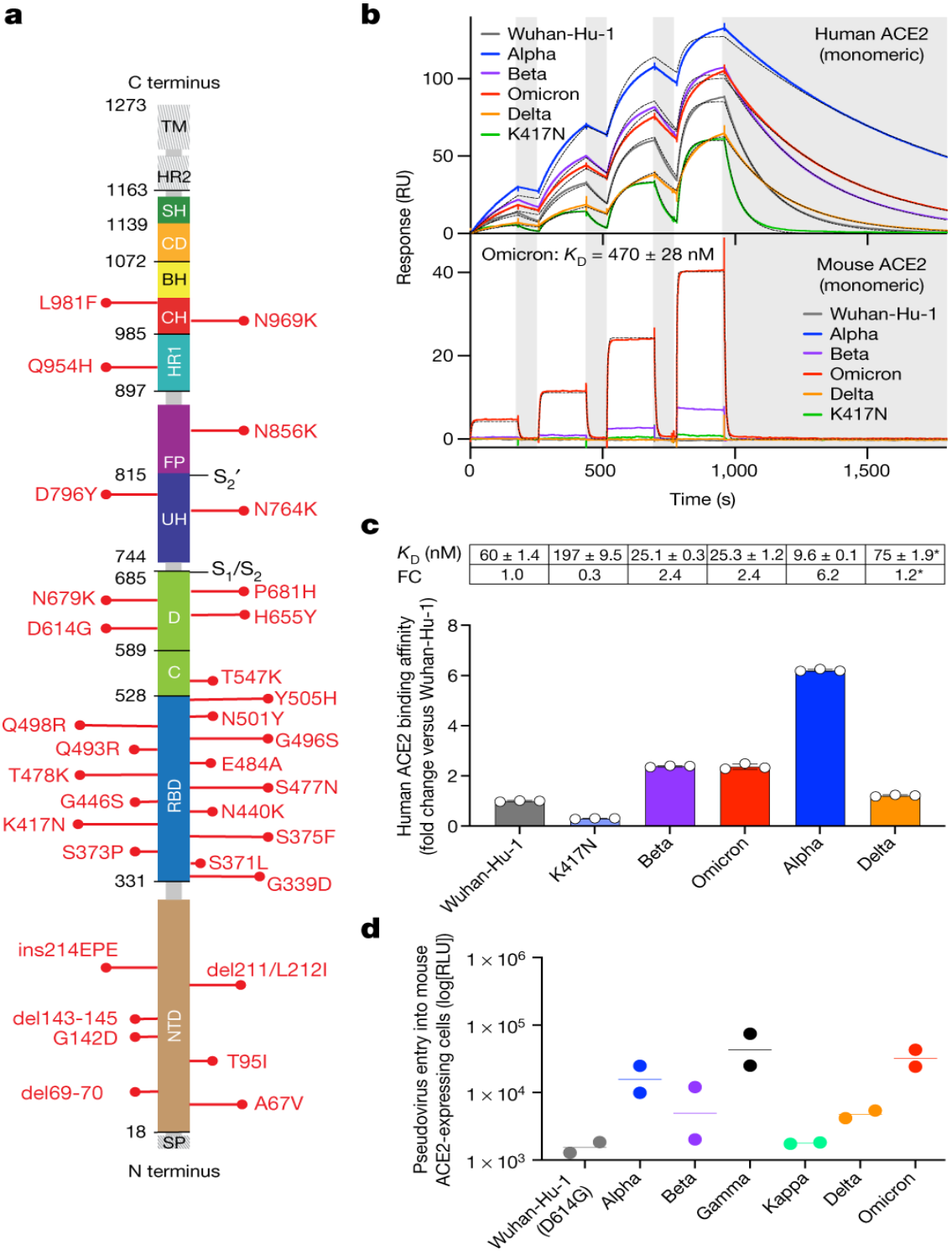


Figure 13.1: The Omicron RBD shows increased binding to human ACE2 and gains binding to mouse ACE2.

a, Omicron mutations are shown in a primary structure of the SARS-CoV-2 spike protein, with domains and cleavage sites highlighted. BH, beta hairpin; C, C domain; CD, connector domain; CH, central helix; D, domain D; FP, fusion peptide;HR1/2, heptad repeat 1/2; SH, stem helix; SP, signal peptide; TM, transmembrane domain;UH, upstream helix. **b**, Single-cycle kinetics SPR analysis of ACE2 binding to six RBD variants. ACE2 is injected successively at 11, 33, 100 and 300 nM (human) or 33, 100, 300 and 900 nM (mouse). Dashed black curves show fits to a 1:1 binding model. White and grey stripes indicate association and dissociation phases, respectively. K_D, dissociation constant; RU, response units. **c**, Quantification of human ACE2 binding data (mean ± s.d. of three replicates). Asterisks indicate that Delta was measured in a separate experiment with a different chip surface and capture tag. Delta fold change (FC) is calculated relative to the affinity of Wuhan-Hu-1 measured in parallel (91 ± 1.6 nM). **d**, Entry of Wuhan-Hu-1, Alpha, Beta, Delta, Gamma, Kappa and Omicron spike VSV pseudoviruses into mouse ACE2-expressing HEK293T cells. RLU, luciferase relative light units. Shown are two biological replicates (technical triplicates). Lines, geometric mean.

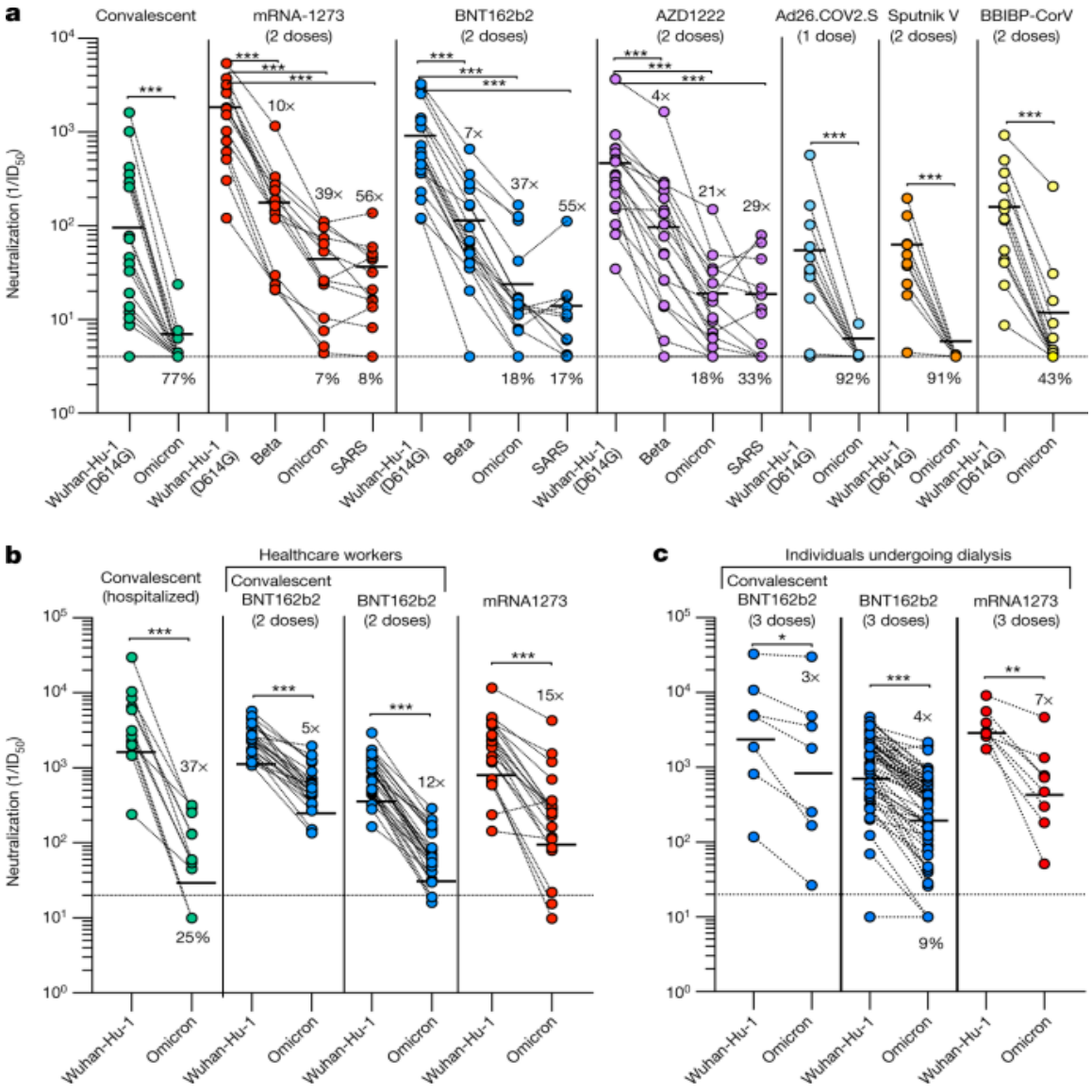


Figure 13.2: Neutralization of Omicron SARS-CoV-2 VSV pseudovirus by plasma from convalescent and vaccinated individuals.

Plasma neutralizing activity in convalescent or vaccinated individuals (mRNA-1273, BNT162b2, AZD1222, Ad26.COV2.S (single dose), Sputnik V and BBIBP-CorV). **a**, Pairwise neutralizing antibody titres (half-maximum inhibitory dose; ID_{50}) against Wuhan-Hu-1 (D614G), Beta and Omicron VOCs, and SARS-CoV. Vero E6-TMPRSS2 cells were used as target cells. Data are the geometric mean of $n = 3$ biologically independent experiments except for SARS-CoV for which $n = 1$. **b**, Pairwise neutralizing antibody titres of plasma (ID_{50}) against Wuhan-Hu-1 and Omicron. Data are the geometric mean of $n = 2$ biologically independent experiments. **c**, Plasma neutralizing activity in individuals undergoing dialysis who received three doses of either the BNT162b2 or mRNA-1273. Pairwise neutralizing antibody titres of plasma (ID_{50}) against Wuhan-Hu-1 and Omicron. One representative experiment out of two is shown. Vero E6

cells were used as target cells in **b**, **c**, Line, geometric mean of 1/ID_{50} titres. Shown is the percentage of samples that lost detectable neutralization against Omicron or SARS-CoV, excluding samples with 1/ID_{50} below the limit of detection. The demographics of enrolled donors are provided in Extended Data Table 1. Statistical significance is set as $P < 0.05$ and P values are indicated with asterisks (*$P = 0.033$; **$P = 0.002$; ***$P < 0.001$), using a paired two-sided t-test (Wilcoxon rank test).

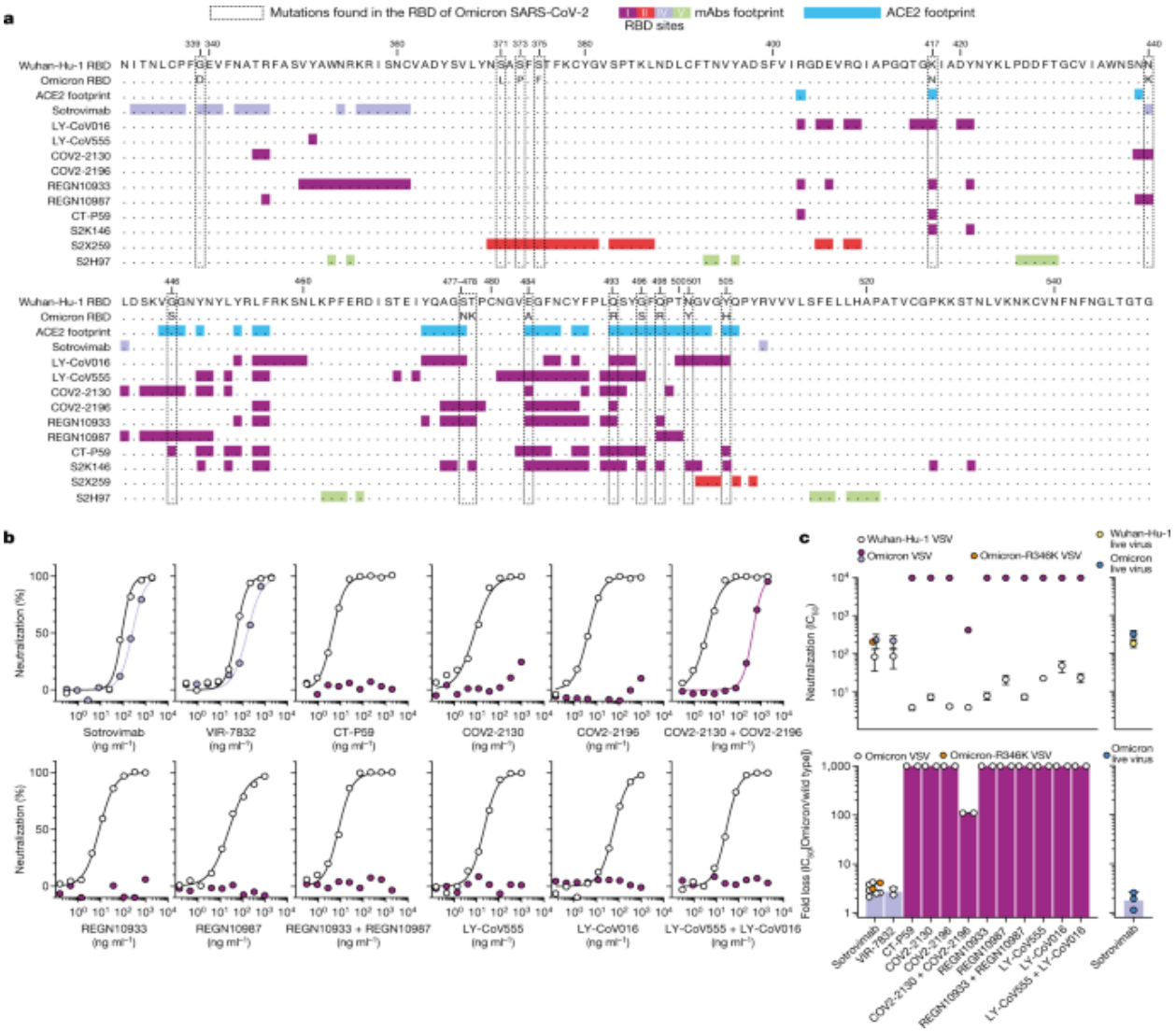


Figure 13.3: Neutralization of Omicron SARS-CoV-2 VSV pseudovirus by clinical-stage monoclonal antibodies.

a, RBD sequence of SARS-CoV-2 Wuhan-Hu-1 with highlighted footprints of ACE2 (light blue) and monoclonal antibodies (mAbs; coloured according to the RBD antigenic site recognized). The Omicron RBD is also shown, and amino acid substitutions are boxed. **b**, Neutralization of SARS-CoV-2 VSV pseudoviruses displaying Wuhan-Hu-1 (white) or Omicron (coloured as in **Fig. 13.4b**) spike proteins by clinical-stage monoclonal antibodies. Data are representative of at least two independent

experiments. Shown is the mean of two technical replicates. **c**, Geometric mean half-maximum inhibitory concentration (IC_{50}) values for Omicron (coloured as in **Fig. 13.4b**) and Wuhan-Hu-1 (white) (top), and geometric mean fold change (bottom). Vero E6 cells were used as target cells. Shown in blue (right) is neutralization of authentic virus by sotrovimab (WA1/2020 versus hCoV-19/USA/WI-WSLH-221686/2021). Non-neutralizing IC_{50} titres and fold change were set to 10^4 and 10^3, respectively. Orange dots for sotrovimab indicate neutralization of Omicron VSV pseudovirus carrying R346K (Omicron-R346K). Data are representative of $n = 2$ biologically independent experiments for most monoclonal antibodies; for sotrovimab against Omicron VSV $n = 6$ and for Omicron authentic virus $n = 3$.

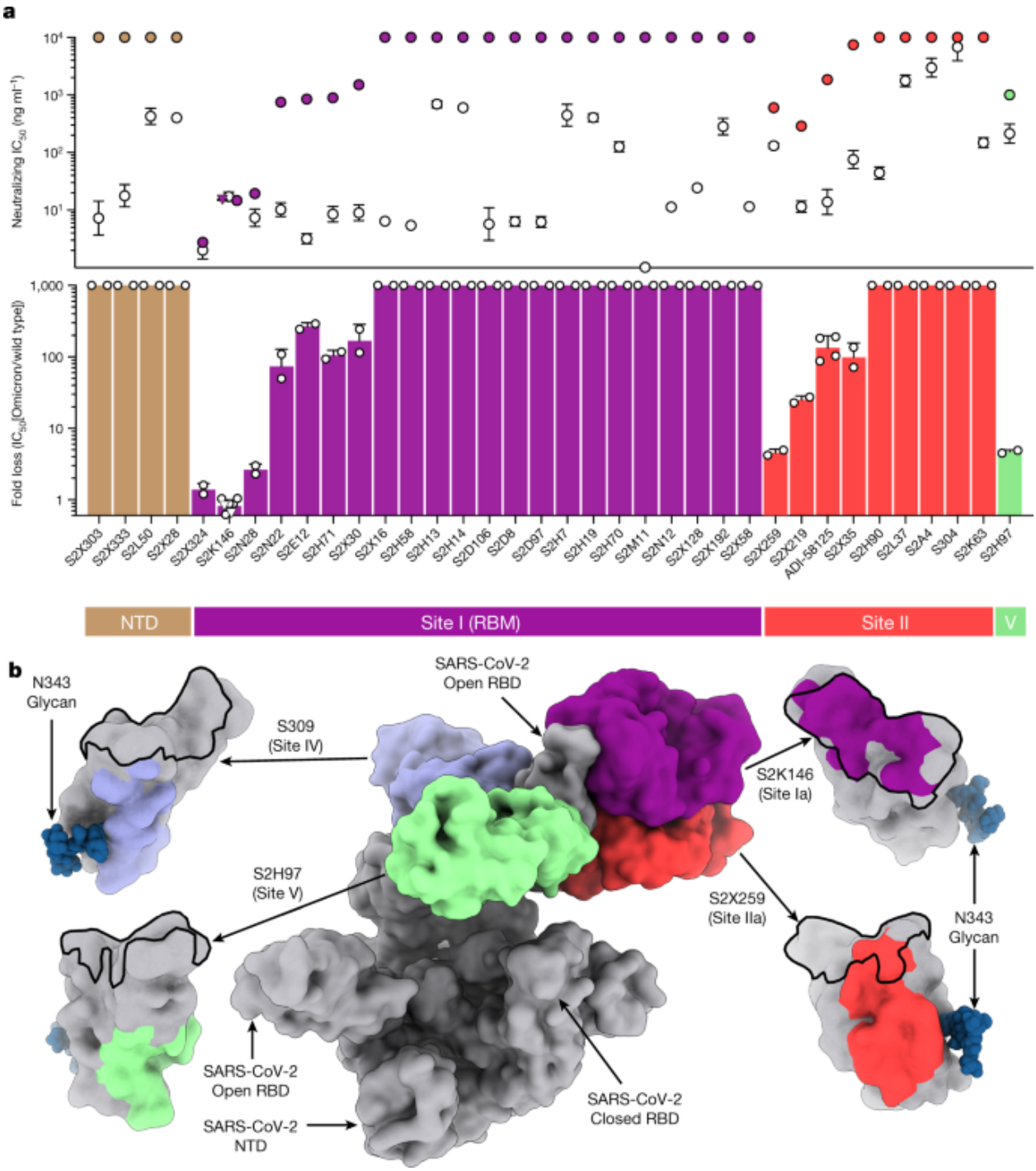


Figure 13.4: Neutralization of Omicron SARS-CoV-2 VSV pseudovirus by monoclonal antibodies.

a, Top, mean IC_{50} values for Omicron (coloured as in **b**) and Wuhan-Hu-1 (white). Bottom, mean fold change for 4 NTD monoclonal antibodies and 32 RBD monoclonal antibodies. Non-neutralizing IC_{50} titres and fold change were set to 10^4 and 10^3, respectively. Triangles for S2K146 indicate neutralization of Omicron carrying R346K. Vero E6 cells were used as target cells. Data are representative of $n = 2$ biologically

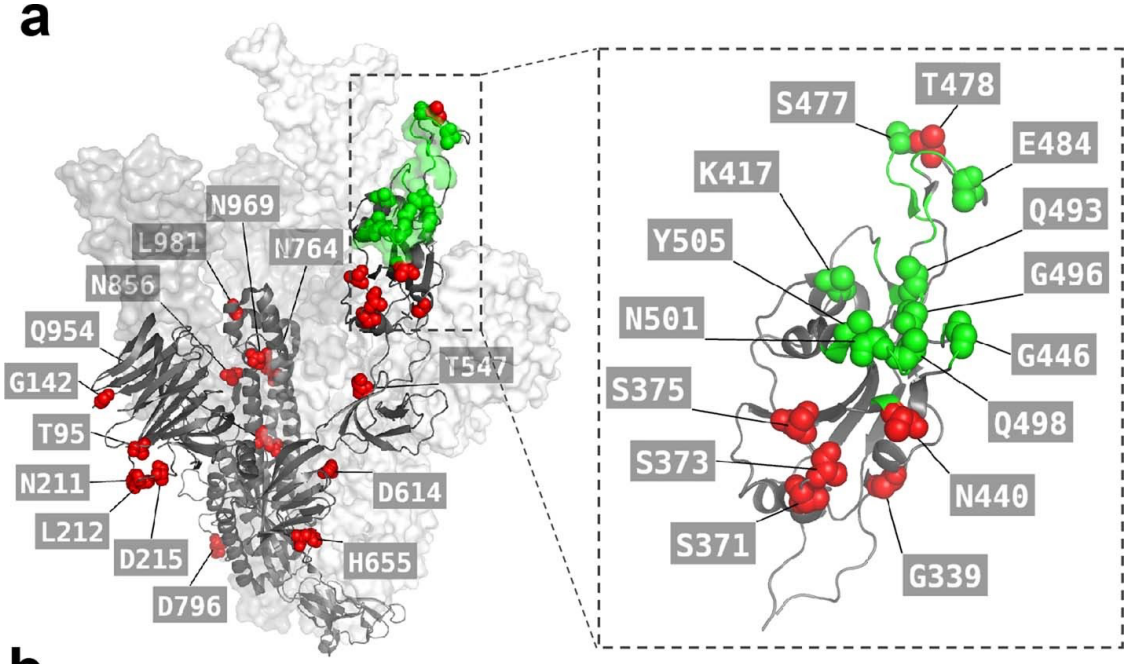
independent experiments (except for S2K146, for which $n = 6$). **b**, The RBD sites targeted by four monoclonal antibodies that cross-neutralize Omicron are annotated and representative antibodies (the Fv region) bound to spike proteins are shown as a composite. Coloured surfaces on the RBD depict the epitopes and the RBM is shown as a black outline.

	PANGO	WHO label	Signal peptide and NTD	RBD	S1/S2	S2
VOC	B.1.1.7	Alpha (α)	H69-/ V70-/ Y144-	N501Y	A570D/D614G/P681H	T716I/S982A/D1118H
	B.1.351	Beta (β)	L18F/D80A/D215G/242-244 del	K417N/E484K/N501Y	D614G	A701V
	P.1	Gamma (γ)	L18F/T20N/P26S/D138Y/R190S	K417T/E484K/N501Y	D614G/H655Y	T1027I/V1176F
	B.1.617.2	Delta (δ)	T19R/G142D/E156G/F157-/R158-	L452R/T478K	D614G/P681R	D950N
	B.1.1.529	Omicron (ο)	A67V / Δ69-70 / T95I / G142D/Δ143-145 Δ211/L212I / ins214EPE	G339D / S371L / S373P / S375F / K417N / N440K / G446S / S477N / T478K / E484A / Q493R / G496S / Q498R / N501Y / Y505H	T547K / D614G/ H655Y/ N679K/ P681H	N764K/ D796Y/ N856K/ Q954H/ N969K/ L981F
VOI	C.37	Lambda (λ)	G75V/T76I/246-252del	L452Q/F490S		T859N
	B.1.621	Mu (μ)	T95I/Y144T/Y145S/ins146N	R346K/E484K/N501K	D614G/P681H	D950N
VUM	B.1.617.1	Kappa (κ)	T95I/G142D/E154K	L452R/E484Q	D614G/P681R	Q1071H
	B.1.526	Iota (ι)	L5F/T95I/D253G	E484K	D614G	A701V
	B.1.525	Eta (η)	Q52R/A67V/H69-/V70-/Y144-	E484K	D614G/Q677H	F888L
Former	B.1.429	Epsilon (ε)	S13I/W152C	L452R	D614G	
	P.2	Zeta (ζ)		E484K	D614G	
	P.3	Theta (θ)	141-143del/242-244del	E484K/N501Y	P681H	S1101Y/E1092K/V1176F

Extended Data Figure 13-1: Schematic of mutations landscape in SARS-CoV-2 VOCs, variants of interest and variants under monitoring.

D, deletion; ins, insertion; VOI, variant of interest; VUM, variant under monitoring.

a



b

Sum of counts (N=12696)	G339	R346	S371	S373	S375	K417	N440	G446	L452	S477	T478	E484	Q493	G496	Q498	N501	Y505	Wu-1
7410	D		L	P	F					N	K	A	R	S	R	Y	H	Omicron VOC, as of December 20, 2021
2044	D		L	P	F	N	K	S		N	K	A	R	S	R	Y	H	
590	D	K	L	P	F					N	K	A	R	S	R	Y	H	
308	D																	
257	D	K	L	P	F	N	K	S		N	K	A	R	S	R	Y	H	
201	D		L	P	F					N	K	A	R	S	R	Y	H	
173						N	K	S		N	K	A	R	S	R	Y	H	
120	D		L	P	F	N				N	K	A	R	S	R	Y	H	
89	D		L	P	F				R	N	K	A	R	S	R	Y	H	
55																		
47	D						K	S		N	K	A	R	S	R	Y	H	
46	D		L	P	F		K	S		N	K	A	R	S	R	Y	H	
40	D		L	P	F		K	S		N	K	A	R	S	R	Y	H	
25	D		L	P	F	N												
22	D	K																
20	D		L	P	F					N	K	A	R	S	R	Y	H	
prevalence (%)	**99.7**	**8.9**	**99.4**	**99.0**	**98.8**	**91.8**	**93.8**	**94.5**	**3.9**	**99.4**	**99.6**	**99.4**	**99.6**	**99.5**	**99.6**	**99.6**	**99.6**	
Alpha (α)																Y		VOC
Beta (β)						N						K				Y		
Gamma (γ)						T						K				Y		
Delta (δ)						*N*			R		K							
Epsilon (ε)									R									VOI
Zeta (ζ)												K						
Eta (η)												K						
Theta (θ)												K				Y		
Iota (ι)												K						
Kappa (κ)									R		Q							
Lambda (λ)									Q									
Mu (μ)		K										K				Y		

Extended Data Figure 13-2: Amino acid substitutions and their prevalence in the Omicron RBD.

a, SARS-CoV-2 spike protein in fully open conformation (PDB: 7K4N) with positions of mutated residues in Omicron highlighted on one protomer in green or red spheres in or outside the ACE2 footprint (ACE2), respectively. RBM is defined by a 6 Å cut-off in the RBD-ACE2 interface (Thomson, 2021). Not all Omicron mutations are shown. **b**, Substitutions and their prevalence in Omicron sequences reported in GISAID as of 20 December 2021 (ambiguous amino acid substitutions are indicated with strikethrough cells). Shown are also the prevalence in Omicron (%) and substitutions found in other variants. K417N mutation in Delta is found only in a fraction of sequences.

Sum of counts (N=12696)	A67	H69	V70	T95	L141	G142	V143	Y144	Y145	I210	N211	L212	R214	Wu-1
10271	V	-	-	I		D	-	-	-		-	I	REPE	Omicron, as of Dec. 20, 2021
353	V	-	-	I		D	-	-	-					
336	V	-	-	I		D	-	-	-		-	I		
153	V			I		X								
61	V	-	-	I		D	-	-	-	IIV	R			
43	V	-	-	I	F						-	I	REPE	
42	V			I										
36	V	-	-	I	F						-	I	REPE	
33	V	-	-	I							-	I	REPE	
25	V	-	-	I							-	I		
23	V	-	-	I		D	-	-	-					
prevalence (%)	99.9	99.9	99.9	99.8	0.8	98.4	99.4	99.5	99.4	0.6	99.7	99.7	93.1	
Alpha (α)		-	-					-						VOC
Beta (β)														
Gamma (γ)														
Delta (δ)				/		D								
Epsilon (ε)														VOI
Zeta (ζ)														
Eta (η)	V	-	-					-						
Theta (θ)					-	-	-							
Iota (ι)				I										
Kappa (κ)				I		D								
Lambda (λ)														
Mu (μ)				I				T						

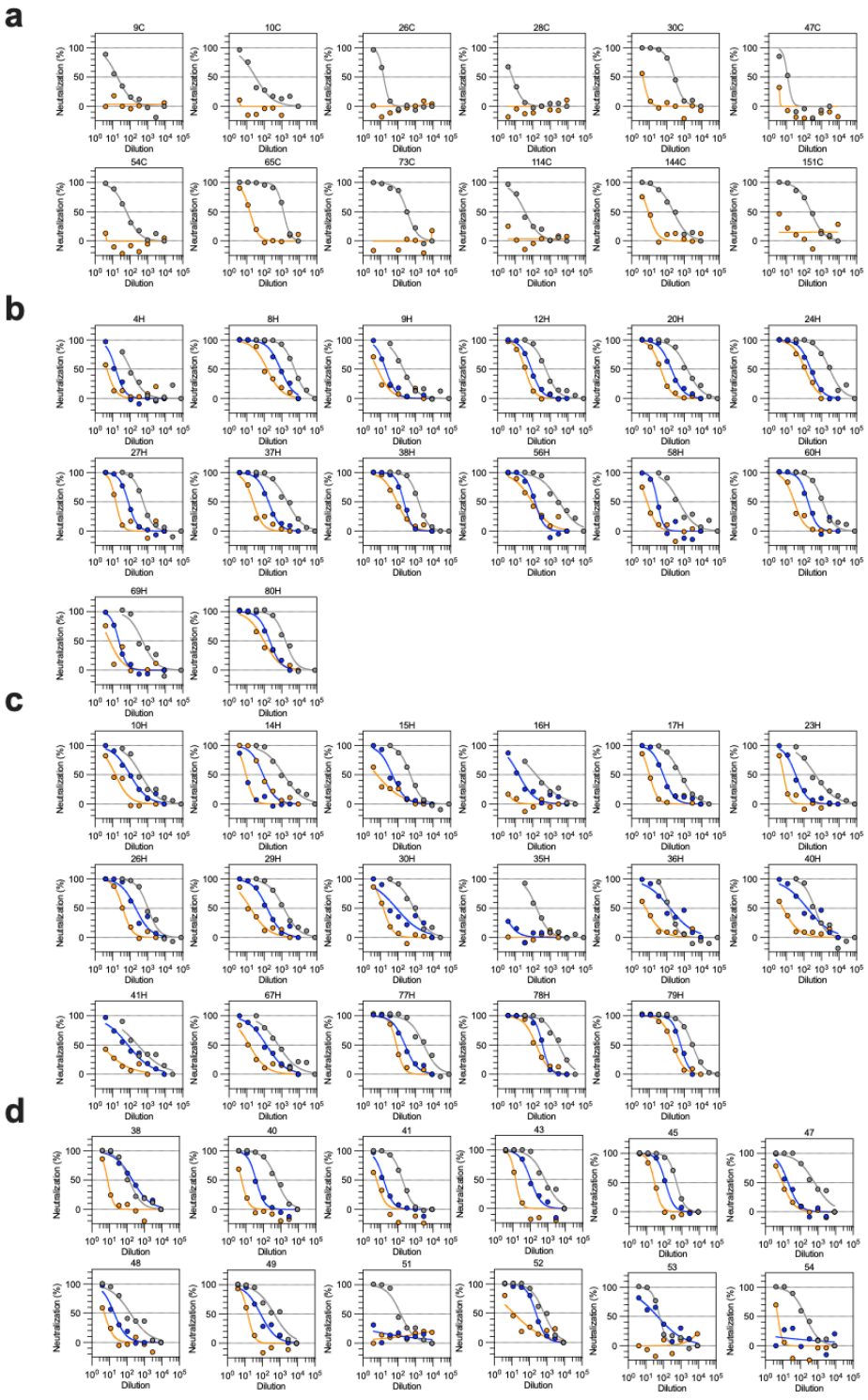
Extended Data Figure 13-3: Amino acid substitutions and their prevalence in the Omicron NTD.

Sequences reported in GISAID as of 20 December 2021 (ambiguous amino acid substitutions are marked with strikethrough cells). Shown are also the substitutions found in other variants.

Sum of counts (N=12696)	N679	P681	A701	N764	D796	N856	Q954	N969	L981	I1081	D1084	V1264	Wu-1
6039	K	H		K	Y	K	H	K	F				Omicron , as of Dec. 20, 2021
1817	K	H	V	K	Y	K	H	K	F				
1018	K	H			Y	K	H	K	F				
1006	K	H			Y	K	H	K	F				
563	K	H	V		Y	K	H	K	F				
516	K	H		K	Y	K	H	K	F	V			
174	K	H			Y	K	H	K	F	V			
92	K	H			Y	K	H	K	F	V			
89	K	H											
38	K	H		K	Y	K	H	K	F			M	
37	K	H		K	Y	K	H	K	F				
27	K	H			Y	K	H	K	F				
21	K	H				K	H	K	F				
20					Y	K	H	K	F				
20	K	H		K	Y	K	H	K	F		E		
prevalence (%)	99.7	99.7	22.9	99.3	99.8	99.8	99.9	99.8	99.8	7	0.2	0.5	
Alpha (α)		H		I									VOC
Beta (β)			V										
Gamma (γ)													
Delta (δ)		R											
Epsilon (ε)													VOI
Zeta (ζ)													
Eta (η)		H											
Theta (θ)													
Iota (ι)			V										
Kappa (κ)		R											
Lambda (λ)													
Mu (μ)		H											

Extended Data Figure 13-4: Amino acid substitutions and their prevalence in Omicron S2.

Sequences reported in GISAID as of 20 December 2021 (ambiguous amino acid substitutions are marked with strikethrough cells). Shown are also the substitutions found in other variants.



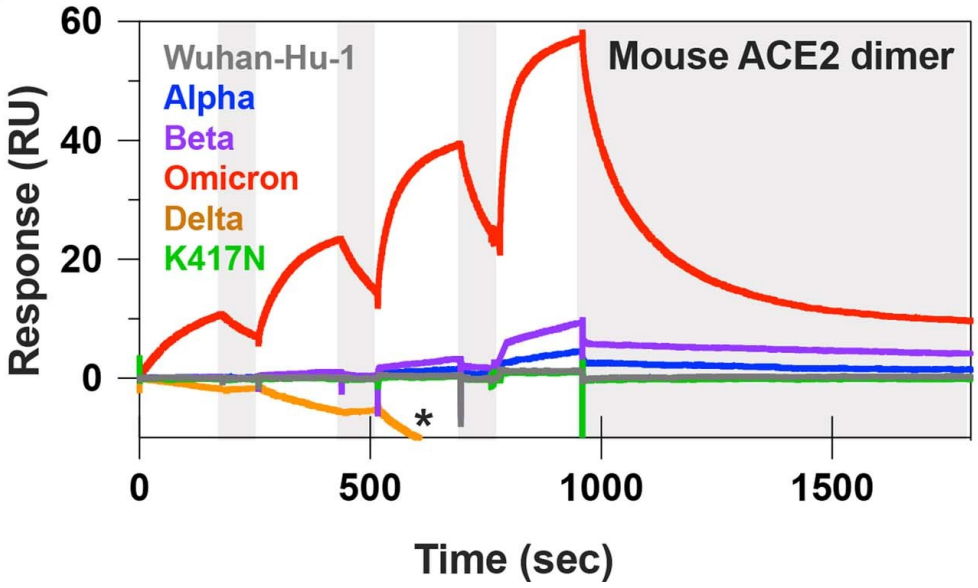
Extended Data Figure 13-5: Characteristics of emergent mutations of Omicron.

a, Shared mutations of Omicron with other sarbecovirus and with VOC. **b**, Since the beginning of the pandemic there is a progressive coalescence of Omicron-defining mutations into non-Omicron haplotypes that may carry as many as 10 of the Omicron-defining mutations. **c**, Pango lineages (dots) rarely carry more than 10-15 lineage-defining mutations. **d**, Exceptionally, some non-Omicron haplotypes may carry up to a maximum 19 Omicron-defining mutations. Shown are selected exceptional haplotypes. Spike G142D and Y145del may also be noted as G142del and Y145D.

a

Analyte	Capture format	RBD variant	ka (1/Ms)	kd (1/s)	KD (M)	Rmax (RU)	RBD capture level (RU)
Monomeric Human ACE2	StrepTactin: TwinStrep Tag	**WT**	1.07E+05	0.00627	5.86E-08	101.9	57
		K417N	9.16E+04	0.01763	1.93E-07	98.6	65
		Beta	1.18E+05	0.00293	2.49E-08	111.2	60
		Omicron	8.87E+04	0.00228	2.57E-08	109.1	73
		Alpha	1.17E+05	0.00111	9.56E-09	131.2	64
	Streptavidin :Biotin	**WT**	6.99E+04	0.00644	9.23E-08	82.92	67
		Delta	5.53E+04	0.00425	7.68E-08	79.65	65
Monomeric Mouse ACE2	StrepTactin: TwinStrep Tag	**Omicron**	3.69E+05	0.1782	4.83E-07	95.75	72

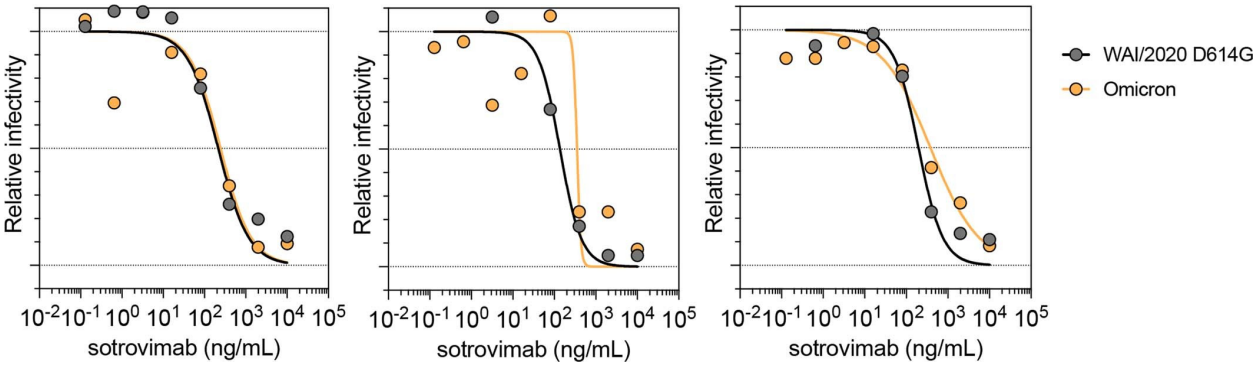
b



Extended Data Figure 13-6: SPR analysis of human and mouse ACE2.

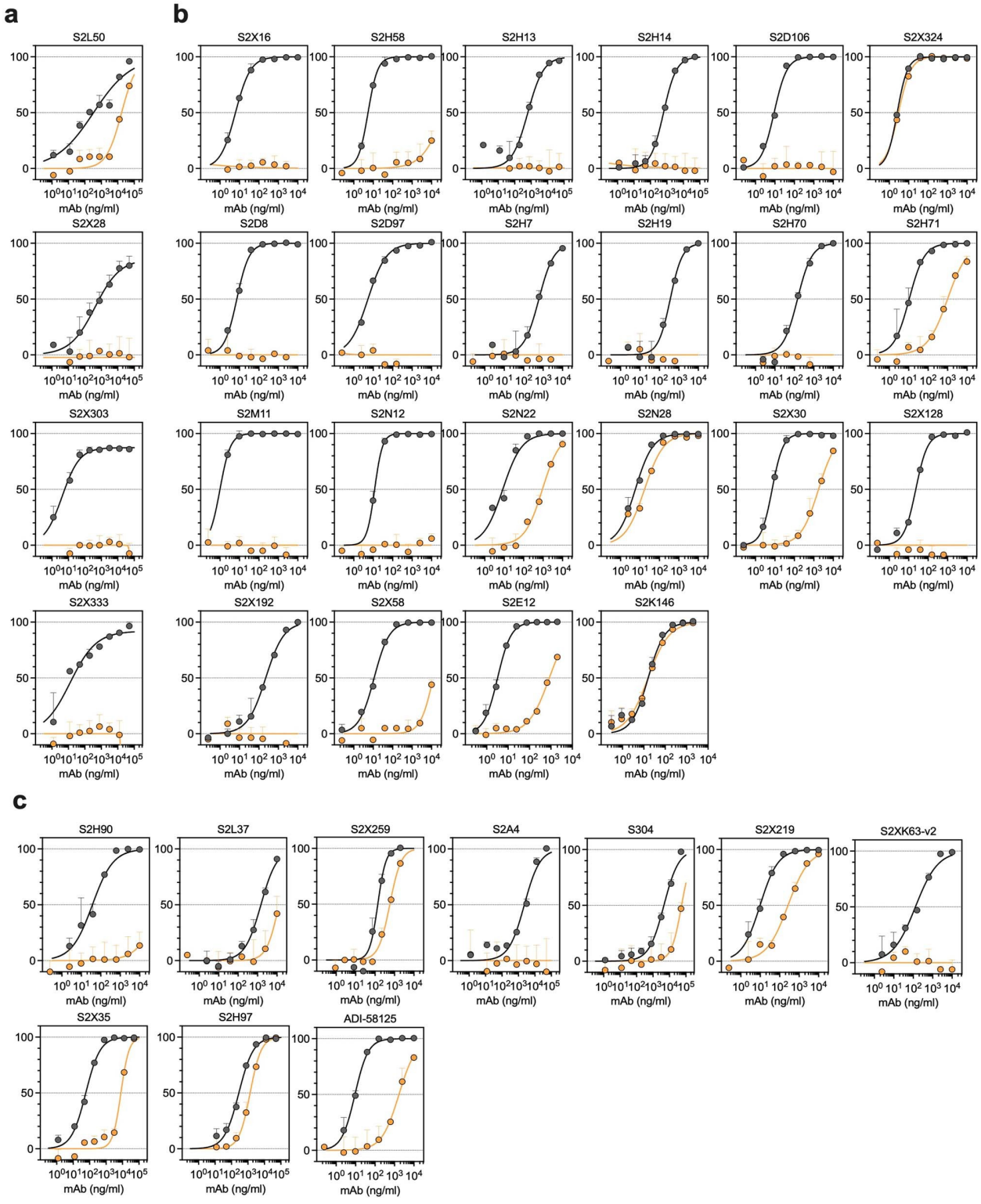
a, Full fit results for one representative replicate from each quantifiable SPR dataset with a monomeric analyte (1:1 binding model). **b**, Single-cycle kinetics SPR analysis of dimeric mouse ACE2 binding to six RBD variants. Dimeric ACE2 is injected successively at 33, 100, 300, and 900 nM. White and grey stripes indicate association and dissociation phases, respectively. The asterisk indicates where high concentrations of dimeric mouse ACE2 are non-specifically binding to the sensor chip surface (Delta

experiment was performed separately from the other RBD variants, with a different capture tag and chip surface).



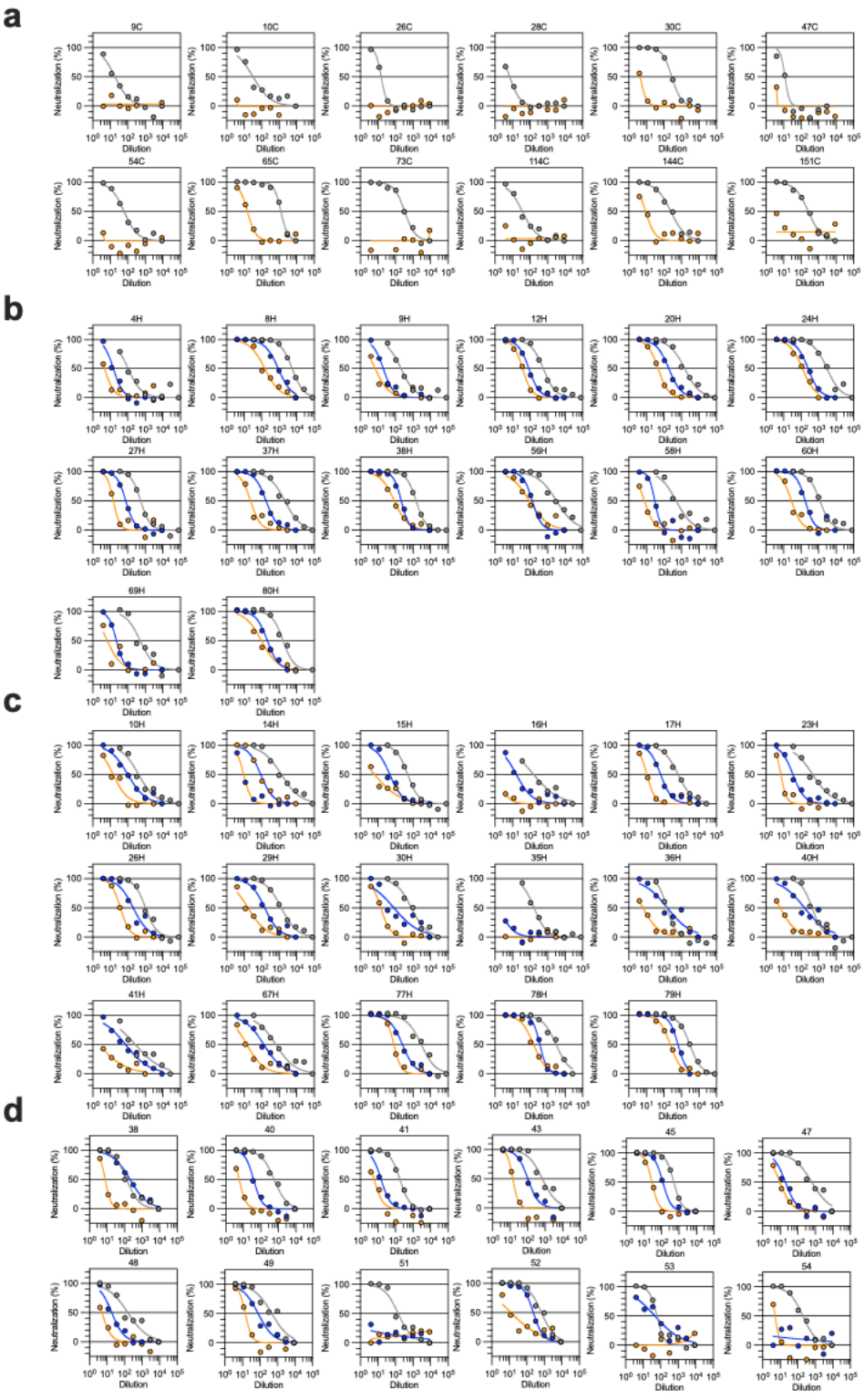
Extended Data Figure 13-7: Neutralization of SARS-CoV-2 Omicron live virus by sotrovimab in Vero-TMPRSS2 cells.

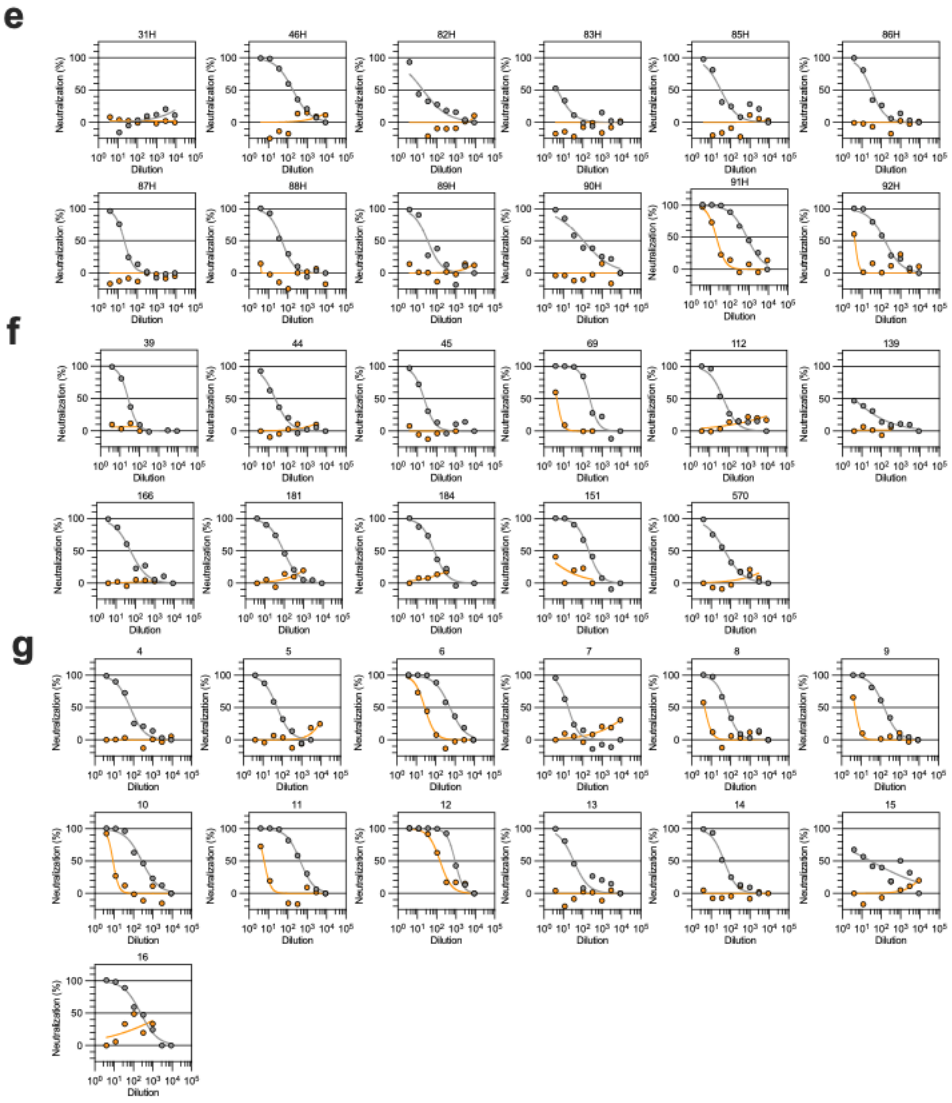
a–**f**, Neutralization curves in Vero-TMPRSS2 cells comparing the sensitivity of SARS-CoV-2 strains with sotrovimab with WA1/2020 D614G and hCoV-19/USA/WI-WSLH-221686/2021 (an infectious clinical isolate of Omicron from a symptomatic individual in the United States). Shown are three independent experiments performed in technical duplicate.



Extended Data Figure 13-8: Neutralization of WT (D614) and Omicron SARS-CoV-2 Spike pseudotyped virus by a panel of 36 monoclonal antibodies.

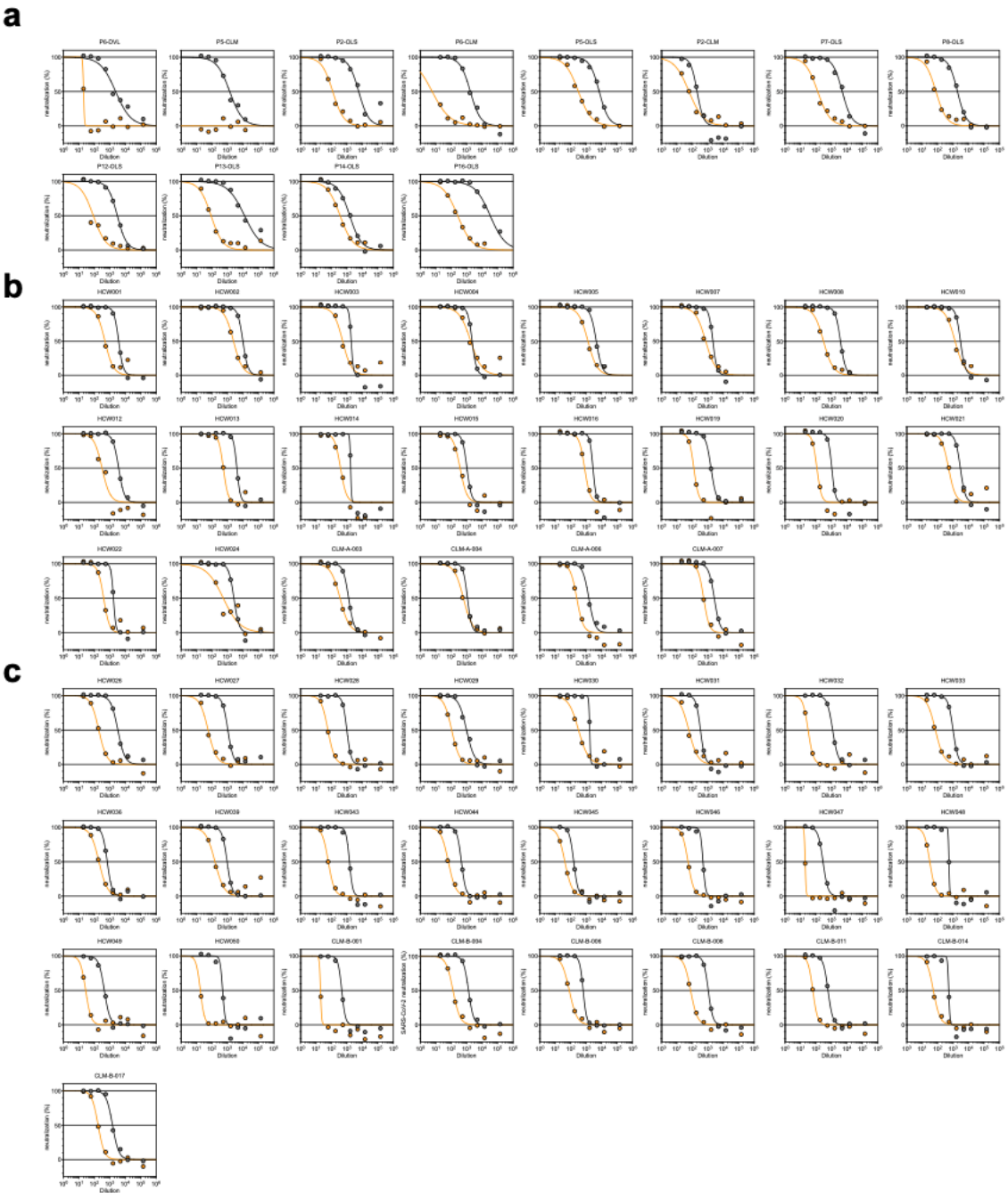
a–**c**, Neutralization of SARS-CoV-2 VSV pseudoviruses carrying wild-type D614 (grey) or Omicron (orange) spike protein by NTD-targeting (**a**) and RBD-targeting (**b**, **c**) monoclonal antibodies (**b**, site I; **c**, sites II and V). Data are representative of two independent experiments. Shown is the mean of 2 technical replicates.


a
9C
10C
26C
28C
30C
47C
54C
65C
73C
114C
144C
151C
b
4H
8H
9H
12H
20H
24H
27H
37H
38H
56H
58H
60H
69H
80H
c
10H
14H
15H
16H
17H
23H
26H
29H
30H
35H
36H
40H
41H
67H
77H
78H
79H
d
38
40
41
43
45
47
48
49
51
52
53
54
Neutralization (%)
Dilution



Supplemental Figure 13.1: Neutralization of WT and Omicron S pseudotyped SARS-CoV-2 virus neutralization by plasma from COVID-19 convalescent and vaccinated individuals.

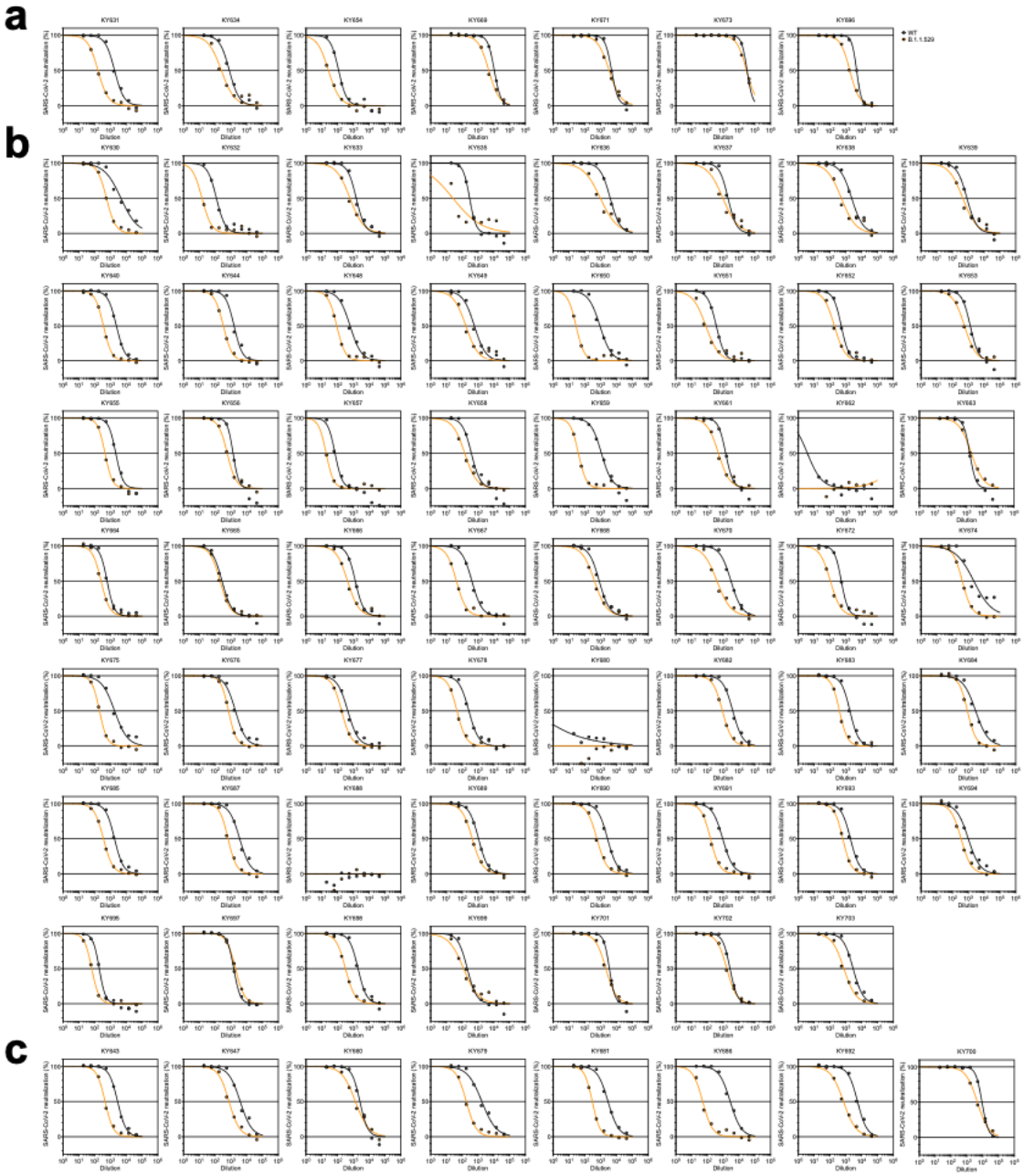
Neutralization of SARS-CoV-2 pseudotyped VSV carrying Wuhan-Hu-1 D614G (grey), Beta (blue) or Omicron (orange) S protein by plasma from convalescent (**a**) or vaccinated individuals (**b**, mRNA-1273; **c**, BNT162b2; **d**, ChAdOx1; **e**, Ad26.COV2.S; **f**, Sputnik V; **g**, BBIBP-CorV) as shown in **Fig. 2a**. Data are representative of *n = 2* independent experiments.



Supplemental Figure 13.2: Neutralization of WT and Omicron S pseudotyped SARS-CoV-2 virus neutralization by plasma from COVID-19 convalescent and vaccinated individuals.

a-c, Neutralization of SARS-CoV-2 pseudotyped VSV carrying Wuhan-Hu-1 or Omicron S protein by plasma from convalescent individuals 2-4 weeks after infection by WT SARS- CoV-2 (a, 11 out of 12 individuals were hospitalized for COVID-19), and

previously infected (b) or naïve (c) individuals, 2-4 weeks after receiving the second dose of BNT162b2 mRNA vaccine. Data are representative of *n = 2* independent experiments.



Supplemental Figure 13.3: Neutralization of WT and Omicron S pseudotyped SARS-CoV-2 by plasma from vaccinated dialysis patients.

a-c, Neutralization of SARS-CoV-2 pseudotyped VSV carrying Wuhan-Hu-1 or Omicron S protein by plasma from convalescent dialysis patients 2-4 weeks after vaccination with 3 doses of BNT162b2 (a), naïve dialysis patients 2-4 weeks after vaccination with 3

doses of BNT162b2 (b) or mRNA1273 (c). Data are representative of *n = 2* independent experiments.

Extended Data Table 13.1: Demographics of enrolled donors.

2-4 weeks after infection/2nd vaccine dose	Dating of SARS-CoV-2 infection	Figure	Nr.	Females	Males	Age (average, range)
Wild type SARS-CoV-2-infected convalescent			**29**	**10**	**19**	**56, 34-73**
Ospedale Luigi Sacco	Mar-Apr 2020	2b	11	1	10	56, 34-73
Swiss volunteers	Mar 2020	2b	1		1	52, 52-52
HAARVI (University of Washington)	Mar-Apr 2020	2a	17	9	8	51, 25-78
Previously infected BNT162b2-vaccinated			**29**	**19**	**10**	**39, 26-56**
Clinica Luganese Moncucco	Mar-Nov 2020	2b	4	3	1	38, 27-54
Ente Ospedaliero Cantonale (EOC)	Mar 2020-Jan 2021	2b	18	14	4	39, 26-56
EOC, dialysis pts	Mar 2020-Jan 2021	2c	7	2	5	69, 48-87
Naïve BNT162b2-vaccinated			**99**	**49**	**50**	**43, 24-67**
Clinica Luganese Moncucco		2b	7	4	3	42, 28-50
Ente Ospedaliero Cantonale (EOC)		2b	18	13	5	43, 24-67
EOC, dialysis pts		2c	55	22	33	74, 29-97
HAARVI (University of Washington)		2a	17	10	7	45, 22-76
Naïve mRNA-1273-vaccinated			**40**	**25**	**15**	
Innovative Research, Novi Michigan (1 week after 2nd dose)		2b	20	14	6	58, 34-74
EOC, dialysis pts		2c	8	2	6	85, 81-92
HAARVI (University of Washington)		2a	14	9	5	47, 23-79
Naïve ChAdOx1-vaccinated						
INGM, Ospedale Maggio Policlinico		2a	17	13	4	38, 29-51
Naïve Sputnik V-vaccinated						
Hospital de Clínicas José de San San Martin, Buenos Aires		2a	11	7	4	42, 30-58
Naïve BBIBP-CorV-vaccinated						
Aga Khan University		2a	13	9	4	30, 25-39

1-19 weeks after 1st vaccine dose			Nr.	Females	Males	Age (average, range)
Naïve Ad26.COV2.S-vaccinated						
HAARVI (University of Washington)		2a	12	6	6	33, 23-60

Total			**250**	**138**	**112**	

Extended Data Table 13.2: Properties of tested monoclonal antibodies.

mAb	Domain (site)	VH usage	Source (DSO)	IC50 Wu-Hu-1 (ng/ml)	IC50 Omicron (ng/ml)	PDB/EMD	Ref.
sotrovimab	RBD (IV)	3-23	SARS-CoV immune	90.6	260	6WPS, 7JX3	2-4,10,51,72,73
					209 (R346K)		
				179 (WA1/2020)	320 (hCoV-19/USA/WI-WSLH-221686/2021)		
VIR-7832	RBD (IV)	3-23	SARS-CoV immune	53.2	165	6WPS, 7JX3	2-4,10,51,72,73
CT-P59	RBD (I/RBM)	N/A	SARS-CoV-2 immune	4.3	>10'000	7CM4	74,75
COV2-2130	RBD (I/RBM)	3-15	SARS-CoV-2 immune	8.1	2772	7L7E	46,48
COV2-2196	RBD (I/RBM)	1-58	SARS-CoV-2 immune	4.3	>10'000	7L7E, 7L7D	46,48
2130+2196				3.8	418		
REGN10933	RBD (I/RBM)	3–11	SARS-CoV-2 huIg mice	8.9	>10'000	6XDG	38-42
REGN10987	RBD (I/RBM)	3–30	SARS-CoV-2 immune	25.1	>10'000	6XDG	38-42
103933+10987				7.2	>10'000		
LY-CoV555	RBD (I/RBM)	1-69	SARS-CoV-2 immune	21.3	>10'000	7KMG	44,45,76,77
LY-CoV016	RBD (I/RBM)	3-66	SARS-CoV-2 immune	59.2	>10'000	7C01	43
555+016				23	>10'000		
S2D106	RBD (I/RBM)	1-69	Hosp. (98)	9.1	>10'000	7R7N	4,49
S2D8	RBD (I/RBM)	3-23	Hosp. (49)	7.3	>10'000		49
S2D97	RBD (I/RBM)	2-5	Hosp. (98)	5.3	>10'000		49
S2E12	RBD (I/RBM)	1-58	Hosp. (51)	3.7	896	7K4N, 7R6X	4,49,51,53
S2H14	RBD (I/RBM)	3-15	Sympt. (17)	625	>10'000	7JX3	4,15,49
S2H19	RBD (I/RBM)	3-15	Sympt. (45)	361	>10'000		49
S2H58	RBD (I/RBM)	1-2	Sympt. (45)	5.4	>10'000		4,49
S2H7	RBD (I/RBM)	3-66	Sympt. (17)	607	>10'000		49
S2H70	RBD (I/RBM)	1-2	Sympt. (45)	145	>10'000		49
S2H71	RBD (I/RBM)	2-5	Sympt. (45)	10.6	993		49
S2M11	RBD (I/RBM)	1-2	Hosp. (46)	1.0	>10'000	7K43	10,49,53
S2N12	RBD (I/RBM)	4-39	Hosp. (51)	11.8	10.8		49
S2N22	RBD (I/RBM)	3-23	Hosp. (51)	8.4	919		49
S2N28	RBD (I/RBM)	3-30	Hosp. (51)	5.8	17.1		49
S2X128	RBD (I/RBM)	1-69-2	Sympt. (75)	23.2	>10'000		49
S2X16	RBD (I/RBM)	1-69	Sympt. (48)	6.2	>10'000		4,49
S2X192	RBD (I/RBM)	1-69	Sympt. (75)	223	>10'000		49
S2X30	RBD (I/RBM)	1-69	Sympt. (48)	7.2	1750		49
S2X324	RBD (I/RBM)	2-5	Sympt. (125)	2.6	3.0		26
S2X58	RBD (I/RBM)	1-46	Sympt. (48)	11.1	>10'000	EMD-24607	4,49
S2K146	RBD (I/RBM)	3-43	Sympt. (35)	14.2	12.6	pending	1
S2H13	RBD (I/RBM)	3-7	Sympt. (17)	628	>10'000	7JV4	4,15
ADI-58125	RBD (II)	3-23	SARS-CoV immune	9.3	1703		78
S2H90	RBD (II)	4-61	Sympt. (81)	37	>10'000		49
S2K63v2	RBD (II)	3-30	Sympt. (118)	129	>10'000		26
S2L37	RBD (II)	3-13	Hosp. (51)	1496	>10'000		26
S2X259	RBD (II)	1-69	Sympt. (75)	129	588	7RA8, 7M7W	3
S2X35	RBD (II)	1-18	Sympt. (48)	58.6	7999	7R6W	4,15
S2X219	RBD (II)	3-53	Sympt. (75)	9.8	268.3		
S304	RBD (II)	3-13	SARS-CoV immune	4603	>10'000	7JX3	4,15
S2A4	RBD (II)	3-7	Hosp. (24)	2285	>10'000	7JVC	15
S2H97	RBD (V)	5-51	Sympt. (81)	280	1368	7M7W	4
S2L50	NTD (i)	4-59	Hosp. (52)	338	>10'000		54
S2X28	NTD (i)	3-30	Sympt. (48)	423	>10'000	EMD-23584	54
S2X303	NTD (i)	2-5	Sympt. (125)	4.5	>10'000	7SOF, 7SOE	10,54
S2X333	NTD (i)	3-33	Sympt. (125)	13	>10'000	7LXW, 7LXY	10,51,54

Table S13.1: Characteristics of single point mutations present in Omicron RBD relative to Wuhan-Hu-1 RBD.

	RBD expr.	hACE2 binding	hACE2 binding
		FACS	SPR
	log fold-change	log fold-change	fold-change
G339D	0.3	0.06	
S371L	-0.61	-0.14	
S373P	-0.22	-0.08	
S375F	-1.81	-0.55	
K417N	0.1	-0.45	0.3
N440K	-0.12	0.07	1.2
G446S	-0.4	-0.2	
S477N	0.06	0.06	
T478K	0.02	0.02	1.2
E484A	-0.23	-0.07	1.0 for E484K
Q493K	0.03	0.05	
G496S	0.12	-0.63	
Q498R	-0.1	-0.06	
N501Y	-0.14	0.24	6.2
Y505H	0.16	-0.71	
Total:	-2.84 (692-fold loss)	-2.39 (245-fold loss)	
K417N/ E484K/ N501Y			2.4

Expression and FACS hACE2 binding (data from Starr, 2020)

CHAPTER 14. AUTOMATIC AND ACCURATE LIGAND STRUCTURE DETERMINATION GUIDED BY CRYO-ELECTRON MICROSCOPY MAPS.

Advances in cryo-electron microscopy (cryoEM) and deep-learning guided protein structure prediction have expedited structural studies of protein complexes. However, methods for accurately determining ligand conformations are lacking. In this manuscript, we develop EMERALD, a tool for automatically determining ligand structures guided by medium-resolution cryoEM density. We show this method is robust at predicting ligands along with sur- rounding side chains in maps as low as 4.5 Å local resolution. Combining this with a measure of placement confidence and running on all protein/ligand structures in the EMDB, we show that 57% of ligands replicate the deposited model, 16% confidently find alternate conformations, 22% have ambiguous density where multiple conformations might be present, and 5% are incorrectly placed. For five cases where our approach finds an alternate con- formation with high confidence, high-resolution crystal structures validate our placement. EMERALD and the resulting analysis should prove critical in using cryoEM to solve protein-ligand complexes.

Adapted from: Muenks, A., Zepeda, S., Zhou, G. *et al.* Automatic and accurate ligand structure determination guided by cryo-electron microscopy maps. *Nat Commun* 14, 1164 (2023). https://doi.org/10.1038/s41467-023-36732-5

14.1 Chapter Introduction

Recent advancements in both microscope hardware and computational processing have led to cryo-electron microscopy (cryoEM) emerging as a mainstream method for biomolecular structure determination. While in ideal cases cryoEM data

approaches atomic resolutions (Yip, 2020; Nakane, 2020; Merk, 2020) most structures determined by cryoEM are in the 3–5 Å resolution range. At these resolutions, model building is time-consuming, error prone, and often ambiguous. To assist this process, methods have been developed to automatically build de novo poly- peptide chains into EM data(Wang, 2015; Terwilliger, 2020; Terashi, 2018; He, 2021), and with the advent of AlphaFold 2, high-quality starting models can oftentimes be obtained from sequence information alone(Jumper, 2021; Baek, 2021). While these methods help build protein models into cryoEM density, tools for automatic fitting of small molecule ligands into cryoEM data are limited. Given the widespread adoption of cryoEM in academia and in industry to support translational studies of drug targets, the ability to accurately model ligand- bound structures is paramount.

There are numerous automated tools from X-ray crystallography for modeling small molecule ligands(Emsley, 2004; Oldfield, 2001; Zwart, 2004; Terwilliger, 2006). However, their methodology is unproven for use in interpreting all but the highest-resolution cryoEM maps. Traditional ligand fitting methods rely on shape and topological features of density maps to match(Emsley, 2004)or build(Oldfield, 2001; Zwart, 2004; Terwilliger, 2006) the ligand into density. But as resolution decreases below 3 Å, the topological features these methods rely on become less defined, and their accuracy in modeling ligands within 1 Å RMSD of a reference ligand falls below 20%(Terwilliger, 2006; Evrad, 2007). While these software packages have been updated to consider cryoEM data, the updates focus on protein modeling without reported updates to small molecule modeling(Casanale, 2020; Chojnowski, 2021) or focus on small molecule refinement instead of automatic model building(van Zundert, 2021).

Along with map features, chemical force fields have provided an energetic approach to accurately fit ligands into their respective density. Two approaches—GemSpot(Robertson, 2020) and MDFF(Vant, 2020)—utilize the ligand- docking software GLIDE to model ligands into cryoEM data. However, both require user input in either selecting models during the protocol or choosing a starting configuration, limiting the automation and applicability of these approaches. The protein modeling software Rosetta recently incorporated a small molecule force field, RosettaGenFF, which accurately models the energetics of arbitrary biomolecules in a manner balanced against Rosetta's protein force field(Park, 2021). Combining this energy model with a genetic-algorithm (GA) optimization method allowing for full receptor side chain flexibility, GALigandDock, yielded superior performance in ligand docking accuracy compared to other state-of-the-art methods.

Here, we leverage the docking power of RosettaGenFF and GA optimization to overcome the challenges of modeling small molecules at near-atomic resolution. We integrate cryoEM density data with the physically realistic force field of RosettaGenFF to create RosettaEMERALD (EM Maps ERoded for Automatic Ligand Docking) for robust ligand modeling into cryoEM maps with no user input during the protocol. We evaluate the performance of EMERALD on all non-ion- mediated ligand-bound protein structures deposited in the EMDB21 and compare our results to their respective deposited structures and high-resolution crystal structures when available.

14.2 Results

An overview of EMERALD is illustrated in **Fig. 14.1**. GALigandDock places ligands in a protein pocket by iteratively refining a pool of 100 conformations, selecting the best 100 models at each generation using predicted energy. To enable this method

to use cryoEM density, two changes were integral: density-guided initial ligand placement and the use of density in model selection at each round. Our initial placement (fully described in Methods) first models density as a pseudo-atomic skeleton (**Fig. 14.1b**). When generating the initial population of ligands, ligands are placed at the center of the skeleton and restrained to points in the skeleton. At each iteration, the population of ligand conformers along with their surrounding flexible side chains are further optimized against the sum of a weighted density correlation and the Rosetta-GenFF energy (**Fig. 14.1c**) and finally refined in Rosetta to minimize the energy of the models (**Fig. 14.1d**). The full protocol generates a structure in 30–120 min, depending on the size of the ligand and the cryoEM map.

To test EMERALD, we ran our docking protocol on all ligands with 25 or fewer rotatable torsion angles present in deposited cryoEM structures determined at a minimum of 6 Å nominal resolution. This yielded 1053 ligands to be placed. For each model, we ran three independent trajectories, and we analyzed the resulting models using three different criteria: (a) agreement of the deposited model to the lowest energy predicted structure; (b) density fit and number of protein/ligand hydrogen bonds; and (c) convergence of the three trajectories. This last criterion is used to evaluate the confidence in a predicted model.

The results of these docking trajectories are summarized in **Fig. 14.2**. In 57% of the cases, our density-guided docking produced a top model within 1 Å RMSD (considering all non-hydrogen atoms in the ligand) of the deposited model after energy minimization (match, **Fig. 14.2a**). While an RMSD cutoff of 2 Å has traditionally been used for docking success, the lack of confidence in the low-resolution reference models

and inability of RMSD to consider receptor contacts led us to divide results further by density correlation and hydrogen bond contacts. There were 401 cases (38%) where EMERALD produced a model with an RMSD value >1Å, and the model was similar or better than the deposited model in both metrics (non-match, similar or better quality). The smallest group belonged to 48 cases where the deposited model was not recapitulated, but the EMERALD model had a worse density fit or fewer hydrogen bonds than the deposited model (non-match, worse quality, 5%). Modeling accuracy decreases as ligand flexibility increases and as the local resolution of the map surrounding the ligand worsens (**Fig. 14.2b, c**). Also, we found that incorporating EM data in GALigandDock is necessary for recapitulating deposited ligand structures with high success rates (**Fig. S14.1**).

Because of the low resolution of the density maps, it is difficult to interpret the quality of docked poses from density fit and receptor interactions alone. To instill more confidence in docking results, we analyzed the convergence among the top-ranked ligand poses across three replicates (**Fig. 14.2d–f**). Of the cases within 1 Å RMSD, 2 or more of the trajectories converge for 81% of cases, further strengthening the quality of the matched cases (**Fig. 14.2d**). Moreover, only 23% of the worse-quality cases converge on the same ligand model (**Fig. 14.2e**). Given how well trajectory convergence agrees with these categories, it can serve as a proxy for confidence when our docked model differs from the reference model in ambiguous cases. 42% of the ambiguous cases have similar top models across our trajectories (**Fig. 14.2f**), giving us confidence in an alternative model to the deposited structure for those entries.

Our dataset includes 15 of the 20 cases benchmarked for the GemSpot pipeline18, with five cases filtered out of the dataset for being peptides or having inter-residue bonds like ion coordination. For 13 of the 15 ligands, EMERALD produced a ligand within 1 Å of the deposited structure, with nine of those placements assessed as confident. For the other two cases, our models disagreed with the deposited model; GemSpot also found solutions different from the deposited model in these two cases.

14.3 Crystal models confirm alternate conformations for EM data

To cross-validate our results—particularly in cases where we found a different solution than the deposited model—we looked for all models with a corresponding high-resolution crystal structure (see Methods). We identified 100 cases where EMERALD converged on a ligand placement and a corresponding high-resolution crystal structure was available. The converged docked model was within 1 Å RMSD of the ligand modeled in the crystal structure for 67% of cases, while 58% of the deposited EM models were within this distance. Considering cases where the model predicted from EMERALD and the reference EM model differ, there were six cases where the EMERALD model was within 1 Å RMSD to the crystal structure while the EM model was not, three cases where the EM model was within 1 Å of the crystal structure but the EMERALD model was not, and eight cases where both models differed from the crystal structure by more than 1 Å. In addition, in five of the six cases where our model predicts the crystal structure, our ligand model improves density correlation by at least 0.03, compared to the deposited cryoEM model.

We show docked models supported by crystal structures in **Fig. 14.3** to highlight the quality of our protocol. These examples include: (a) the hippocampal AMPA receptor with the antagonist MPQX22, where our model makes additional hydrogen

bond and π-stacking interactions with the ligand, matching the crystal structure(Sobolevsky, 2009) (**Fig. 14.3a**); (b) NBQX in an AMPA receptor(Zhang, 2021), where the ligand is flipped, better matching the density, and making bidentate interactions with a nearby arginine residue (**Fig. 14.3b**); (c) DNMDP bound to the SLFN12-PDE3A complex(Garvie, 2021), where small changes better match the crystal structure (**Fig. 14.3c**); (d) an ADP molecule in ClpB disaggregase(Yin, 2021) (**Fig. 14.3d**), where the phosphate groups recapitulate the crystal structure; and (e) a glutamate ligand in the AMPA glutamate receptor (Zhang, 2021), which was missing an oxygen atom in the deposited structure; when the full glutamate molecule is docked, the carboxylates are placed in a configuration matching the crystal structure (Krintel, 2012) (**Fig. 14.3e**).

There were three cases where our docking protocol found a ligand different than the crystal structure, while the EM model matched the crystal structure closely. All three cases were different maps of the same system, a folate molecule bound to MERS-CoV(Park, 2019; Sauer, 2021). In all 3, the EMERALD model and the crystal structure only differ in the placement of a flexible arm with high B-factors in the crystallographic data (**Fig. S14.2**)(Pallensen, 2017). These results lend more support for EMERALD convergence as a confidence metric, which we used to further find instances of alternate ligand conformations.

14.4 Docked poses reveal plausible alternate conformations

Even without crystal structures for reference, trajectory con- vergence and improved ligand density fit provide confidence in other docked poses. In the case of an antimicrobial bound multiple transferable resistance (Mtr) pump(Lyu, 2020), our protocol converges on an ampicillin molecule that is flipped so that its phenyl group is now in a

pocket of unassigned density (**Fig. 14.4a, b**). While the deposited model places the phenyl group sandwiched between two phenylalanine residues (**Fig. 14.4a**), our docked model packs the group near a cluster of hydrophobic residues known to interact with other antibiotics31 (**Fig. 14.4b**). In addition, nearby arginine, serine, and threonine residues have been suggested to generally coordinate ligands binding to the pump(Lyu, 2020); our model has the carboxyl group positioned to make interactions with these residues directly or possibly through bridging water molecules. While it is likely that an antibiotic would bind non-specifically to this site, EMERALD ranks our presented orientation the highest across all three trajectories, and there is a large predicted energy gap (about 10kcal/mol) between the converged conformation and the best- scoring conformation with the phenyl group outside this hydrophobic pocket, suggesting that this pose is strongly favored by EMERALD.

Another instance of improving density fit and receptor interactions is a lipid phosphatidylinositol 4,5-bisphosphate (PIP2) bound to transient receptor potential melastatin member 8 (TRPM8)(Yin, 2019). The EMERALD docked model correlates with the map 10% better than the deposited model, placing all the phosphate groups into density and placing the likely disordered glycerol backbone and beginning of the lipid tails in weaker density (**Fig. 14.4c, d**). Moreover, the 4,5 phosphate groups of our docked model make more interactions with basic residues that bind PIP2 in other structures of TRPM832. While the start of the acyl chains are oriented away from the transmembrane region of the protein, this is likely occurring because the chains are truncated. Considering the phosphate placements in the deposited model do not appear

in the top 20 lowest-energy models for any trajectory and the reasons above, EMERALD predicts a more accurate model of PIP2 binding.

Additional cases with confident alternative models are shown in **Fig. 14.4e–h**. For the ATP analog in a structure of the ATP11C flippase(Nakanishi, 2020) the gamma phosphate sticks out of density in the deposited model (**Fig. 14.4e**) but is modeled into the density and interacting with a nearby lysine residue in the docked model (**Fig. 14.4f**). Finally, our EMERALD model of a small molecule GO52 bound to the CD4-HIV-1 Env SOSIP complex34 confidently fits the amide and piperidine groups into the map better than the deposited map, while keeping the hydrophobic interactions as the deposited model (**Fig. 14.4g, h**).

We next identified cases where: (a) the EMERALD model and deposited structure were different, and (b) half maps were available in the EMDB. For these cases, models were refined into one half map and validated against the other using real-space density correlation. When comparing the deposited and EMERALD models (**Fig. S14.3a**), we found two instances where EMERALD's model fits the validation map worse (**Fig. S14.3b–e**), seven cases where it fits the validation map better (one of which is shown in **Fig. 14.4d**), and saw equivalent quality for the remaining 53 cases.

14.5 Low-confidence unmatched cases show pseudo-symmetry or weak density

While our analysis confidently discovers alternate ligand models, 58% of docked molecules with similar quality to the deposited model have medium or low confidence. We found that small molecules that have pseudo-symmetry or have flexible moieties represent these low- confidence cases because of the challenges they provide from their often noisy and inconclusive density. In some instances, two or more replicates of EMERALD agree on a substructure of the molecule (dark blue, **Fig. 14.5a, b**)(Wang,

2020), but differ in a rotamer of a functional group or a flexible group (light blue, **Fig. 14.5a, b**). For other ligands, ambiguous density leads to little agreement among the reference model and low- energy Rosetta models (**Fig. 14.5c, d**). The authors for the allosteric modulator of a dopamine receptor note the lack of confidence in the deposited structure(Xiao, 2021), but have mutagenesis studies to confirm the conformation modeled (**Fig. 14.5c**)(Hao, 2019). However, one model found with EMERALD aligns with their opposing model and fulfills an unexplained region of density in the deposited model (**Fig. 14.5d**). Altogether, these entries show the difficulty in interpreting cryoEM data at medium to low resolution leading to ambiguous density explanations for a single map, and the limits to automated ligand docking using our protocol.

14.6 Cases with worse ligand models show poor initial sampling

To learn what improvements could be made to EMERALD in the future, we looked at instances where EMERALD predicts a ligand with worse metrics than the reference model. We found that these cases often had density that is discontinuous or noisy, leading to incorrect skeletonization. For a ubiquinone binding electron transport protein(Li, 2021), the density skeleton only finds density near the head group (**Fig. 14.4c**). Without a complete skeleton, the initial population struggles to find the deposited conformation, placing the head group exposed to solvent (**Fig. S14.4b**). In this case, if the 2.63 Å data is instead truncated at 4.0 Å resolution, the density becomes more continuous, and the skeleton generated by EMERALD matches the ligand conformation much more closely (**Fig. S14.4d**). With a complete skeleton, the docked model is no longer worse than the deposited model. The head group of the lowest-energy model makes the same hydrogen bond interactions as the deposited model, and the docked model improves density correlation by 0.03 (**Fig. S14.4e**). This underscores the

importance of the initial sampling step, especially when evaluating ligands with a large number of rotatable bonds, and identifies areas for future upgrades in EMERALD.

14.7 Blind modeling of linoleic acid

To demonstrate our protocol's utility in structure determination, we used EMERALD to create a model for linoleic acid bound in a previously undetermined protein structure. Determining this model manually would be an arduous task considering high flexibility of the ligand (**Fig. 14.6a**). Despite the difficulty of modeling the suspected ligand, EMERALD predicts a small molecule conformation that fits the density, makes an anchoring electrostatic interaction with a neigh- boring arginine residue, and introduces little torsional strain throughout the hydrophobic tail (**Fig. 14.6b**). This placement is supported by the structure of linoleic acid bound to a related protein39. Creating the model required no user input once ligand restraint files were made, and the ease and accuracy when modeling linoleic acid prove the value of EMERALD for structure determination.

14.8 Discussion

Here, we show a method EMERALD that is capable of accurately and automatically producing deposition-ready small molecule models into cryoEM maps without human bias during modeling. After being benchmarked on over one thousand ligand-bound entries in the EMDB, EMERALD identifies a confident solution in 62% of entries, in some cases identifying alternate models supported by crystal structures and map validation. Moreover, we show this fully automated protocol determining the conformation of linoleic acid in a previously unsolved structure.

The method should be generally applicable to most ligands with fewer than 25 rotatable bonds; larger ligands have too large of a search space for this algorithm to

effectively sample. Discontinuous or noisy density also proved challenging, though modified map processing to improve density connectivity was shown to rescue at least one of these cases. Our current approach only models a single ligand at a time, which complicates density assignment for structures with ligands close together like electron transport proteins. Finally, Rosetta's poor handling of metal ions precludes modeling ions as cofactors or as ligands themselves, leaving a significant group of proteins unanalyzed(Andreini, 2008).

Currently, our method requires the modeler to know the identification and approximate binding location of the ligand, a non-trivial task when studying novel protein-ligand complexes. For more utility during model building, expanding our method to recognize potential unmodeled ligand blobs and quickly assess possible ligands to deter- mine identity would be beneficial. As is, however, EMERALD offers an automatic tool for ligand modeling that will prove helpful for the now common scenario of ligand-bound structure determination through cryoEM, and EMERALD will serve as a valuable addition to the toolkit of Rosetta EM modeling methods (Wang, 2015; Frenz, 2017; Frenz, 2019) for model building under one software package.

14.9 Methods

Creating the protein-ligand dataset

All single-particle EMDB entries with an associated ligand bound structure at 6 Å nominal resolution or better as of September 03, 2021 were obtained. Given the specificity of trying to model ions and glycans, structures with only these types of ligands were excluded from the dataset. In addition, the set had several cases with

small molecules in close proximity. To simplify the docking situation, entries with two or more ligands within the binding pocket as defined in our docking protocol were also eliminated from the set. To only have entries with complete macromolecule-ligand complex models that fit the EM density well, structures with a density correlation below 0.4 or that left large regions of density unmodeled were dropped. When considering the first instance of a unique ligand for each EMDB entry, there were a total of 1704 total cases to process for docking.

File preparation for docking

For accurate ligand docking, small molecules need proper protonation states and partial charges. However, the protonation state assigned can depend on the protonation assignment method. To determine the most likely protonation state for a small molecule, we calculated protonation states with three assignment tools—phenix.elbow (Mortiarty, 2009), openbabel (O'Boyle, 2011), and dimorphite (Ropp, 2019)—and selected the protonation state assigned with two or more methods. If there was no agreement or failures during the assignment, the phenix.elbow assignment was used for modeling. SDF files of the first instance of each unique ligand-entry pair were downloaded from the PDB and used for input. For processing with phenix.elbow, all possible hydrogen atoms were added to the SDF file using openbabel, and then hydrogen atoms were removed to the final protonation assignment using phenix.elbow. To generate the protonation state with openbabel, the hydrogen atoms were simply added to the downloaded SDF file at a pH of 7.4. Instead of adding protons to a structure, dimorphite (which utilizes RDKit46) protonated small molecules as SMILES strings, which were

then converted to a structure via openbabel. All three protonation assignment methods agreed for 794 instances with two methods agreeing for 157 cases (**Fig. S14.5**).
With the protonation state assigned, a mol2 file with AM1-BCC partial charges was generated with antechamber(Wang, 2006; Jakalian, 2000). Finally, a Rosetta specific parameters file was created for each ligand. Receptors were cleaned by eliminating non-macromolecular atoms in the PDB file and replacing modified residues with their unmodified correspondent. The ligand to be docked was added to its position in the deposited structure and randomly translated 0.0–2.0Å in any direction before docking.

Density erosion and alignment

To ensure the quality of ligand conformations in the ligand pool during the genetic algorithm, randomly perturbed ligands were aligned into unmodeled density to generate the initial pool. Voxels in the density map within 10 Å of the center of mass of the ligand but >2.5 Å from an atom in the receptor were searched and eroded in a modified erosion algorithm from previously described methods(Zwart, 2004; Greer, 1974). Briefly, voxels were labeled as 0 if their associated density map value was below a density value threshold and labeled as 1 if above it. The voxels were searched in order of density value, and voxels with neighboring voxels of 0 value were removed from the skeleton. If removing a voxel breaks skeleton continuity or if all of the voxel's neighbors had a value of 0, then the voxel was added back into the skeleton. This process was repeated until a skeleton remained of voxels with high density values.
Blobs of density are often discontinuous and difficult to separate from noise at lower resolution. To account for the low resolution, skeletonizing density was performed in two successive steps with increasing strictness on erosion. On the first pass, peaks in the

density were detected and eroded only considering voxels sharing a face with each other. This keeps connections between density blobs that may be disjointed. The remaining voxels were clustered into potential skeleton networks by separating groups of voxels that are 3 Å away from another group. Only the largest network of voxels was chosen for further erosion to eliminate noisy voxels. The largest group of voxels underwent a second, stricter erosion that considered all voxels that share a face or edge with each other, leading to a pseudo-atomic skeleton.

The skeleton was used during initial ligand conformer generation of the genetic algorithm to ensure a starting pool of ligands that already fit into the density. Small molecules were randomly translated and perturbed in the binding pocket and half of the small molecules in the initial pool were aligned to the skeleton. For alignment, the ligands were centered on the center of mass of the skeleton, and then atom-skeleton point pairs were determined. The shortest distance of an atom-skeleton pair while searching over all pairs was found, and this search was repeated until either all atoms or all skeleton points had a unique pairing. For the coordinates in each atom-skeleton pair, the topped out harmonic function in Eq. (1) was used to restrain ligand atoms:

$$E = 36\delta 1 e x_{ij} 2 = 9\text{Þ } ij\delta 1\text{Þ}$$

where Eij is an energy penalty applied and x is the distance in Angstroms between the atom-skeleton pair i, j. The ligand is aligned into the density over two stages of energy minimization with 20 and 15 short rounds of minimizations with the atom-skeleton restraints updated after each round.

Docking protocol and analysis

An initial population of 100 ligands were generated by randomly perturbing across a six-fold axis and the torsion angles of the ligand to be docked. Half of the initial ligands were aligned to the density as described above, while the other half of the population were selected from the top 50 models of 5000 random ligand conformations to ensure diversity in the initial population. The population initialization contributes the longest to EMERALD's completion time. All side chains within 5 Å plus the radius of the ligand of the initial ligand center of mass were also considered for optimization. The ligand population and nearby side chains were optimized over 10 generations of a genetic algorithm using default parameters in GALigandDock and a scoring function with a high electron density score weight of 100 to evaluate a ligand's fit into density. The top 20 ligand conformers at the end of the GA were further optimized along with nearby macro- molecule atoms using a cartesian minimization in Rosetta. Example scripts for running density-guided ligand docking are provided below.

All entries were run in triplicate and the lowest-energy model for each individual run was further analyzed for docking success. Only cases with 25 or fewer torsion angles were analyzed as the search space of ligands with more torsions becomes difficult to fully explore during a GA. This, along with losing cases from inherent failure during ligand processing, left 1053 cases to analyze. Because of a low confidence in the reference models due to their low resolution, docked models were not directly compared to their respective reference models. Instead, all reference models were relaxed into their EM density map in Rosetta using the cartesian minimization used after the genetic algorithm. Along with a symmetry-independent RMSD value, docked models were

compared to reference models by the number of residues that make hydrogen bonds with the ligand and a density correlation calculated in Rosetta. These metrics were used to categorize docking results as matches (docked pose within 1 Å of relaxed reference model); non- match, similar quality (>1Å RMSD, density correlationdock—density correlationdeposited > 0.025 and hydrogen bondsdock—hydrogen bondsdeposited > −1); or non-match, worse quality (>1 Å RMSD, density correlationdock—density correlationdeposited < −0.025 or hydrogen bondsdock—hydrogen bondsdeposited < −1). Further support for docking success was calculated by determining the convergence of lowest energy ligand models across the triplicate runs. The distance between atom pairs across models were calculated and results were further divided into those with two or more trajectories having their lowest energy models within 1 Å RMSD, more than within 1 Å for 60% of atoms, or within 1 Å for fewer than 60% of atoms.

The resolution of cryoEM maps often varies from the nominal resolution of a map, so to analyze the performance of EMERALD against map resolution, we compared docking results to local resolution rather than nominal resolution. Maps with local resolution calculations were generated with MonoRes via the Xmipp software package(Vilas, 2018). The deposited maps were filtered with a Gaussian kernel with a sigma of 0.02 times the map dimensions. Binary masks were created using the filtered maps by keeping voxels with a value above 0.05 times the maximum voxel value in the filtered map. With the binary masks, local resolution estimate maps for all instances were created. To calculate the local resolution surrounding the modeled ligand, the local resolution of all voxels within 5 Å of the ligand were averaged. Voxels with zero local resolution values were not included in the average. Considering that ligand binding sites

are often less- resolved areas of a map, the nominal resolution was reported if the calculated local resolution of a map around the ligand was better than 1 Å than the nominal resolution since an error likely occurred.

The following command in Rosetta was used for the low-pass filter of map EMDB-30475:

$ROSETTA/main/source/bin/density_tools.default. linuxgccrelease -truncate_hires 4.0 -mapfile emd_30475. Map -truncate_map

Comparison of docked and EM models to crystal structures

For each ligand-protein pair in the EMDB dataset, the PDB was searched for structures solved by X-ray crystallography at 2.6 Å resolution or better with at least 50% sequence identity to the protein and containing the same ligand. Results from the PDB were filtered further to only contain entries with similar ligand binding pockets as the corresponding EM model. The crystal models were aligned to the EM models by aligning all residues within 10 Å of the ligand using match- maker in UCSF Chimera(Meng, 2006). Once aligned, the density correlations of the ligands in the crystal models were calculated in Rosetta. All entries with a pocket-aligned RMSD > 1.5 Å and a ligand density correlation lower than 0.1 of the EM model were discarded for being too unalike. This gave 129 ligand-bound EMDB structures with similar crystal models. The 100 cases from this set where EMERALD converged on the same model were analyzed by RMSD to the ligand in the aligned crystal model.

Half-map validation of docked models

Half maps for the instances with the non-match, similar or better quality designation and EMERALD convergence were obtained from the EMDB when

available, giving 62 cases. Maps were sharpened with phenix.auto_sharpen(Terwilliger, 2018), and the deposited and docked models were refined into the first half map using a dualspace refinement in Rosetta.

The density cross correlations for both ligand refined models were calculated with the first and second half map. If the difference in density correlation with the second map was >0.05 between the two models, the model with the higher correlation was considered better.

Visual analysis and images

Figures of ligand-bound models and their EM maps were created using UCSF Chimera(Pettersen, 2004). Maps displayed in figures were changed to 1.0 Å using the vop command in Chimera for visual consistency. Plotting of data was performed using the ggplot2 package in ref. Wickham, 2016.

Example EMERALD scripts

Example Rosetta XML script for docking:

<ROSETTASCRIPTS> <SCOREFXNS>

<ScoreFunction Name="relaxscore" weights="beta_gen pot">

<Reweight scoretype="elec_dens_fast" Weight="100"> <Reweight scoretype="gen_bonded" Weight="1.0"> <Reweight scoretype="coordinate_constraint" Weight="1.

0">

</ScoreFunction>

</SCOREFXNS>

<MOVERS>

<SetupForDensityScoring Name="setupdens" > <LoadDensityMap Name="loaddens"

mapfile="%%map%%" >

<GALigandDock Name="dock" scorefxn="relaxscore" ngen= "10" npool="100"

rmsdthreshold="1.0" smoothing="0.0" ramp_schedule="0.1,1.0" grid_step="0.325"

padding="5. 0" nativepdb="%%native%%" sidechains="auto" final_exact_

minimize="bbsc" random_oversample="100" use_pharmaco phore="false"

skeleton_threshold_const="5.0" neighbor hood_size="7" sample_ring_conformers="1"

reference_ pool="map">

</MOVERS>

<PROTOCOLS>

<Add mover="setupdens">

<Add mover="loaddens">

<Add mover="dock"> </PROTOCOLS>

<OUTPUT scorefxn="relaxscore">

</ROSETTASCRIPTS>

Example command line for docking

$ROSETTA/main/source/bin/rosetta_scripts.linuxgc crelease \

-in:file:extra_res_fa $ligand_params_file \ -in:file:overwrite_database_params \

-gen_potential \

-database $ROSETTA/main/database \ -score::gen_bonded_params_file

scoring/score_functions/

```
generic_potential/generic_bonded.round6p.txt \ -s $input_model \
-overwrite \
-multi_cool_annealer 10 \
-parser:protocol $xml:file \ -parser:script_vars map=$em_map \ -atom_mask 2 \
-sliding_window 1 \ -edensity::score_sliding_window_context \ -edensity::mapreso $reso
\ -edensity::grid_spacing 1.0 \ -no_autogen_cart_improper
```

14.9 Figures and Tables

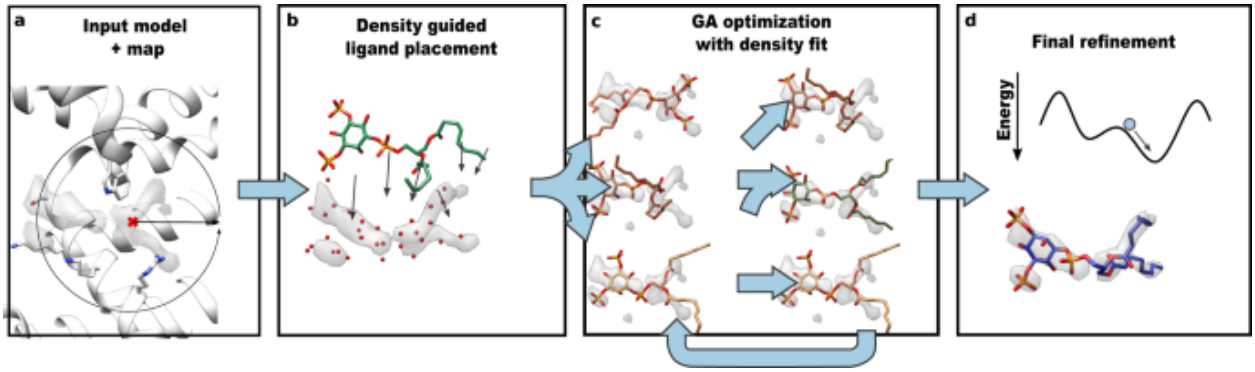


Figure 14.1: Overview of EMERALD docking protocol.

a The cryoEM map, coordinates of the receptor, and the location of the binding site (red cross) are provided as inputs. The binding pocket is calculated depending on the radius of the ligand (circle) to determine boundaries and side chains to consider when modeling. **b** All unmodeled density in the pocket is converted to a pseudo-atomic skeleton (independent of ligand identity), which is used to generate an initial set of ligand conformers. **c** Using a genetic algorithm, the pool of ligand conformers is optimized against Rosetta energy and density fit. The population of ligand conformers evolve over 10 generations with low energy conformations surviving and combining attributes with each other. **d** The 20 poses with lowest energy are refined in Rosetta.

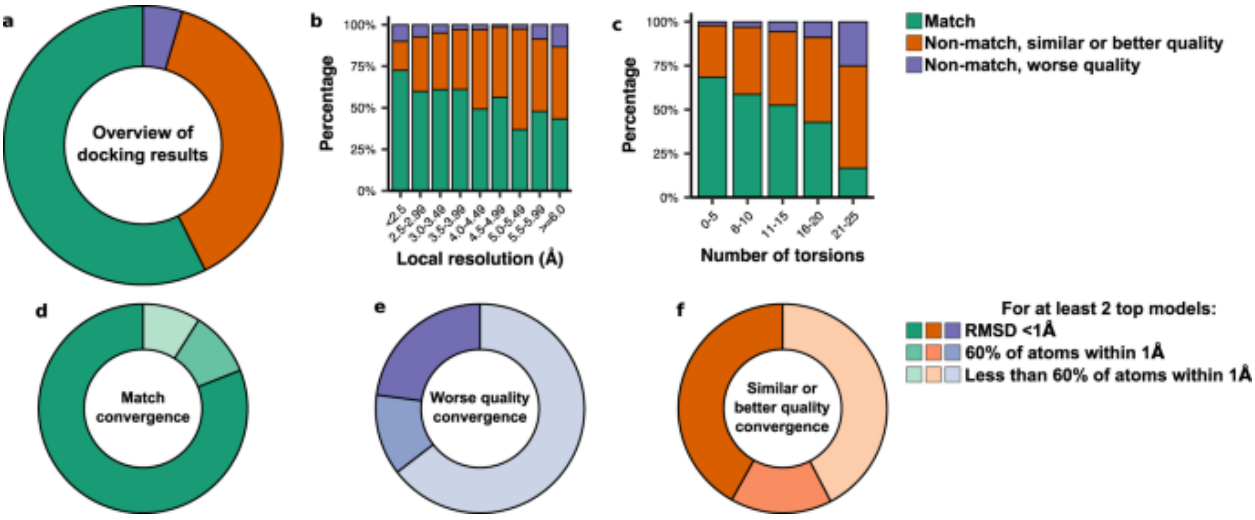


Figure 14.2: Benchmarking EMERALD against the EMDB.

a A comparison of EMERALD models to the deposited structures for 1053 EMDB-deposited complexes. In total, 57% of EMERALD-docked models were placed within 1 Å RMSD of the deposited ligand (match, green); 38% were more than 1 Å RMSD of the

deposited ligand but had similar or better density correlations and numbers of hydrogen bonds (similar or better quality, orange), and 5% were more than 1 Å RMSD from the deposited ligand and had worse density correlations or number of hydrogen bonds (worse quality, blue). **b**, **c** Bins of binding pocket local resolution (**b**) and number of torsion angles in the small molecule (**c**) shown as percentage by docking result. **d**–**f** The convergence of the best ranking models across multiple runs for matches (**d**), worse quality (**e**), and similar or better quality (**f**) cases. The darkest shade had multiple runs converge with all atoms within 1 Å of each other, the middle shade had multiple runs converge with at least 60% of atoms within 1 Å, and the lightest shade had divergent top-scoring models.

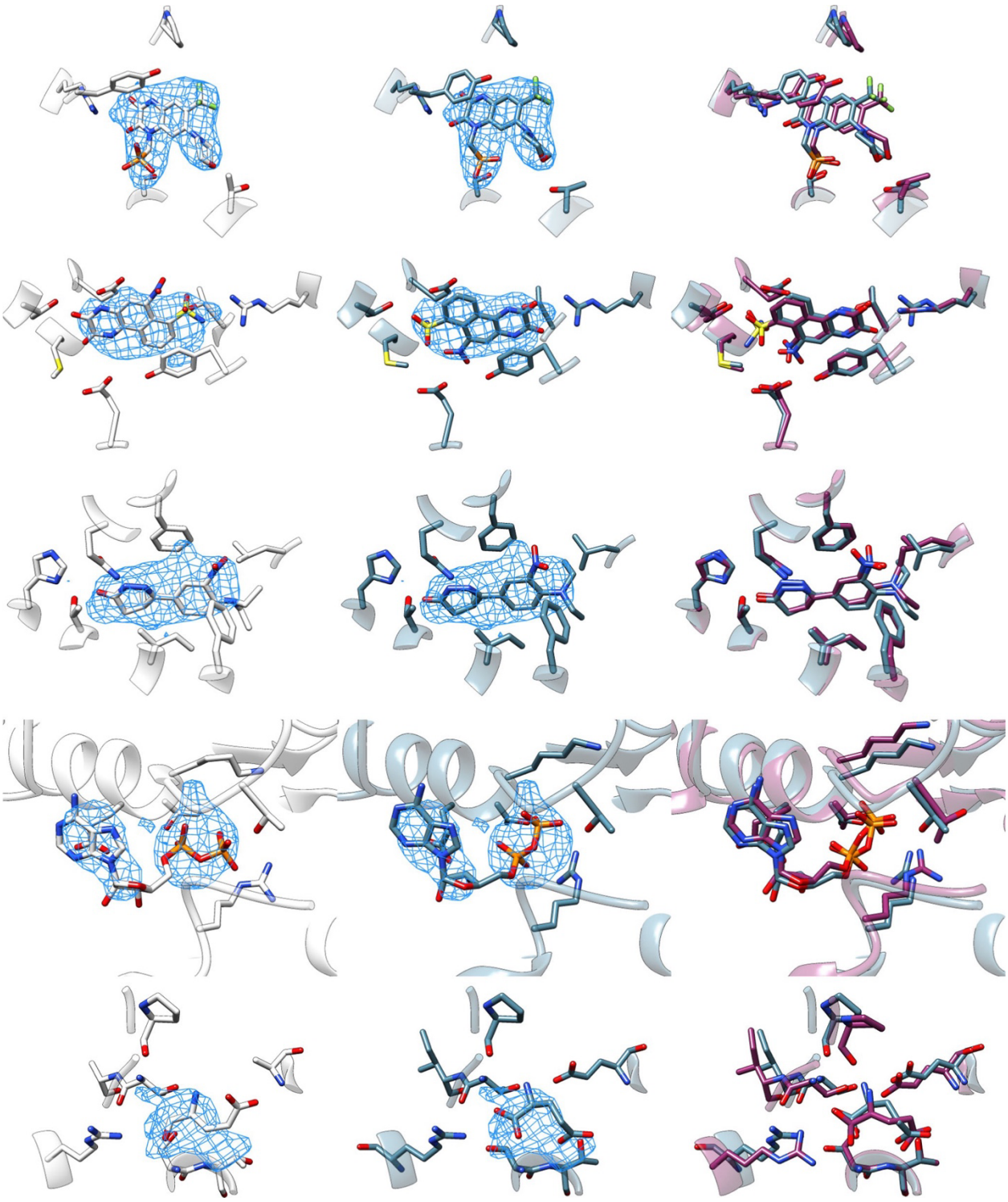


Figure 14.3: Alternate conformations predicted with EMERALD match high-resolution crystal structures.

a–e Comparison of the deposited model (left, white), EMERALD model (center and right, blue), and higher resolution crystal model (right, purple). a Antagonist ZK200775 in AMPA receptor (EMDB: 23292, PDB: 7LEP, local resolution: 3.45 Å) and its

associated crystal model (PDB: 5ZG2). b Molecule NBQX bound to the AMPA receptor (EMDB: 12805, PDB: 7OCE, local resolution: 2.75 Å) and its associated crystal model (PDB: 6FQH). c DNMDP bound to the SLFN12-PDE3A complex (EMDB: 23495, PDB: 7LRD, local resolution: 2.95 Å) and its associated crystal model (PDB: 7KWE). d ADP bound to ClpB (EMDB: 21553, PDB: 6W6E, local resolution: 4.42 Å) and its associated crystal model (PDB: 5LJ8). e Glutamate ligand in an AMPA receptor (EMDB: 12806, PDB: 7OCF, local resolution: 4.26 Å) and its associated crystal model (PDB: 3TKD).

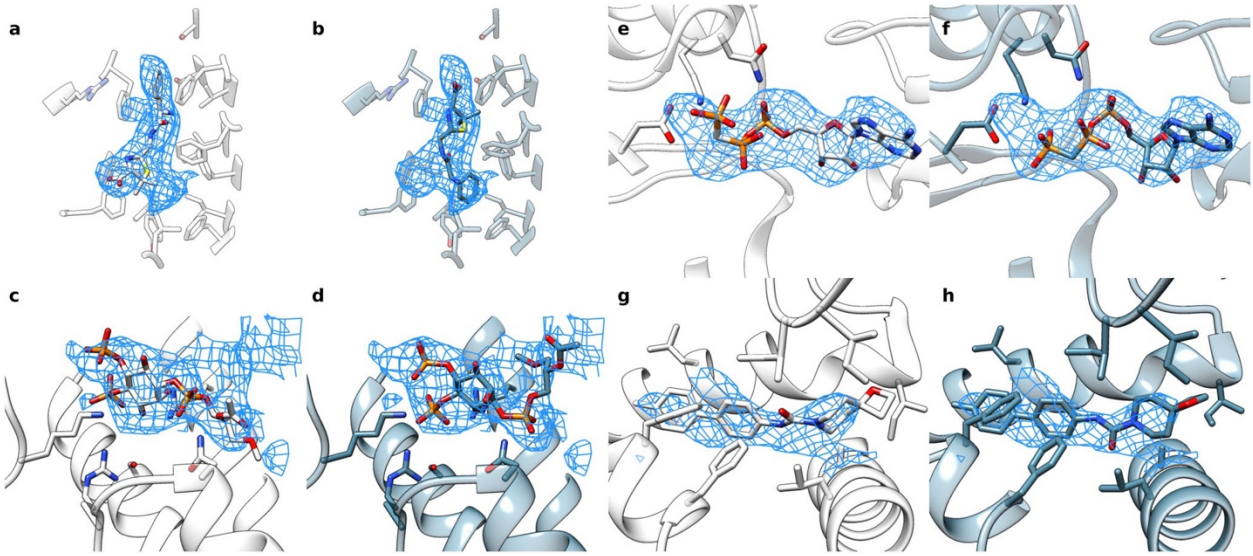


Figure 14.4: Alternate conformations found by EMERALD in cases without crystal structures.

a, b The deposited structure (a) and EMERALD docked model (b) of Ampicillin bound to Mtr pump (EMDB: 21228, PDB: 6VKS, local resolution: 3.84 Å). c, d The deposited structure (c) and EMERALD docked model (d) of PIP2 bound to the TRMP8 ion channel (EMDB: 0487, PDB: 6NR2, local resolution: 4.07 Å). e, f The
groups of our docked model make more interactions with basic resi- dues that bind PIP2 in other structures of TRPM832. While the start of the acyl chains are oriented away from the transmembrane region of the protein, this is likely occurring because the chains are truncated. Considering the phosphate placements in the deposited model do not appear in the top 20 lowest-energy models for any trajectory and the reasons above, EMERALD predicts a more accurate model of PIP2 binding.

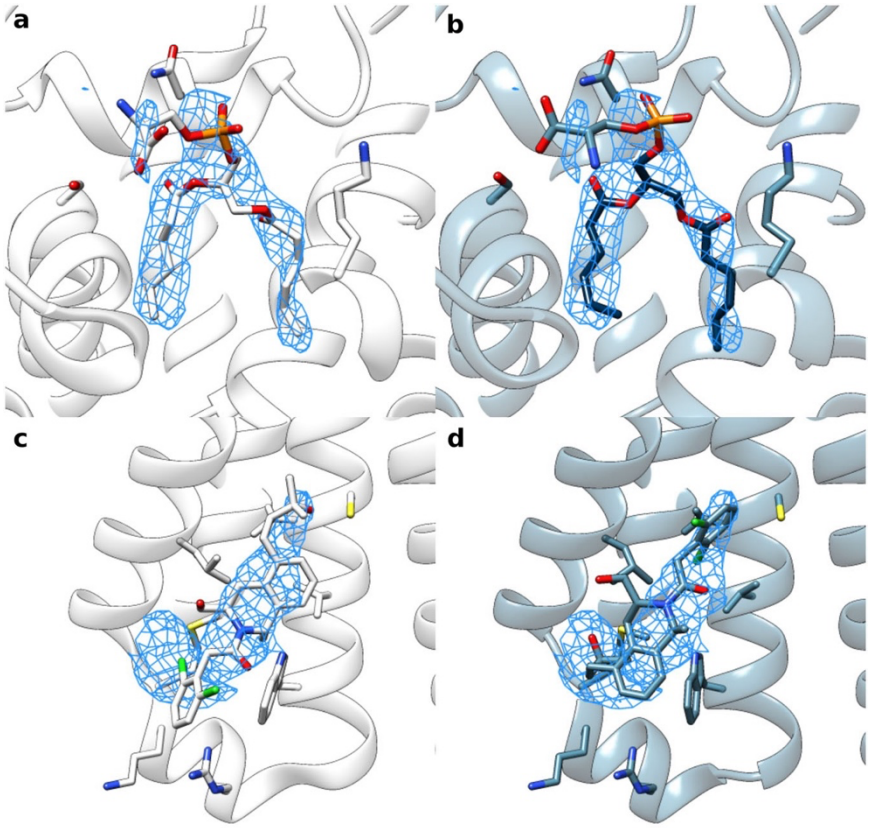


Figure 14.5: Examples of low-confidence docked models where there may be ambi- guity or heterogeneity in the data.

a The deposited model of a phosphoserine lipid in ATPase (EMDB: 21844, PDB: 6WLW, local resolution: 2.75 Å) and the EMERALD-docked model (b) place the fatty acid tails in strong density, but have differences in the head group. The lowest-energy models for all three triplicates

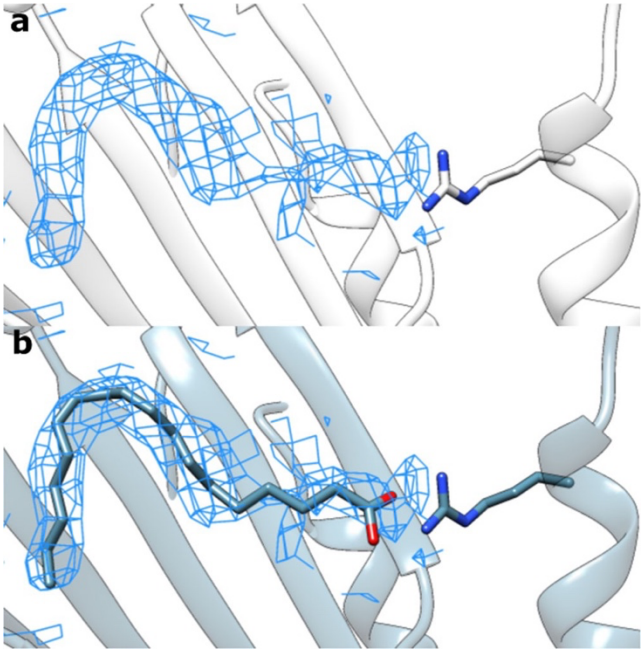


Figure 14.6: Blind modeling of linoleic acid.

a Unmodeled density for linoleic acid (local resolution: 2.9 Å). **b** Output model from density-guided docking. The model makes an anchoring electrostatic interaction with a nearby arginine residue and models the tail strain-free into the density.

CHAPTER 15. CONCLUDING REMARKS

From the studies I have presented here we have learned that ACE2 receptor usage is an ancestral and evolvable trait amongst sarbecoviruses, that has been lost in clade 2 sarbecoviruses (Chapter 2) that even those bat-borne sarbecoviruses that currently lack the ability to use human ACE2 for cellular entry can gain the ability within a few mutations (Chapter 2-4). We elucidate the role of RBD conformation in acting as a barrier to spillover of sarbecoviruses, wherein sarbecoviruses with open RBDs; like SARS-CoV-2 can facilitate cellular entry lower affinity interactions and those with closed RBD's require higher affinities for cellular entry (like RshTT200, and RaTG13) (Chapter 2). We describe receptor usage of coronaviruses across Rhinolophus bat, domesticated animal and wildlife populations (Chapter 2-5). We identify and characterize ACE2 receptor usage farther back on the coronavirus evolutionary tree in merbecoviruses, specifically with HKU5 and describe the very different binding mode used (Chapter 5). We characterize several monoclonal antibodies antigenicity and neutralization of not only bat borne coronaviruses (Chapter 3-8, 10) but also across SARS-CoV-2 variants (Chapter 6-8, 10-13). We describe sarbecovirus targeting nanoparticle vaccination design and protection in mice and non-human primates (Chapter 9). We also describe SARS-CoV-2 Omicron variants discussing structural basis of its altered ACE2 affinity compared to past variants, expanded receptor tropism to include mouse ACE2, antigenicity, fusogenicity, cleavage and alterations to its TMPRSS2 usage (Chapter 11-13). Finally, we describe the development of a ligand fitting software EMERALD and its effectiveness in accurately placing ligands in structural datasets (Chapter 14). A

software that was used in the placement of linoleic acid in our RshTT200 Hexapro Spike structure (Chapter 3). Together this body of work has advanced our understanding of sarbecoviruses in bats, their receptor usage and spillover potential. It has also identified therapeutics that should be stockpiled in case of another sarbecovirus spillover.

REFERENCES

Abdelnabi, R., Boudewijns, R., Foo, C. S., Seldeslachts, L., Sanchez-Felipe, L., Zhang, X., Delang, L., Maes, P., Kaptein, S. J. F., Weynand, B., Vande Velde, G., Neyts, J., & Dallmeier, K. (2021). Comparing infectivity and virulence of emerging SARS-CoV-2 variants in Syrian hamsters. *EBioMedicine*, *68*, 103403.

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500.

Abu Jabal, K., Ben-Amram, H., Beiruti, K., Batheesh, Y., Sussan, C., Zarka, S., & Edelstein, M. (2021). Impact of age, ethnicity, sex and prior infection status on immunogenicity following a single dose of the BNT162b2 mRNA COVID-19 vaccine: real-world evidence from healthcare workers, Israel, December 2020 to January 2021. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, *26*(6), 2100096. https://doi.org/10.2807/1560-7917.ES.2021.26.6.2100096

Adams, L.E., Leist, S.R., Dinnon, K.H., 3rd, West, A., Gully, K.L., Anderson, E.J., Loome, J.F., Madden, E.A., Powers, J.M., Scha€fer, A., et al. (2023). Fc-mediated pan-sarbecovirus protection after alphavirus vector vaccination. Cell Rep. *42*, 112326.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 2), 213–221. https://doi.org/10.1107/S0907444909052925

Addetia, A., Piccoli, L., Case, J.B., Park, Y.J., Beltramello, M., Guarino, B., Dang, H., Pinto, D., Scheaffer, S.M., Sprouse, K., et al. (2023). Therapeutic and vaccine-induced cross-reactive antibodies with effector function against emerging Omicron variants. Preprint at bioRxiv. https:// doi.org/10.1101/2023.01.17.523798.

Addetia, A., Piccoli, L., Case, J.B., Park, Y.-J., Beltramello, M., Guarino, B., Dang, H., de Melo, G.D., Pinto, D., Sprouse, K., et al. (2023). Neutralization, effector function and immune imprinting of Omicron variants. Nature. 10.1038/s41586-023-06487-6.

Addetia, A., Stewart, C., Seo, A.J., Sprouse, K.R., Asiri, A.Y., Al-Mozaini, M., Memish, Z.A., Alshukairi, A.N., and Veesler, D. (2024). Mapping immunodominant sites on the MERS-CoV spike glycoprotein targeted by infection-elicited antibodies in humans. Cell Rep. *43*, 114530.

Aggarwal, A. et al. SARS-CoV-2 Omicron: evasion of potent humoral responses and resistance to clinical immunotherapeutics relative to viral variants of concern. Preprint at https://doi.org/10.1101/2021.12.14.21267772(2021).

Agirre, J., Iglesias-Ferna´ ndez, J., Rovira, C., Davies, G.J., Wilson, K.S., and Cowtan, K.D. (2015). Privateer: software for the conformational validation of carbohydrate structures. Nat. Struct. Mol. Biol. *22*, 833–834.

Alkhovsky, S., Lenshin, S., Romashin, A., Vishnevskaya, T., Vyshemirsky, O., Bulycheva, Y., Lvov, D., and Gitelman, A. (2022). SARS-like coronaviruses in horseshoe bats (Rhinolophus spp.) in Russia, 2020. Viruses *14*. https://doi.org/10.3390/v14010113.

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology*, *37*(4), 420–423. https://doi.org/10.1038/s41587-019-0036-z

Andreano, E., Piccini, G., Licastro, D., Casalino, L., Johnson, N. V., Paciello, I., Dal Monego, S., Pantano, E., Manganaro, N., Manenti, A., Manna, R., Casa, E., Hyseni, I., Benincasa, L., Montomoli, E., Amaro, R. E., McLellan, J. S., & Rappuoli, R. (2020). SARS-CoV-2 escape *in vitro* from a highly neutralizing COVID-19 convalescent plasma. *bioRxiv : the preprint server for biology*, 2020.12.28.424451. https://doi.org/10.1101/2020.12.28.424451

Andreini, C., Bertini, I., Cavallaro, G., Holliday G.L. & Thornton, J.M. Metal ions in biological catalysis: from enzyme databases to general principles. *J Biol Inorg Chem* 13, 1205–1218 (2008).

Andrews, N. et al. Effectiveness of COVID-19 vaccines against the Omicron variant of concern. Preprint at https://doi.org/10.1101/2021.12.14.21267615 (2021).

Anthony, S.J., Gilardi, K., Menachery, V.D., Goldstein, T., Ssebide, B., Mbabazi, R., Navarrete-Macias, I., Liang, E., Wells, H., Hicks, A., et al. (2017). Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus. MBio *8*. https://doi.org/10.1128/mBio.00373-17.

Antinori, A., and Bausch-Jurken, M. (2023). The burden of COVID-19 in the immunocompromised patient: Implications for vaccination and needs for the future. J. Infect. Dis. *228*, S4–S12.

Apaa, T., Withers, A.J., Staley, C., Blanchard, A., Bennett, M., Bremner-Harrison, S., Chadwick, E.A., Hailer, F., Harrison, S.W.R., Loose, M., et al. (2023). Sarbecoviruses of British horseshoe bats; sequence variation and epidemiology. J. Gen. Virol. *104*. 10.1099/jgv.0.001859.

Arora, P., Kempf, A., Nehlmeier, I., Schulz, S.R., Jäck, H.-M., Pöhlmann, S., and Hoffmann, M. (2023). Omicron sublineage BQ.1.1 resistance to monoclonal antibodies. Lancet Infect. Dis. *23*, 22–23.

Arunachalam, P. S., Walls, A. C., Golden, N., Atyeo, C., Fischinger, S., Li, C., Aye, P., Navarro, M. J., Lai, L., Edara, V. V., Röltgen, K., Rogers, K., Shirreff, L., Ferrell, D. E., Wrenn, S., Pettie, D., Kraft, J. C., Miranda, M. C., Kepl, E., Sydeman, C., … Pulendran, B. (2021). Adjuvanting a subunit COVID-19 vaccine to induce protective immunity. *Nature*, *594*(7862), 253–258. https://doi.org/10.1038/s41586-021-03530-2.

Asarnow, D., Palovcak, E., and Cheng, Y. (2019). UCSF pyem v0. 5. Zenodo https://doi. org/10. 5281/zenodo *3576630*, 2019.

Avanzato, V. A., Matson, M. J., Seifert, S. N., Pryce, R., Williamson, B. N., Anzick, S. L., Barbian, K., Judson, S. D., Fischer, E. R., Martens, C., Bowden, T. A., de Wit, E.,

Riedo, F. X., & Munster, V. J. (2020). Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. *Cell*, *183*(7), 1901–1912.e9. https://doi.org/10.1016/j.cell.2020.10.049

Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373, 871-876 (2021).

Bailey, L.J., Sheehy, K.M., Dominik, P.K., Liang, W.G., Rui, H., Clark, M., Jaskolowski, M., Kim, Y., Deneka, D., Tang, W.-J., et al. (2018). Locking the elbow: Improved antibody fab fragments as chaperones for structure determination. J. Mol. Biol. *430*, 337–347.

Bale, J. B., Gonen, S., Liu, Y., Sheffler, W., Ellis, D., Thomas, C., Cascio, D., Yeates, T. O., Gonen, T., King, N. P., & Baker, D. (2016). Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science (New York, N.Y.)*, *353*(6297), 389–394. https://doi.org/10.1126/science.aaf8818

Barad, B. A., Echols, N., Wang, R. Y., Cheng, Y., DiMaio, F., Adams, P. D., & Fraser, J. S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature methods*, *12*(10), 943–946. https://doi.org/10.1038/nmeth.3541

Barnes, C. O., Jette, C. A., Abernathy, M. E., Dam, K. A., Esswein, S. R., Gristick, H. B., Malyutin, A. G., Sharaf, N. G., Huey-Tubman, K. E., Lee, Y. E., Robbiani, D. F., Nussenzweig, M. C., West, A. P., Jr, & Bjorkman, P. J. (2020). SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature*, *588*(7839), 682–687.

Barnes, C. O., West, A. P., Jr, Huey-Tubman, K. E., Hoffmann, M. A. G., Sharaf, N. G., Hoffman, P. R., Koranda, N., Gristick, H. B., Gaebler, C., Muecksch, F., Lorenzi, J. C. C., Finkin, S., Hägglöf, T., Hurley, A., Millard, K. G., Weisblum, Y., Schmidt, F., Hatziioannou, T., Bieniasz, P. D., Caskey, M., … Bjorkman, P. J. (2020). Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell*, *182*(4), 828–842.e16. https://doi.org/10.1016/j.cell.2020.06.025

Bassi, J. et al. Poor neutralization and rapid decay of antibodies to SARS-CoV-2 variants in vaccinated dialysis patients. Preprint at https://doi.org/10.1101/2021.10.05.21264054 (2021).

Baum, A., Fulton, B. O., Wloga, E., Copin, R., Pascal, K. E., Russo, V., Giordano, S., Lanza, K., Negron, N., Ni, M., Wei, Y., Atwal, G. S., Murphy, A. J., Stahl, N., Yancopoulos, G. D., & Kyratsous, C. A. (2020). Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science (New York, N.Y.)*, *369*(6506), 1014–1018. https://doi.org/10.1126/science.abd0831

Baum, A. et al. REGN-COV2 antibodies prevent and treat SARS-CoV-2 infection in rhesus macaques and hamsters. *Science* **370**, 1110–1115 (2020).

Brandyn West, Anna Wec, Michael Doyle, Chengzi Kaku, Pamela Hawn, Lukas Dillinger, Laura Walker NVD200 potently neutralises Omicron and its sublineages. https://invivyd.com/wp-content/uploads/2023/08/ECCMID-2023_NVD200-potently-neutralises-Omicron-and-its-sublineages-_Final-Poster.pdf.

Becker, M. M. et al. Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *Proc. Natl Acad. Sci. USA* 105, 19944–19949 (2008).
Belk, J., Deveau, L. M., Rappazzo, C. G., Walker, L. & Wec, A. PCT application WO/2021/207597: Compounds specific to coronavirus S protein and uses thereof. (Adagio Therapeutics, 2021).
Benatuil, L., Perez, J.M., Belk, J., and Hsieh, C.-M. (2010). An improved yeast transformation method for the generation of very large human antibody libraries. Protein Eng. Des. Sel. *23*, 155–159.
Bepler, T., Kelley, K., Noble, A.J., and Berger, B. (2020). Topaz-Denoise: general deep denoising models for cryoEM and cryoET. Nat. Commun. *11*, 5208.
Bergström, J. J., Xu, H., & Heyman, B. (2017). Epitope-Specific Suppression of IgG Responses by Passively Administered Specific IgG: Evidence of Epitope Masking. *Frontiers in immunology*, *8*, 238. https://doi.org/10.3389/fimmu.2017.00238
Biniossek, M.L., Nägler, D.K., Becker-Pauly, C., and Schilling, O. (2011). Proteomic identification of protease cleavage sites characterizes prime and non-prime specificity of cysteine cathepsins B, L, and S. J. Proteome Res. *10*, 5363–5373.
Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
Bolles, M., Deming, D., Long, K., Agnihothram, S., Whitmore, A., Ferris, M., Funkhouser, W., Gralinski, L., Totura, A., Heise, M., & Baric, R. S. (2011). A double-inactivated severe acute respiratory syndrome coronavirus vaccine provides incomplete protection in mice and induces increased eosinophilic proinflammatory pulmonary response upon challenge. *Journal of virology*, *85*(23), 12201–12215. https://doi.org/10.1128/JVI.06048-11
Boni, M. F. et al. Evolutionary origins of the SARS-CoV-2 Sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat. Microbiol.* 5, 1408–1417 (2020).
Bosch, B. J., Huizinga, E. G. & De Groot, R. J. Human coronaviruses OC43 and HKU1 bind to 9-*O*-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl Acad. Sci. USA* 116, 2681–2690 (2019).
Boudewijns, R., Thibaut, H. J., Kaptein, S. J. F., Li, R., Vergote, V., Seldeslachts, L., Van Weyenbergh, J., De Keyzer, C., Bervoets, L., Sharma, S., Liesenborghs, L., Ma, J., Jansen, S., Van Looveren, D., Vercruysse, T., Wang, X., Jochmans, D., Martens, E., Roose, K., De Vlieger, D., … Dallmeier, K. (2020). STAT2 signaling restricts viral dissemination but drives severe pneumonia in SARS-CoV-2 infected hamsters. *Nature communications*, *11*(1), 5838. https://doi.org/10.1038/s41467-020-19684-y
Bournazos, S., Wang, T. T., & Ravetch, J. V. (2016). The Role and Function of Fcγ Receptors on Myeloid Cells. *Microbiology spectrum*, *4*(6), 10.1128/microbiolspec.MCHD-0045-2016. https://doi.org/10.1128/microbiolspec.MCHD-0045-2016
Bournazos, S., Corti, D., Virgin, H. W., & Ravetch, J. V. (2020). Fc-optimized antibodies elicit CD8 immunity to viral respiratory infection. *Nature*, *588*(7838), 485–490. https://doi.org/10.1038/s41586-020-2838-z
Bowen, J.E., Park, Y.J., Stewart, C., Brown, J.T., Sharkey, W.K., Walls, A.C., Joshi, A., Sprouse, K.R., McCallum, M., Tortorici, M.A., et al. (2022). SARS-CoV-2 spike

conformation determines plasma neutralizing activity elicited by a wide panel of human vaccines. Sci. Immunol. *7*, eadf1421. https://doi.org/10.1126/sciimmunol.adf1421.

Bowen, J.E., Addetia, A., Dang, H.V., Stewart, C., Brown, J.T., Sharkey, W.K., Sprouse, K.R., Walls, A.C., Mazzitelli, I.G., Logue, J.K., et al. (2022). Omicron spike function and neutralizing activity elicited by a comprehensive panel of vaccines. Science *377*, 890–894.

Boyoglu-Barnum, S., Ellis, D., Gillespie, R. A., Hutchinson, G. B., Park, Y. J., Moin, S. M., Acton, O. J., Ravichandran, R., Murphy, M., Pettie, D., Matheson, N., Carter, L., Creanga, A., Watson, M. J., Kephart, S., Ataca, S., Vaile, J. R., Ueda, G., Crank, M. C., Stewart, L., … Kanekiyo, M. (2021). Quadrivalent influenza nanoparticle vaccines induce broad protection. *Nature*, *592*(7855), 623–628. https://doi.org/10.1038/s41586-021-03365-x

Brouwer, P. J. M., Caniels, T. G., van der Straten, K., Snitselaar, J. L., Aldon, Y., Bangaru, S., Torres, J. L., Okba, N. M. A., Claireaux, M., Kerster, G., Bentlage, A. E. H., van Haaren, M. M., Guerra, D., Burger, J. A., Schermer, E. E., Verheul, K. D., van der Velde, N., van der Kooi, A., van Schooten, J., van Breemen, M. J., … van Gils, M. J. (2020). Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science (New York, N.Y.)*, *369*(6504), 643–650. https://doi.org/10.1126/science.abc5902

Buchholz, U.J., Finke, S., and Conzelmann, K.K. (1999). Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. J. Virol. *73*, 251–259.

Cameroni, E., Bowen, J.E., Rosen, L.E., Saliba, C., Zepeda, S.K., Culap, K., Pinto, D., VanBlargan, L.A., De Marco, A., di Iulio, J., et al. (2021). Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. Nature *602*, 664–670. https://doi.org/10.1038/d41586-021- 03825-4.

Cantuti-Castelvetri, L., Ojha, R., Pedro, L. D., Djannatian, M., Franz, J., Kuivanen, S., van der Meer, F., Kallio, K., Kaya, T., Anastasina, M., Smura, T., Levanov, L., Szirovicza, L., Tobi, A., Kallio-Kokko, H., Österlund, P., Joensuu, M., Meunier, F. A., Butcher, S. J., Winkler, M. S., … Simons, M. (2020). Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science (New York, N.Y.)*, *370*(6518), 856–860. https://doi.org/10.1126/science.abd2985

Cao, Y. et al. B.1.1.529 escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature*https://doi.org/10.1038/s41586-021-04385-3 (2021).

Cao, Y., Yisimayi, A., Jian, F., Song, W., Xiao, T., Wang, L., Du, S., Wang, J., Li, Q., Chen, X., et al. (2022). BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. Nature *608*, 593–602.

Cao, Y., Wang, J., Jian, F., Xiao, T., Song, W., Yisimayi, A., Huang, W., Li, Q., Wang, P., An, R., et al. (2022). Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature *602*, 657–663.

Cao, Y., Jian, F., Zhang, Z., Yisimayi, A., Hao, X., Bao, L., Yuan, F., Yu, Y., Du, S., Wang, J., et al. (2022). Rational identification of potent and broad sarbecovirus-neutralizing antibody cocktails from SARS convalescents. Cell Rep. *41*, 111845.

Cao, Y., Jian, F., Wang, J., Yu, Y., Song, W., Yisimayi, A., Wang, J., An, R., Chen, X., Zhang, N., et al. (2023). Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. Nature *614*, 521–529.

Casañal, A., Lohkamp, B., & Emsley, P. (2020). Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein science : a publication of the Protein Society*, *29*(4), 1069–1078. https://doi.org/10.1002/pro.3791

Case, D.A., Metin Aktulga, H., Belfon, K., Ben-Shalom, I., Brozell, S.R., Cerutti, D.S., Cheatham, T.E., III, Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., et al. (2021). Amber 2021 (University of California, San Francisco).

Case, J. B., Rothlauf, P. W., Chen, R. E., Liu, Z., Zhao, H., Kim, A. S., Bloyet, L. M., Zeng, Q., Tahan, S., Droit, L., Ilagan, M. X. G., Tartell, M. A., Amarasinghe, G., Henderson, J. P., Miersch, S., Ustav, M., Sidhu, S., Virgin, H. W., Wang, D., Ding, S., … Whelan, S. P. J. (2020). Neutralizing Antibody and Soluble ACE2 Inhibition of a Replication-Competent VSV-SARS-CoV-2 and a Clinical Isolate of SARS-CoV-2. *Cell host & microbe*, *28*(3), 475–485.e5. https://doi.org/10.1016/j.chom.2020.06.021

Case, J. B., Rothlauf, P. W., Chen, R. E., Kafai, N. M., Fox, J. M., Smith, B. K., Shrihari, S., McCune, B. T., Harvey, I. B., Keeler, S. P., Bloyet, L. M., Zhao, H., Ma, M., Adams, L. J., Winkler, E. S., Holtzman, M. J., Fremont, D. H., Whelan, S. P. J., & Diamond, M. S. (2020). Replication-Competent Vesicular Stomatitis Virus Vaccine Vector Protects against SARS-CoV-2-Mediated Pathogenesis in Mice. *Cell host & microbe*, *28*(3), 465–474.e4. https://doi.org/10.1016/j.chom.2020.07.018

Cathcart, A. L., et al. (2022). "The dual function monoclonal antibodies VIR-7831 and VIR-7832 demonstrate potent in vitro and in vivo activity against SARS-CoV-2." bioRxiv: 2021.2003.2009.434607.

Cele, S. et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature*https:///doi.org/10.1038/s41586-021-04387-1 (2021).

Challen, R., Brooks-Pollock, E., Read, J. M., Dyson, L., Tsaneva-Atanasova, K., & Danon, L. (2021). Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ (Clinical research ed.)*, *372*, n579. https://doi.org/10.1136/bmj.n579

Chao, C.W., Sprouse, K.R., Miranda, M.C., Catanzaro, N.J., Hubbard, M.L., Addetia, A., Stewart, C., Brown, J.T., Dosey, A., Valdez, A., et al. (2024). Protein nanoparticle vaccines induce potent neutralizing antibody responses against MERS-CoV. bioRxiv. https://doi.org/10.1101/2024.03.13.584735.

Chen, J., Yang, X., Si, H., Gong, Q., Que, T., Li, J., Li, Y., Wu, C., Zhang, W., Chen, Y., et al. (2023). A bat MERS-like coronavirus circulates in pangolins and utilizes human DPP4 and host proteases for cell entry. Cell *186*, 850–863.e16.

Chen, P. et al. SARS-CoV-2 neutralizing antibody LY-CoV555 in outpatients with Covid-19. *N. Engl. J. Med.* **384**, 229–237 (2021).

Chen, R. E. et al. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**, 717–726 (2021).

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting

and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy *135*, 24–35.
Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.
Chen, Y., Lu, S., Jia, H., Deng, Y., Zhou, J., Huang, B., Yu, Y., Lan, J., Wang, W., Lou, Y., Qin, K., & Tan, W. (2017). A novel neutralizing monoclonal antibody targeting the N-terminal domain of the MERS-CoV spike protein. *Emerging microbes & infections*, *6*(5), e37. https://doi.org/10.1038/emi.2017.18
Chi, X., Yan, R., Zhang, J., Zhang, G., Zhang, Y., Hao, M., Zhang, Z., Fan, P., Dong, Y., Yang, Y., Chen, Z., Guo, Y., Zhang, J., Li, Y., Song, X., Chen, Y., Xia, L., Fu, L., Hou, L., Xu, J., … Chen, W. (2020). A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science (New York, N.Y.)*, *369*(6504), 650–655. https://doi.org/10.1126/science.abc6952
Chiodo, F. et al. (2020) Novel ACE2-Independent Carbohydrate-Binding of SARS-CoV-2 Spike Protein to Host Lectins and Lung Microbiota. bioRxiv 2020.05.13.092478; doi: https://doi.org/10.1101/2020.05.13.092478
Choi, B., Choudhary, M. C., Regan, J., Sparks, J. A., Padera, R. F., Qiu, X., Solomon, I. H., Kuo, H. H., Boucau, J., Bowman, K., Adhikari, U. D., Winkler, M. L., Mueller, A. A., Hsu, T. Y., Desjardins, M., Baden, L. R., Chan, B. T., Walker, B. D., Lichterfeld, M., Brigl, M., … Li, J. Z. (2020). Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *The New England journal of medicine*, *383*(23), 2291–2293. https://doi.org/10.1056/NEJMc2031364
Chojnowski, G., Sobolev, E., Heuser, P. & Lamzin, V.S. The accuracy of protein models automatically built into cryo-EM maps with ARP/wARP. *Acta Cryst.* D77, 142-150 (2021).
Clausen, T. M., Sandoval, D. R., Spliid, C. B., Pihl, J., Perrett, H. R., Painter, C. D., Narayanan, A., Majowicz, S. A., Kwong, E. M., McVicar, R. N., Thacker, B. E., Glass, C. A., Yang, Z., Torres, J. L., Golden, G. J., Bartels, P. L., Porell, R. N., Garretson, A. F., Laubach, L., Feldman, J., … Esko, J. D. (2020). SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. *Cell*, *183*(4), 1043–1057.e15. https://doi.org/10.1016/j.cell.2020.09.033
Cohen, A.A., Gnanapragasam, P.N.P., Lee, Y.E., Hoffman, P.R., Ou, S., Kakutani, L.M., Keeffe, J.R., Wu, H.J., Howarth, M., West, A.P., et al. (2021). Mosaic nanoparticles elicit cross-reactive immune responses to zoonotic coronaviruses in mice. Science *371*, 735–741.
Cohen, A.A., van Doremalen, N., Greaney, A.J., Andersen, H., Sharma, A., Starr, T.N., Keeffe, J.R., Fan, C., Schulz, J.E., Gnanapragasam, P.N.P., et al. (2022). Mosaic RBD nanoparticles protect against challenge by diverse sarbecoviruses in animal models. Science *377*, eabq0839.
Cohen, A. A., Yang, Z., Gnanapragasam, P. N. P., Ou, S., Dam, K. A., Wang, H., & Bjorkman, P. J. (2021). Construction, characterization, and immunization of nanoparticles that display a diverse array of influenza HA trimers. *PloS one*, *16*(3), e0247963. https://doi.org/10.1371/journal.pone.0247963

Collier, D. A., De Marco, A., Ferreira, I. A. T. M., Meng, B., Datir, R. P., Walls, A. C., Kemp, S. A., Bassi, J., Pinto, D., Silacci-Fregni, C., Bianchi, S., Tortorici, M. A., Bowen, J., Culap, K., Jaconi, S., Cameroni, E., Snell, G., Pizzuto, M. S., Pellanda, A. F., Garzoni, C., … Gupta, R. K. (2021). Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature*, *593*(7857), 136–141. https://doi.org/10.1038/s41586-021-03412-7

Consortium, Chinese SARS Molecular Epidemiology (2004). Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science *303*, 1666–1669.

Copin, R. et al. The monoclonal antibody combination REGEN-COV protects against SARS-CoV-2 mutational escape in preclinical and human studies. *Cell* **184**, 3949–3961 (2021).

Corbett, K. S., Edwards, D. K., Leist, S. R., Abiona, O. M., Boyoglu-Barnum, S., Gillespie, R. A., Himansu, S., Schäfer, A., Ziwawo, C. T., DiPiazza, A. T., Dinnon, K. H., Elbashir, S. M., Shaw, C. A., Woods, A., Fritch, E. J., Martinez, D. R., Bock, K. W., Minai, M., Nagata, B. M., Hutchinson, G. B., … Graham, B. S. (2020). SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature*, *586*(7830), 567–571. https://doi.org/10.1038/s41586-020-2622-0

Corbett, K. S., Flynn, B., Foulds, K. E., Francica, J. R., Boyoglu-Barnum, S., Werner, A. P., Flach, B., O'Connell, S., Bock, K. W., Minai, M., Nagata, B. M., Andersen, H., Martinez, D. R., Noe, A. T., Douek, N., Donaldson, M. M., Nji, N. N., Alvarado, G. S., Edwards, D. K., Flebbe, D. R., … Graham, B. S. (2020). Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *The New England journal of medicine*, *383*(16), 1544–1555. https://doi.org/10.1056/NEJMoa2024671

Corey, L., Beyrer, C., Cohen, M. S., Michael, N. L., Bedford, T., & Rolland, M. (2021). SARS-CoV-2 Variants in Patients with Immunosuppression. *The New England journal of medicine*, *385*(6), 562–566. https://doi.org/10.1056/NEJMsb2104756

Corman, V.M., Kallies, R., Philipps, H., Göpner, G., Müller, M.A., Eckerle, I., Brünink, S., Drosten, C., and Drexler, J.F. (2014). Characterization of a novel betacoronavirus related to middle East respiratory syndrome coronavirus in European hedgehogs. J. Virol. *88*, 717–724.

Corti, D., Purcell, L. A., Snell, G. & Veesler, D. Tackling COVID-19 with neutralizing monoclonal antibodies. *Cell* **184**, 3086–3108 (2021).

Crawford, K.H.D., and Bloom, J.D. (2019). alignparse: a Python package for parsing complex features from high-throughput long-read sequencing. J. Open Source Softw. *4*, 4. https://doi.org/10.21105/joss.01915.

Crawford, K.H.D., Eguia, R., Dingens, A.S., Loes, A.N., Malone, K.D., Wolf, C.R., Chu, H.Y., Tortorici, M.A., Veesler, D., Murphy, M., et al. (2020). Protocol and reagents for pseudotyping lentiviral particles with SARS-CoV-2 spike protein for neutralization assays. Viruses *12*. https://doi.org/10.3390/v12050513.

Croll T. I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta crystallographica. Section D, Structural biology*, *74*(Pt 6), 519–530. https://doi.org/10.1107/S2059798318002425

Crook, J.M., Murphy, I., Carter, D.P., Pullan, S.T., Carroll, M., Vipond, R., Cunningham, A.A., and Bell, D. (2021). Metagenomic identification of a new sarbecovirus from horseshoe bats in Europe. Sci. Rep. *11*, 14723.

Cui, J., Li, F. & Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 17, 181–192 (2019).

Dacon, C., Tucker, C., Peng, L., Lee, C.-C.D., Lin, T.-H., Yuan, M., Cong, Y., Wang, L., Purser, L., Williams, J.K., et al. (2022). Broadly neutralizing antibodies target the coronavirus fusion peptide. Science *377*, 728–735.

Dacon, C., Peng, L., Lin, T.-H., Tucker, C., Lee, C.-C.D., Cong, Y., Wang, L., Purser, L., Cooper, A.J.R., Williams, J.K., et al. (2023). Rare, convergent antibodies targeting the stem helix broadly neutralize diverse betacoronaviruses. Cell Host Microbe *31*, 97-111.e12.

Dadonaite, B., Brown, J., McMahon, T.E., Farrell, A.G., Figgins, M.D., Asarnow, D., Stewart, C., Lee, J., Logue, J., Bedford, T., et al. (2024). Spike deep mutational scanning helps predict success of SARS-CoV-2 clades. Nature *631*, 617–626.

Daly, J. L., Simonetti, B., Klein, K., Chen, K. E., Williamson, M. K., Antón-Plágaro, C., Shoemark, D. K., Simón-Gracia, L., Bauer, M., Hollandi, R., Greber, U. F., Horvath, P., Sessions, R. B., Helenius, A., Hiscox, J. A., Teesalu, T., Matthews, D. A., Davidson, A. D., Collins, B. M., Cullen, P. J., … Yamauchi, Y. (2020). Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science (New York, N.Y.)*, *370*(6518), 861–865. https://doi.org/10.1126/science.abd3072

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*. 10.1093/gigascience/giab008.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Dauparas, J., Anishchenko, I., Bennett, N., Bai, H., Ragotte, R.J., Milles, L.F., Wicky, B.I.M., Courbet, A., de Haas, R.J., Bethel, N., et al. (2022). Robust deep learning–based protein sequence design using ProteinMPNN. Science *378*, 49–56.

Davies, N. G., Abbott, S., Barnard, R. C., Jarvis, C. I., Kucharski, A. J., Munday, J. D., Pearson, C. A. B., Russell, T. W., Tully, D. C., Washburne, A. D., Wenseleers, T., Gimma, A., Waites, W., Wong, K. L. M., van Zandvoort, K., Silverman, J. D., CMMID COVID-19 Working Group, COVID-19 Genomics UK (COG-UK) Consortium, Diaz-Ordaz, K., Keogh, R., … Edmunds, W. J. (2021). Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science (New York, N.Y.)*, *372*(6538), eabg3055. https://doi.org/10.1126/science.abg3055

Davies, N. G., Jarvis, C. I., CMMID COVID-19 Working Group, Edmunds, W. J., Jewell, N. P., Diaz-Ordaz, K., & Keogh, R. H. (2021). Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature*, *593*(7858), 270–274. https://doi.org/10.1038/s41586-021-03426-1

Delaune, D. et al. A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *Nat. Commun.* 12, 6563 (2021).

Dejnirattisai, W., Zhou, D., Ginn, H. M., Duyvesteyn, H. M. E., Supasa, P., Case, J. B., Zhao, Y., Walter, T. S., Mentzer, A. J., Liu, C., Wang, B., Paesen, G. C., Slon-Campos, J., López-Camacho, C., Kafai, N. M., Bailey, A. L., Chen, R. E., Ying,

B., Thompson, C., Bolton, J., … Screaton, G. R. (2021). The antigenic anatomy of SARS-CoV-2 receptor binding domain. *Cell*, *184*(8), 2183–2200.e22. https://doi.org/10.1016/j.cell.2021.02.032

Demogines, A., Farzan, M. & Sawyer, S. L. Evidence for ACE2-utilizing coronaviruses (CoVs) related to severe acute respiratory syndrome CoV in bats. *J. Virol.* 86, 6350–6353 (2012).

Deng, X., Garcia-Knight, M. A., Khalid, M. M., Servellita, V., Wang, C., Morris, M. K., Sotomayor-González, A., Glasner, D. R., Reyes, K. R., Gliwa, A. S., Reddy, N. P., Sanchez San Martin, C., Federman, S., Cheng, J., Balcerek, J., Taylor, J., Streithorst, J. A., Miller, S., Sreekumar, B., Chen, P. Y., … Chiu, C. Y. (2021). Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell*, *184*(13), 3426–3437.e8. https://doi.org/10.1016/j.cell.2021.04.025

de Wit, E., Feldmann, F., Cronin, J., Jordan, R., Okumura, A., Thomas, T., Scott, D., Cihlar, T., and Feldmann, H. (2020). Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection. Proc. Natl. Acad. Sci. U. S. A. *117*, 6771–6776.

di Iulio, J., Bartha, I., Wong, E.H.M., Yu, H.-C., Lavrenko, V., Yang, D., Jung, I., Hicks, M.A., Shah, N., Kirkness, E.F., et al. (2018). The human noncoding genome defined by genetic diversity. Nat. Genet. *50*, 333–337.

DiMaio, F., Song, Y., Li, X., Brunner, M. J., Xu, C., Conticello, V., Egelman, E., Marlovits, T., Cheng, Y., & Baker, D. (2015). Atomic-accuracy models from 4.5-Å cryo-electron microscopy data with density-guided iterative local refinement. *Nature methods*, *12*(4), 361–365. https://doi.org/10.1038/nmeth.3286

Dinnon, K. H. III et al. A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures. *Nature* 586, 560–566 (2020).

Dong, J., Zost, S. J., Greaney, A. J., Starr, T. N., Dingens, A. S., Chen, E. C., Chen, R. E., Case, J. B., Sutton, R. E., Gilchuk, P., Rodriguez, J., Armstrong, E., Gainza, C., Nargi, R. S., Binshtein, E., Xie, X., Zhang, X., Shi, P. Y., Logue, J., Weston, S., … Crowe, J. E., Jr (2021). Genetic and structural basis for SARS-CoV-2 variant neutralization by a two-antibody cocktail. *Nature microbiology*, *6*(10), 1233–1244. https://doi.org/10.1038/s41564-021-00972-2

Drexler, J.F., Gloza-Rausch, F., Glende, J., Corman, V.M., Muth, D., Goettsche, M., Seebens, A., Niedrig, M., Pfefferle, S., Yordanov, S., et al. (2010). Genomic characterization of severe acute respiratory syndrome related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. J. Virol. *84*, 11336–11349.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. *348*, 1967–1976.

Eastman, P., Galvelis, R., Peláez, R.P., Abreu, C.R.A., Farr, S.E., Gallicchio, E., Gorenko, A., Henry, M.M., Hu, F., Huang, J., et al. (2023). OpenMM 8: Molecular dynamics simulation with machine learning potentials. 10.48550/ARXIV.2310.03121.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Eick, G. N., Bridgham, J. T., Anderson, D. P., Harms, M. J. & Thornton, J. W. Robustness of reconstructed ancestral protein functions to statistical uncertainty. *Mol. Biol. Evol.* 34, 247–261 (2017).
Ellis, D., Lederhofer, J., Acton, O.J. *et al.* Structure-based design of stabilized recombinant influenza neuraminidase tetramers. *Nat Commun* **13**, 1825 (2022). https://doi.org/10.1038/s41467-022-29416-z
Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst.* D60, 2126-2132 (2004).
Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.
Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. PLoS One *3*, e3647.
Erasmus, J. H., Khandhar, A. P., O'Connor, M. A., Walls, A. C., Hemann, E. A., Murapa, P., Archer, J., Leventhal, S., Fuller, J. T., Lewis, T. B., Draves, K. E., Randall, S., Guerriero, K. A., Duthie, M. S., Carter, D., Reed, S. G., Hawman, D. W., Feldmann, H., Gale, M., Jr, Veesler, D., … Fuller, D. H. (2020). An *Alphavirus*-derived replicon RNA vaccine induces SARS-CoV-2 neutralizing antibody and T cell responses in mice and nonhuman primates. *Science translational medicine*, *12*(555), eabc9396. https://doi.org/10.1126/scitranslmed.abc9396
Evans, R.A., Dube, S., Lu, Y., Yates, M., Arnetorp, S., Barnes, E., Bell, S., Carty, L., Evans, K., Graham, S., et al. (2023). Impact of COVID-19 on immunocompromised populations during the Omicron era: insights from the observational population-based INFORM study. Lancet Reg. Health Eur. *35*, 100747.
Evans, T.S., Tan, C.W., Aung, O., Phyu, S., Lin, H., Coffey, L.L., Toe, A.T., Aung, P., Aung, T.H., Aung, N.T., et al. (2023). Exposure to diverse sarbe- coviruses indicates frequent zoonotic spillover in human communities in- teracting with wildlife. Int. J. Infect. Dis. *131*, 57–64. https://doi.org/10. 1016/j.ijid.2023.02.015.
Evrard, G.X., Langer, G.G., Perrakis, A. & Lamzin, V.S. Assessment of automatic ligand building in ARP/wARP. *Acta Cryst.* D63, 108-117 (2007).
Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D. D. S., Mishra, S., Crispim, M. A. E., Sales, F. C. S., Hawryluk, I., McCrone, J. T., Hulswit, R. J. G., Franco, L. A. M., Ramundo, M. S., de Jesus, J. G., Andrade, P. S., Coletti, T. M., Ferreira, G. M., Silva, C. A. M., Manuli, E. R., Pereira, R. H. M., … Sabino, E. C. (2021). Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science (New York, N.Y.)*, *372*(6544), 815–821. https://doi.org/10.1126/science.abh2644
Fass, J., Sivak, D., Crooks, G., Beauchamp, K., Leimkuhler, B., and Chodera, J. (2018). Quantifying configuration-sampling error in Langevin simulations of complex molecular systems. Entropy (Basel) *20*, 318.
FDA. Fact sheet for healthcare providers: emergency use authorization (EUA) of sotrovimab. https://www.fda.gov/media/149534/download (2021).
FDA (2024). Fact sheet for healthcare providers: emergency use authorization of Pemgarda (pemivibart). www.fda.gov/media/177067/download.
Fenwick, C., Turelli, P., Duhoo, Y., Lau, K., Herate, C., Marlin, R., Lamrayah, M., Campos, J., Esteves-Leuenberger, L., Farina, A., et al. (2023). Broadly potent

anti-SARS-CoV-2 antibody shares 93% of epitope with ACE2 and provides full protection in monkeys. J. Infect. *87*, 524–537.
Fischer, W., Giorgi, E. E., Chakraborty, S., Nguyen, K., Bhattacharya, T., Theiler, J., Goloboff, P. A., Yoon, H., Abfalterer, W., Foley, B. T., Tegally, H., San, J. E., de Oliveira, T., Network for Genomic Surveillance in South Africa (NGS-SA), Gnanakaran, S., & Korber, B. (2021). HIV-1 and SARS-CoV-2: Patterns in the evolution of two pandemic pathogens. *Cell host & microbe*, *29*(7), 1093–1110. https://doi.org/10.1016/j.chom.2021.05.012
Fraser, B.J., Beldar, S., Seitova, A., Hutchinson, A., Mannar, D., Li, Y., Kwon, D., Tan, R., Wilson, R.P., Leopold, K., et al. (2022). Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. Nat. Chem. Biol. *18*, 963–971.
Frenz, B., Walls, A., Egelman, E., Veesler, D. & DiMaio, F. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nat. Methods* 14, 797–800 (2017).
Frenz, B., Rämisch, S., Borst, A. J., Walls, A. C., Adolf-Bryfogle, J., Schief, W. R., Veesler, D., & DiMaio, F. (2019). Automatically Fixing Errors in Glycoprotein Structures with Rosetta. *Structure (London, England : 1993)*, *27*(1), 134–139.e3. https://doi.org/10.1016/j.str.2018.09.006
Freitas, A.R.R., Beckedorff, O.A., de Go´es Cavalcanti, L.P., Siqueira, A.M., de Castro, D.B., da Costa, C.F., Lemos, D.R.Q., and Barros, E.N.C. (2021). The emergence of novel SARS-CoV-2 variant P.1 in Amazonas (Brazil) was temporally associated with a change in the age and sex profile of COVID-19 mortality: a population based ecological study. The Lancet Regional Health – Americas *1*. https://doi.org/10. 1016/j.lana.2021.100021.
Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics *28*, 3150–3152.
Fuerst, T.R., Niles, E.G., Studier, F.W., and Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. *83*, 8122–8126.
Fujino, T., Nomoto, H., Kutsuna, S., Ujiie, M., Suzuki, T., Sato, R., Fujimoto, T., Kuroda, M., Wakita, T., and Ohmagari, N. (2021). Novel SARS-CoV-2 variant in travelers from brazil to Japan. Emerg. Infect. Dis. *27*, 1243–1245.
Gao, C., Zeng, J., Jia, N., Stavenhagen, K., Matsumoto, Y., Zhang, H., Li, J., Hume, A. J., Mühlberger, E., van Die, I., Kwan, J., Tantisira, K., Emili, A., & Cummings, R. D. (2020). SARS-CoV-2 Spike Protein Interacts with Multiple Innate Immune Receptors. *bioRxiv : the preprint server for biology*, 2020.07.29.227462.
Garvie, C.W. et al. Structure of PDE3A-SLFN12 complex reveals requirements for activation of SLFN12 RNase. *Nat. Commun.* 12, 4375 (2021).
Ge, X.Y., Li, J.L., Yang, X.L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., et al. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature *503*, 535–538.
Gelderblom HR. Structure and Classification of Viruses. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at

Galveston; 1996. Chapter 41. Available from: https://www.ncbi.nlm.nih.gov/books/NBK8174/
Gentile, I., Foggia, M., Silvitelli, M., Sardanelli, A., Cattaneo, L., Viceconte, G., and Federico II COVID team (2023). Early combination with remdesivir, nirmatrelvir/ritonavir and sotrovimab for the treatment of COVID-19 in immunocompromised hosts. bioRxiv. 10.1101/2023.09.07.23295202.
Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318.
Giroglou, T., Cinatl, J., Jr, Rabenau, H., Drosten, C., Schwalbe, H., Doerr, H. W., & von Laer, D. (2004). Retroviral vectors pseudotyped with severe acute respiratory syndrome coronavirus S protein. *Journal of virology*, *78*(17), 9007–9015. https://doi.org/10.1128/JVI.78.17.9007-9015.2004
Goddard, T. D., Huang, C. C., & Ferrin, T. E. (2007). Visualizing density maps with UCSF Chimera. *Journal of structural biology*, *157*(1), 281–287. https://doi.org/10.1016/j.jsb.2006.06.010
Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25.
Gorrec, F. (2009). The MORPHEUS protein crystallization screen. J. Appl. Crystallogr. *42*, 1035–1042.
Gottlieb, R. L., Nirula, A., Chen, P., Boscia, J., Heller, B., Morris, J., Huhn, G., Cardona, J., Mocherla, B., Stosor, V., Shawa, I., Kumar, P., Adams, A. C., Van Naarden, J., Custer, K. L., Durante, M., Oakley, G., Schade, A. E., Holzer, T. R., Ebert, P. J., … Skovronsky, D. M. (2021). Effect of Bamlanivimab as Monotherapy or in Combination With Etesevimab on Viral Load in Patients With Mild to Moderate COVID-19: A Randomized Clinical Trial. *JAMA*, *325*(7), 632–644. https://doi.org/10.1001/jama.2021.0202
Graham, R. *et al*. (2013) A decade after SARS: strategies for controlling emerging coronaviruses. *Nature*. 11: 836-848.
Greaney, A. J., Starr, T. N., Gilchuk, P., Zost, S. J., Binshtein, E., Loes, A. N., Hilton, S. K., Huddleston, J., Eguia, R., Crawford, K. H. D., Dingens, A. S., Nargi, R. S., Sutton, R. E., Suryadevara, N., Rothlauf, P. W., Liu, Z., Whelan, S. P. J., Carnahan, R. H., Crowe, J. E., Jr, & Bloom, J. D. (2021). Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell host & microbe*, *29*(1), 44–57.e9. https://doi.org/10.1016/j.chom.2020.11.007
Greaney, A.J., Loes, A.N., Gentles, L.E., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., and Bloom, J.D. (2021). Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection. Sci. Transl. Med. *13*. https://doi.org/10.1126/scitranslmed.abi9915.
Greaney, A.J., Starr, T.N., Eguia, R.T., Loes, A.N., Khan, K., Karim, F., Cele, S., Bowen, J.E., Logue, J.K., Corti, D., et al. (2021). A SARS- CoV-2 variant elicits an antibody response with a shifted immunodominance hierarchy. Preprint at bioRxiv. https://doi.org/10.1101/2021.10. 12.464114.

Greaney, A. J., Loes, A. N., Crawford, K. H. D., Starr, T. N., Malone, K. D., Chu, H. Y., & Bloom, J. D. (2021). Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell host & microbe*, *29*(3), 463–476.e6. https://doi.org/10.1016/j.chom.2021.02.003

Greer, J. Three-dimensional pattern recognition: An approach to automated interpretation of electron density maps of proteins. *J. Mol. Bio.* 82, 279-301 (1974).

Group, A.-T. L.-C. S. et al. A neutralizing monoclonal antibody for hospitalized patients with Covid-19. *N. Engl. J. Med.* **384**, 905–914 (2021).

Guan, Y. *et al.* (2003). Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China. *Science.* **302**:276-278.

Gui *et al*. (2017). Electron microscopy studies of the coronavirus ribonucleoprotein complex. *Protein Cell*.**8(3)**:219-224.

Gupta, A. et al. Early treatment for Covid-19 with SARS-CoV-2 neutralizing antibody sotrovimab. *N. Engl. J. Med.***385**, 1941–1950 (2021).

Guo, H., Hu, B.J., Yang, X.L., Zeng, L.P., Li, B., Ouyang, S., and Shi, Z.L. (2020). Evolutionary arms race between virus and host drives genetic di- versity in bat severe acute respiratory syndrome-related coronavirus spike genes. J. Virol. *94*. https://doi.org/10.1128/JVI.00902-20.

Guo, H. et al. Identification of a novel lineage bat SARS-related coronaviruses that use bat ACE2 receptor. *Emerg. Microbes Infect.* 10, 1507–1514 (2021).

Haagmans, B.L., Al Dhahiry, S.H., Reusken, C.B., Raj, V.S., Galiano, M., Myers, R., Godeke, G.J., Jonges, M., Farag, E., Diab, A., et al. (2014). Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. Lancet Infect. Dis. *14*, 140–145.

Hansen, J., Baum, A., Pascal, K. E., Russo, V., Giordano, S., Wloga, E., Fulton, B. O., Yan, Y., Koon, K., Patel, K., Chung, K. M., Hermann, A., Ullman, E., Cruz, J., Rafique, A., Huang, T., Fairhurst, J., Libertiny, C., Malbec, M., Lee, W. Y., … Kyratsous, C. A. (2020). Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science (New York, N.Y.)*, *369*(6506), 1010–1014. https://doi.org/10.1126/science.abd0827

Hanson-Smith, V., Kolaczkowski, B. & Thornton, J. W. Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. *Mol. Biol. Evol.* 27, 1988–1999 (2010).

Hao, J. et al. Synthesis and pharmacological characterization of 2-(2,6-dichlorophenyl)-1-((1S,3R)-5-(3-hydroxy-3-methylbutyl)-3-(hydroxymethyl)-1-methyl-3,4-dihydroisoquinolin-2(1H)-yl)ethan-1-one (LY3154207), a potent, subtype selective, and orally available positive allosteric modulator of the human dopamine D1 receptor. *J. Med. Chem*. 62, 8711-8732 (2019).

Harbison, A.M., Fogarty, C.A., Phung, T.K., Satheesan, A., Schulz, B.L., and Fadda, E. (2022). Fine-tuning the spike: role of the nature and topol- ogy of the glycan shield in the structure and dynamics of the SARS-CoV- 2 S. Chem. Chem. Sci. *13*, 386–395.

Hassan, A. O., Kafai, N. M., Dmitriev, I. P., Fox, J. M., Smith, B. K., Harvey, I. B., Chen, R. E., Winkler, E. S., Wessel, A. W., Case, J. B., Kashentseva, E., McCune, B. T., Bailey, A. L., Zhao, H., VanBlargan, L. A., Dai, Y. N., Ma, M., Adams, L. J., Shrihari, S., Danis, J. E., … Diamond, M. S. (2020). A Single-Dose Intranasal ChAd Vaccine Protects Upper and Lower Respiratory Tracts against SARS-CoV-2. *Cell*, *183*(1), 169–184.e13. https://doi.org/10.1016/j.cell.2020.08.026

Hauser, K., Essuman, B., He, Y., Coutsias, E., Garcia-Diaz, M., and Simmerling, C. (2016). A human transcription factor in search mode. Nucleic Acids Res. *44*, 63–74.

He, B. et al. Identification of diverse alphacoronaviruses and genomic characterization of a novel severe acute respiratory syndrome-like coronavirus from bats in China. *J. Virol.* 88, 7070–7082 (2014).

He, J. & Huang, S.Y. Full-length de novo protein structure determination from cryo-EM maps using deep learning. *Bioinformatics* 37, 3480-3490 (2021).

He, W.-T., Ji, X., He, W., Dellicour, S., Wang, S., Li, G., Zhang, L., Gilbert, M., Zhu, H., Xing, G., et al. (2020). Genomic Epidemiology, Evolution, and Transmission Dynamics of Porcine Deltacoronavirus. Mol. Biol. Evol. *37*, 2641–2654.

Hodcroft, E.B., Zuber, M., Nadeau, S. *et al.* Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**, 707–712 (2021). https://doi.org/10.1038/s41586-021-03677-y

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kru€ger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell *181*, 271–280.e8.

Hoffman, M., Kleine-Weber, H., and Pohlmann, S. (2020) A Multibasic Cleavage Site in the Spike protein of SARS-CoV-2 Is Essential for infection of Human Lung cells. *Mol. Cell.* **78**: 779-784.

Hoffmann, M., Mösbauer, K., Hofmann-Winkler, H., Kaul, A., Kleine-Weber, H., Krüger, N., Gassen, N. C., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). Chloroquine does not inhibit infection of human lung cells with SARS-CoV-2. *Nature*, *585*(7826), 588–590. https://doi.org/10.1038/s41586-020-2575-3

Hoffmann, M. et al. The Omicron variant is highly resistant against antibody-mediated neutralization—implications for control of the COVID-19 pandemic. *Cell* https://doi.org/10.1016/j.cell.2021.12.032 (2021).

Hou, Y. J., Chiba, S., Halfmann, P., Ehre, C., Kuroda, M., Dinnon, K. H., 3rd, Leist, S. R., Schäfer, A., Nakajima, N., Takahashi, K., Lee, R. E., Mascenik, T. M., Graham, R., Edwards, C. E., Tse, L. V., Okuda, K., Markmann, A. J., Bartelt, L., de Silva, A., Margolis, D. M., … Baric, R. S. (2020). SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science (New York, N.Y.)*, *370*(6523), 1464–1468. https://doi.org/10.1126/science.abe8499

Hsieh, C.L., Goldsmith, J.A., Schaub, J.M., DiVenere, A.M., Kuo, H.C., Javanmardi, K., Le, K.C., Wrapp, D., Lee, A.G., Liu, Y., et al. (2020). Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. Science *369*, 1501–1505. https://doi.org/10.1126/science.abd0826.

Hsieh, C. L., Werner, A. P., Leist, S. R., Stevens, L. J., Falconer, E., Goldsmith, J. A., Chou, C. W., Abiona, O. M., West, A., Westendorf, K., Muthuraman, K., Fritch, E.

J., Dinnon, K. H., 3rd, Schäfer, A., Denison, M. R., Chappell, J. D., Baric, R. S., Graham, B. S., Corbett, K. S., & McLellan, J. S. (2021). Stabilized coronavirus spike stem elicits a broadly protective antibody. *Cell reports*, *37*(5), 109929. https://doi.org/10.1016/j.celrep.2021.109929.

Hsueh, F.-C., Wu, C.-N., Lin, M.Y.-C., Hsu, F.-Y., Lin, C.-F., Chang, H.-W., Lin, J.-H., Liu, H.-F., Chiou, M.-T., Chan, K.R., et al. (2021). Phylodynamic analysis and spike protein mutations in porcine deltacoronavirus with a new variant introduction in Taiwan. Virus Evol *7*, veab096.

Hu, B., Zeng, L.P., Yang, X.L., Ge, X.Y., Zhang, W., Li, B., Xie, J.Z., Shen, X.R., Zhang, Y.Z., Wang, N., et al. (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog. *13*, e1006698.

Huo, J., Zhao, Y., Ren, J., Zhou, D., Duyvesteyn, H. M. E., Ginn, H. M., Carrique, L., Malinauskas, T., Ruza, R. R., Shah, P. N. M., Tan, T. K., Rijal, P., Coombes, N., Bewley, K. R., Tree, J. A., Radecke, J., Paterson, N. G., Supasa, P., Mongkolsapaya, J., Screaton, G. R., … Stuart, D. I. (2020). Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell host & microbe*, *28*(3), 445–454.e6. https://doi.org/10.1016/j.chom.2020.06.010

Huang, C., et al. "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China." *The Lancet.* **395**, 497-506. (2020).

Imai, M., Iwatsuki-Horimoto, K., Hatta, M., Loeber, S., Halfmann, P. J., Nakajima, N., Watanabe, T., Ujie, M., Takahashi, K., Ito, M., Yamada, S., Fan, S., Chiba, S., Kuroda, M., Guan, L., Takada, K., Armbrust, T., Balogh, A., Furusawa, Y., Okuda, M., … Kawaoka, Y. (2020). Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(28), 16587–16595. https://doi.org/10.1073/pnas.2009799117.

Icho, S., Rujas, E., Muthuraman, K., Tam, J., Liang, H., Landreth, S., Liao, M., Falzarano, D., Julien, J.-P., and Melnyk, R.A. (2022). Dual Inhibition of Vacuolar-ATPase and TMPRSS2 Is Required for Complete Blockade of SARS-CoV-2 Entry into Cells. Antimicrob. Agents Chemother. *66*, e0043922.

Ithete, N.L., Stoffberg, S., Corman, V.M., Cottontail, V.M., Richards, L.R., Schoeman, M.C., Drosten, C., Drexler, J.F., and Preiser, W. (2013). Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. Emerg. Infect. Dis. *19*, 1697–1699.

Jackson, L. A., Anderson, E. J., Rouphael, N. G., Roberts, P. C., Makhene, M., Coler, R. N., McCullough, M. P., Chappell, J. D., Denison, M. R., Stevens, L. J., Pruijssers, A. J., McDermott, A., Flach, B., Doria-Rose, N. A., Corbett, K. S., Morabito, K. M., O'Dell, S., Schmidt, S. D., Swanson, P. A., 2nd, Padilla, M., … mRNA-1273 Study Group (2020). An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. *The New England journal of medicine*, *383*(20), 1920–1931. https://doi.org/10.1056/NEJMoa2022483

Jakalian, A., Bush, B.L., Jack, B.D. & Bayly, C.I. Fast, efficient generation of high-quality atomic charges. AM1-BCC Model: I. Method. *J. Comp. Chem*. 21, 132-146 (2000).

Jette, C.A., Cohen, A.A., Gnanapragasam, P.N.P., Muecksch, F., Lee, Y.E., Huey-Tubman, K.E., Schmidt, F., Hatziioannou, T., Bieniasz, P.D., Nussenzweig, M.C., et al. (2021). Broad cross-reactivity across sarbecoviruses exhibited by a subset of COVID-19 donor-derived neutralizing antibodies. Cell Rep. *36*, 109760.

Ji, W., Peng, Q., Fang, X., Li, Z., Li, Y., Xu, C., Zhao, S., Li, J., Chen, R., Mo, G., et al. (2022). Structures of a deltacoronavirus spike protein bound to porcine and human receptors. Nat. Commun. *13*, 1467.

Jiaming, L., Yanfeng, Y., Yao, D., Yawei, H., Linlin, B., Baoying, H., Jinghua, Y., Gao, G. F., Chuan, Q., & Wenjie, T. (2017). The recombinant N-terminal domain of spike proteins is a potential vaccine against Middle East respiratory syndrome coronavirus (MERS-CoV) infection. *Vaccine*, *35*(1), 10–18. https://doi.org/10.1016/j.vaccine.2016.11.064

Jian, F., Feng, L., Yang, S., Yu, Y., Wang, L., Song, W., Yisimayi, A., Chen, X., Xu, Y., Wang, P., et al. (2023). Convergent evolution of SARS-CoV-2 XBB lineages on receptor-binding domain 455–456 synergistically enhances antibody evasion and ACE2 binding. PLoS Pathog. *19*, e1011868.

Jian, F., Yang, S., Yu, Y., Song, W., Yisimayi, A., Chen, X., Xu, Y., Wang, P., Yu, L., Wang, J., et al. (2023). Convergent evolution of SARS-CoV-2 XBB lineages on receptor-binding domain 455-456 enhances antibody evasion and ACE2 binding. bioRxiv, 2023.08.30.555211. 10.1101/2023.08.30.555211.

Jones, B. E., Brown-Augsburger, P. L., Corbett, K. S., Westendorf, K., Davies, J., Cujec, T. P., Wiethoff, C. M., Blackbourne, J. L., Heinz, B. A., Foster, D., Higgs, R. E., Balasubramaniam, D., Wang, L., Zhang, Y., Yang, E. S., Bidshahri, R., Kraft, L., Hwang, Y., Žentelis, S., Jepson, K. R., … Falconer, E. (2021). The neutralizing antibody, LY-CoV555, protects against SARS-CoV-2 infection in nonhuman primates. *Science translational medicine*, *13*(593), eabf1906. https://doi.org/10.1126/scitranslmed.abf1906

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935

Joung, I.S., and Cheatham, T.E., 3rd (2008). Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. J. Phys. Chem. B *112*, 9020–9041.

Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., Yu, J., Shan, S., Zhou, B., Song, S., Tang, X., Yu, J., Lan, J., Yuan, J., Wang, H., Zhao, J., Zhang, S., Wang, Y., Shi, X., Liu, L., … Zhang, L. (2020). Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature*, *584*(7819), 115–119. https://doi.org/10.1038/s41586-020-2380-z

Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021).

Kabsch W. (2010). XDS. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 2), 125–132. https://doi.org/10.1107/S0907444909047337.

Kan, B., Wang, M., Jing, H., Xu, H., Jiang, X., Yan, M., Liang, W., Zheng, H., Wan, K., Liu, Q., et al. (2005). Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. J. Virol. *79*, 11892–11900.

Kaname, Y., Tani, H., Kataoka, C., Shiokawa, M., Taguwa, S., Abe, T., Moriishi, K., Kinoshita, T., and Matsuura, Y. (2010). Acquisition of complement resistance through incorporation of CD55/decay-accelerating factor into viral particles bearing baculovirus GP64. J. Virol. *84*, 3210–3219.

Kanekiyo, M., Joyce, M. G., Gillespie, R. A., Gallagher, J. R., Andrews, S. F., Yassine, H. M., Wheatley, A. K., Fisher, B. E., Ambrozak, D. R., Creanga, A., Leung, K., Yang, E. S., Boyoglu-Barnum, S., Georgiev, I. S., Tsybovsky, Y., Prabhakaran, M. S., Andersen, H., Kong, W. P., Baxa, U., Zephir, K. L., … Graham, B. S. (2019). Mosaic nanoparticle display of diverse influenza virus hemagglutinins elicits broad B cell responses. *Nature immunology*, *20*(3), 362–372. https://doi.org/10.1038/s41590-018-0305-x

Kang, B.H., Lax, B.M., and Wittrup, K.D. (2022). Yeast surface display for protein engineering: Library generation, screening, and affinity maturation. Methods Mol. Biol. *2491*, 29–62.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Keech, C., Albert, G., Cho, I., Robertson, A., Reed, P., Neal, S., Plested, J. S., Zhu, M., Cloney-Clark, S., Zhou, H., Smith, G., Patel, N., Frieman, M. B., Haupt, R. E., Logue, J., McGrath, M., Weston, S., Piedra, P. A., Desai, C., Callahan, K., … Glenn, G. M. (2020). Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. *The New England journal of medicine*, *383*(24), 2320–2332. https://doi.org/10.1056/NEJMoa2026920

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649.

Khoury, D. S., Cromer, D., Reynaldi, A., Schlub, T. E., Wheatley, A. K., Juno, J. A., Subbarao, K., Kent, S. J., Triccas, J. A., & Davenport, M. P. (2021). Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nature medicine*, *27*(7), 1205–1211. https://doi.org/10.1038/s41591-021-01377-8

Kim, C. et al. A therapeutic neutralizing antibody targeting receptor binding domain of SARS-CoV-2 spike protein. *Nat. Commun.* **12**, 288 (2021).

Kimanius, D., Forsberg, B. O., Scheres, S. H., & Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife*, *5*, e18722. https://doi.org/10.7554/eLife.18722

Kirchdoerfer, R. et al. (2016) Pre-fusion structure of a human coronavirus spike protein. Nature. 2016;531(7592):118-121. doi:10.1038/nature17200.

Kirchdoerfer, R.N., Wang, N., Pallesen, J., Wrapp, D., Turner, H.L., Cottrell, C.A., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2018). Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. Sci. Rep. *8*, 15701.

Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L., and Woods, R.J. (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J. Comput. Chem. *29*, 622–655.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H., and Frost, S.D.W. (2006). GARD: a genetic algorithm for recombination detection. Bioinformatics *22*, 3096–3098.

Kokic, G., Hillen, H.S., Tegunov, D., Dienemann, C., Seitz, F., Schmitzova, J., Farnung, L., Siewert, A., Höbartner, C., and Cramer, P. (2021). Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. Nat. Commun. *12*, 279.

Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E. E., Bhattacharya, T., Foley, B., Hastie, K. M., Parker, M. D., Partridge, D. G., Evans, C. M., Freeman, T. M., de Silva, T. I., Sheffield COVID-19 Genomics Group, McDanal, C., Perez, L. G., Tang, H., … Montefiori, D. C. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell*, *182*(4), 812–827.e19. https://doi.org/10.1016/j.cell.2020.06.043.

Krammer, F., Srivastava, K., Alshammary, H., Amoako, A. A., Awawda, M. H., Beach, K. F., Bermúdez-González, M. C., Bielak, D. A., Carreño, J. M., Chernet, R. L., Eaker, L. Q., Ferreri, E. D., Floda, D. L., Gleason, C. R., Hamburger, J. Z., Jiang, K., Kleiner, G., Jurczyszak, D., Matthews, J. C., Mendez, W. A., … Simon, V. (2021). Antibody Responses in Seropositive Persons after a Single Dose of SARS-CoV-2 mRNA Vaccine. *The New England journal of medicine*, *384*(14), 1372–1374. https://doi.org/10.1056/NEJMc2101667.

Krintel, C. et al. Thermodynamics and structural analysis of positive allosteric modulation of the ionotropic glutamate receptor GluA2. *Biochem J.* 441, 173-178 (2012).

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assem- blies from crystalline state. J. Mol. Biol. *372*, 774–797.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., et al. (2003). A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. *348*, 1953–1966.

Kuhn RJ, Strauss JH. Enveloped viruses. Adv Protein Chem. 2003;64:363-77. doi: 10.1016/s0065-3233(03)01010-6. PMID: 13677053.

Kuhn, J. H., Li, W., Choe, H., & Farzan, M. (2004). Angiotensin-converting enzyme 2: a functional receptor for SARS coronavirus. *Cellular and molecular life sciences : CMLS*, *61*(21), 2738–2743. https://doi.org/10.1007/s00018-004-4242-5.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol. Evol. *35*, 1547–1549.

Kupferschmidt K. (2021). Where did 'weird' Omicron come from?. *Science (New York, N.Y.)*, *374*(6572), 1179. https://doi.org/10.1126/science.acx9738

Lam, T.T., Jia, N., Zhang, Y.W., Shum, M.H., Jiang, J.F., Zhu, H.C., Tong, Y.G., Shi, Y.X., Ni, X.B., Liao, Y.S., et al. (2020). Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. Nature *583*, 282–285. https://doi.org/10.1038/s41586-020-2169-0.

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., et al. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature *581*, 215–220. https://doi.org/10.1038/s41586-020-2180-5.

Landrum, G. RDKit: Open-source cheminformatics. *RDKit* https://www.rdkit.org [https://www.rdkit.org] (2020).
Latinne, A. et al. Origin and cross-species transmission of bat coronaviruses in China. *Nat. Commun.* 11, 4235 (2020).
Lau, S. K. P. et al. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl Acad. Sci. USA* 102, 14040–14045 (2005).
Lau, S.K.P., Luk, H.K.H., Wong, A.C.P., Fan, R.Y.Y., Lam, C.S.F., Li, K.S.M., Ahmed, S.S., Chow, F.W.N., Cai, J.-P., Zhu, X., et al. (2019). Identification of a Novel Betacoronavirus (Merbecovirus) in Amur Hedgehogs from China. Viruses *11*. https://doi.org/10.3390/v11110980.
Lau, S. *et al.* (2020) Attenuated SARS-CoV-2 Variants with deletions at the S1/S2 Junction. *Emerg Microbes Infect.* **9(1)**: 837-842.
Lawson, C.L., et al. EMDataBank unified data resource for 3DEM. *Nucleic Acids Res*. 44, D396-D403 (2016).
Lednicky, J.A., Tagliamonte, M.S., White, S.K., Elbadry, M.A., Alam, M.M., Stephenson, C.J., Bonny, T.S., Loeb, J.C., Telisma, T., Chavannes, S., et al. (2021). Independent infections of porcine deltacoronavirus among Haitian children. Nature *600*, 133–137.
Lee, F.S., Anderson, A.G., and Olafson, B.D. (2023). Benchmarking TriadAb using targets from the second antibody modeling assessment. Protein Eng. Des. Sel. *36*. 10.1093/protein/gzad013.
Lee, J., Zepeda, S.K., Park, Y.-J., Taylor, A.L., Quispe, J., Stewart, C., Leaf, E.M., Treichel, C., Corti, D., King, N.P., et al. (2023). Broad receptor tropism and immunogenicity of a clade 3 sarbecovirus. Cell Host Microbe *31*, 1961–1973.e11.
Leimkuhler, B., and Matthews, C. (2013). Robust and efficient configurational molecular sampling via Langevin dynamics. J. Chem. Phys. *138*, 174102.
Leimkuhler, B., and Matthews, C. (2016). Efficient molecular dynamics using geodesic integration and solvent–solute splitting. Proc. Math. Phys. Eng. Sci. *472*, 20160138.
Leist, S. R., Dinnon, K. H., 3rd, Schäfer, A., Tse, L. V., Okuda, K., Hou, Y. J., West, A., Edwards, C. E., Sanders, W., Fritch, E. J., Gully, K. L., Scobey, T., Brown, A. J., Sheahan, T. P., Moorman, N. J., Boucher, R. C., Gralinski, L. E., Montgomery, S. A., & Baric, R. S. (2020). A Mouse-Adapted SARS-CoV-2 Induces Acute Lung Injury and Mortality in Standard Laboratory Mice. *Cell*, *183*(4), 1070–1085.e12. https://doi.org/10.1016/j.cell.2020.09.050.
Lempp, F. A., Soriaga, L. B., Montiel-Ruiz, M., Benigni, F., Noack, J., Park, Y. J., Bianchi, S., Walls, A. C., Bowen, J. E., Zhou, J., Kaiser, H., Joshi, A., Agostini, M., Meury, M., Dellota, E., Jr, Jaconi, S., Cameroni, E., Martinez-Picado, J., Vergara-Alert, J., Izquierdo-Useros, N., … Corti, D. (2021). Lectins enhance SARS-CoV-2 infection and influence neutralizing antibodies. *Nature*, *598*(7880), 342–347. https://doi.org/10.1038/s41586-021-03925-1.
Letko, M., Marzi, A., and Munster, V. (2020). Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B beta- coronaviruses. Nat. Microbiol. *5*, 562–569.

Letko, M. (2024). Functional assessment of cell entry and receptor use for merbecoviruses. bioRxiv, 2024.03.13.584892. https://doi.org/10.1101/2024.03.13.584892.
Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. *52*, W78–W82.
Li, F., Li, W., Farzan, M., and Harrison, S.C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science *309*, 1864–1868.
Li, J. et al. Cryo-EM structures of *Escherichia coli* cytochrome *bo3* reveal bound phospholipids and ubiquinone-8 in a dynamic substrate binding site. *Proc Natl Acad Sci USA* 118, e2106750118 (2021).
Li, Q., Wu, J., Nie, J., Zhang, L., Hao, H., Liu, S., Zhao, C., Zhang, Q., Liu, H., Nie, L., Qin, H., Wang, M., Lu, Q., Li, X., Sun, Q., Liu, J., Zhang, L., Li, X., Huang, W., & Wang, Y. (2020). The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell*, *182*(5), 1284–1294.e9. https://doi.org/10.1016/j.cell.2020.07.012
Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature *426*, 450–454.
Li, W. *et al*. (2004). Efficient Replication of Severe Acute Respiratory Syndrome Coronavirus in Mouse Cells Is Limited by Murine Angiotensin-Converting Enzyme 2. *J Virol.***78**: 11429-11433.
Li, W. et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310, 676–679 (2005a).
Li, W. et al. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J.* 24, 1634–1643 (2005b).
Li, W., Chen, Y., Prévost, J., Ullah, I., Lu, M., Gong, S.Y., Tauzin, A., Gasser, R., Vézina, D., Anand, S.P., et al. (2022). Structural basis and mode of action for two broadly neutralizing antibodies against SARS-CoV-2 emerging variants of concern. Cell Rep. *38*, 110210.
Liebschner, D., Afonine, P.V., Baker, M.L., Bunko´czi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. *75*, 861–877.
Lieu, R., Antonysamy, S., Druzina, Z., Ho, C., Kang, N.R., Pustilnik, A., Wang, J., and Atwell, S. (2020). Rapid and robust antibody Fab fragment crystallization utilizing edge-to-edge beta-sheet packing. PLoS One *15*, e0232311.
Liu, L., Wang, P., Nair, M. S., Yu, J., Rapp, M., Wang, Q., Luo, Y., Chan, J. F., Sahi, V., Figueroa, A., Guo, X. V., Cerutti, G., Bimela, J., Gorman, J., Zhou, T., Chen, Z., Yuen, K. Y., Kwong, P. D., Sodroski, J. G., Yin, M. T., … Ho, D. D. (2020). Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature*, *584*(7821), 450–456.
Liu, P., Huang, M.-L., Guo, H., Si, J.-Y., Chen, Y.-M., Wang, C.-L., Yu, X., Shi, L.-L., Xiong, Q., Ma, C.-B., et al. (2024). Engineering customized viral receptors for

various coronaviruses. bioRxiv, 2024.03.03.583237. https://doi.org/10.1101/2024.03.03.583237.
Liu, Z., VanBlargan, L. A., Bloyet, L. M., Rothlauf, P. W., Chen, R. E., Stumpf, S., Zhao, H., Errico, J. M., Theel, E. S., Liebeskind, M. J., Alford, B., Buchser, W. J., Ellebedy, A. H., Fremont, D. H., Diamond, M. S., & Whelan, S. P. J. (2021). Landscape analysis of escape variants identifies SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *bioRxiv : the preprint server for biology*, 2020.11.06.372037. https://doi.org/10.1101/2020.11.06.372037
Liu, Z., VanBlargan, L. A., Bloyet, L. M., Rothlauf, P. W., Chen, R. E., Stumpf, S., Zhao, H., Errico, J. M., Theel, E. S., Liebeskind, M. J., Alford, B., Buchser, W. J., Ellebedy, A. H., Fremont, D. H., Diamond, M. S., & Whelan, S. P. J. (2021). Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell host & microbe*, *29*(3), 477–488.e4. https://doi.org/10.1016/j.chom.2021.01.014
Liu, K. et al. Binding and molecular basis of the bat coronavirus RaTG13 virus to ACE2 in humans and other species. *Cell* 184, 3438–3451 (2021).
Low, J.S., Jerak, J., Tortorici, M.A., McCallum, M., Pinto, D., Cassotta, A., Foglierini, M., Mele, F., Abdelnabi, R., Weynand, B., et al. (2022). ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies. Science *377*, 735–742.
Lu, G., Hu, Y., Wang, Q., Qi, J., Gao, F., Li, Y., Zhang, Y., Zhang, W., Yuan, Y., Bao, J., et al. (2013). Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature *500*, 227–231.
Lyu, M. et al. Cryo-EM structures of a gonococcal multidrug efflux pump illuminate a mechanism of drug recognition and resistance. *mBio* 11, e00996-20 (2020).
Ma, C., Liu, C., Xiong, Q., Gu, M., Shi, L., Wang, C., Si, J., Tong, F., Liu, P., Huang, M., et al. (2023). Broad host tropism of ACE2-using MERS-related coronaviruses and determinants restricting viral recognition. Cell Discov *9*, 57.
Ma, C.-B., Liu, C., Park, Y.-J., Tang, J., Chen, J., Xiong, Q., Lee, J., Stewart, C., Asarnow, D., Brown, J., et al. (2024). Multiple independent acquisitions of ACE2 usage in MERS-related coronaviruses. bioRxiv, 2023.10.02.560486. https://doi.org/10.1101/2023.10.02.560486.
Mackin, S.R., Desai, P., Whitener, B.M., Karl, C.E., Liu, M., Baric, R.S., Edwards, D.K., Chicz, T.M., McNamara, R.P., Alter, G., et al. (2023). Fc-γR-dependent antibody effector functions are required for vaccine- mediated protection against antigen-shifted variants of SARS-CoV-2. Nat. Microbiol. *8*, 569–580.
Madhi, S. A., Baillie, V., Cutland, C. L., Voysey, M., Koen, A. L., Fairlie, L., Padayachee, S. D., Dheda, K., Barnabas, S. L., Bhorat, Q. E., Briner, C., Kwatra, G., Ahmed, K., Aley, P., Bhikha, S., Bhiman, J. N., Bhorat, A. E., du Plessis, J., Esmail, A., Groenewald, M., … Wits-VIDA COVID Group (2021). Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *The New England journal of medicine*, *384*(20), 1885–1898. https://doi.org/10.1056/NEJMoa2102214

Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). Ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. J. Chem. Theory Comput. *11*, 3696–3713.

Makarenkov, V., Mazoure, B., Rabusseau, G., & Legendre, P. (2021). Horizontal gene transfer and recombination analysis of SARS-CoV-2 genes helps discover its close relatives and shed light on its origin. *BMC ecology and evolution*, *21*(1), 5. https://doi.org/10.1186/s12862-020-01732-2

Makowski, E.K., Kinnunen, P.C., Huang, J., Wu, L., Smith, M.D., Wang, T., Desai, A.A., Streu, C.N., Zhang, Y., Zupancic, J.M., et al. (2022). Co-optimization of therapeutic antibody affinity and specificity using machine learning models that generalize to novel mutational space. Nat. Commun. *13*, 3788.

Makowski, E.K., Chen, H., Lambert, M., Bennett, E.M., Eschmann, N.S., Zhang, Y., Zupancic, J.M., Desai, A.A., Smith, M.D., Lou, W., et al. (2022). Reduction of therapeutic antibody self-association using yeast-display selections and machine learning. MAbs *14*, 2146629.

Martinez, D. R., Schäfer, A., Leist, S. R., De la Cruz, G., West, A., Atochina-Vasserman, E. N., Lindesmith, L. C., Pardi, N., Parks, R., Barr, M., Li, D., Yount, B., Saunders, K. O., Weissman, D., Haynes, B. F., Montgomery, S. A., & Baric, R. S. (2021). Chimeric spike mRNA vaccines protect against Sarbecovirus challenge in mice. *Science (New York, N.Y.)*, *373*(6558), 991–998. https://doi.org/10.1126/science.abi4506

Martinez, D. R., Schaefer, A., Gobeil, S., Li, D., De la Cruz, G., Parks, R., Lu, X., Barr, M., Manne, K., Mansouri, K., Edwards, R. J., Yount, B., Anasti, K., Montgomery, S. A., Shen, S., Zhou, T., Kwong, P. D., Graham, B. S., Mascola, J. R., Montefiori, D. C., … Baric, R. S. (2021). A broadly neutralizing antibody protects against SARS-CoV, pre-emergent bat CoVs, and SARS-CoV-2 variants in mice. *bioRxiv : the preprint server for biology*, 2021.04.27.441655. https://doi.org/10.1101/2021.04.27.441655

Mason, D.M., Friedensohn, S., Weber, C.R., Jordi, C., Wagner, B., Meng, S.M., Ehling, R.A., Bonati, L., Dahinden, J., Gainza, P., et al. (2021). Optimization of therapeutic antibodies by predicting antigen specificity from antibody sequence via deep learning. Nat. Biomed. Eng. *5*, 600–612.

Martinez, D. R., Schäfer, A., Gobeil, S., Li, D., De la Cruz, G., Parks, R., Lu, X., Barr, M., Stalls, V., Janowska, K., Beaudoin, E., Manne, K., Mansouri, K., Edwards, R. J., Cronin, K., Yount, B., Anasti, K., Montgomery, S. A., Tang, J., Golding, H., … Baric, R. S. (2022). A broadly cross-reactive antibody neutralizes and protects against sarbecovirus challenge in mice. *Science translational medicine*, *14*(629), eabj7125. https://doi.org/10.1126/scitranslmed.abj7125

Mayo, S.L., Olafson, B.D., and Goddard, W.A. (1990). DREIDING: a generic force field for molecular simulations. J. Phys. Chem. *94*, 8897–8909.

Maziarz, R. T., Sridharan, P., Slater, S., Meyers, G., Post, M., Erdman, D. D., Peret, T. C., & Taplitz, R. A. (2010). Control of an outbreak of human parainfluenza virus 3 in hematopoietic stem cell transplant recipients. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation*, *16*(2), 192–198. https://doi.org/10.1016/j.bbmt.2009.09.014

McCallum, M., Walls, A. C., Bowen, J. E., Corti, D., & Veesler, D. (2020). Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nature structural & molecular biology*, *27*(10), 942–949. https://doi.org/10.1038/s41594-020-0483-8

McCallum, M., De Marco, A., Lempp, F.A., Tortorici, M.A., Pinto, D., Walls, A.C., Beltramello, M., Chen, A., Liu, Z., Zatta, F., et al. (2021). N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. Cell *184*, 2332–2347.e16.

McCallum, M., Bassi, J., De Marco, A., Chen, A., Walls, A.C., Di Iulio, J., Tortorici, M.A., Navarro, M.J., Silacci-Fregni, C., Saliba, C., et al. (2021). SARS-CoV-2 immune evasion by the B.1.427/B.1.429 variant of concern. Science *373*, 648–654. https://doi.org/10.1126/science.abi7994.

McCallum, M. et al. Molecular basis of immune evasion by the Delta and Kappa SARS-CoV-2 variants. *Science* **374**, 1621–1626 (2021).

McCallum, M., Czudnochowski, N., Rosen, L.E., Zepeda, S.K., Bowen, J.E., Walls, A.C., Hauser, K., Joshi, A., Stewart, C., Dillen, J.R., et al. (2022). Structural Basis of SARS-CoV-2 Omicron Immune Evasion and Receptor Engagement. Science, eabn8652.

McCallum, M., Park, Y.-J., Stewart, C., Sprouse, K.R., Addetia, A., Brown, J., Tortorici, M.A., Gibson, C., Wong, E., Ieven, M., et al. (2024). Human coronavirus HKU1 recognition of the TMPRSS2 host receptor. Cell. https://doi.org/10.1016/j.cell.2024.06.006.

McCarthy, K. R., Rennick, L. J., Nambulli, S., Robinson-McCarthy, L. R., Bain, W. G., Haidar, G., & Duprex, W. P. (2021). Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science (New York, N.Y.)*, *371*(6534), 1139–1142. https://doi.org/10.1126/science.abf6950

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography*, *40*(Pt 4), 658–674. https://doi.org/10.1107/S0021889807021206

McMahan, K., Yu, J., Mercado, N. B., Loos, C., Tostanoski, L. H., Chandrashekar, A., Liu, J., Peter, L., Atyeo, C., Zhu, A., Bondzie, E. A., Dagotto, G., Gebre, M. S., Jacob-Dolan, C., Li, Z., Nampanya, F., Patel, S., Pessaint, L., Van Ry, A., Blade, K., … Barouch, D. H. (2021). Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature*, *590*(7847), 630–634. https://doi.org/10.1038/s41586-020-03041-6

Menachery, V.D., Dinnon, K.H., 3rd, Yount, B.L., Jr, McAnarney, E.T., Gralinski, L.E., Hale, A., Graham, R.L., Scobey, T., Anthony, S.J., Wang, L., et al. (2020). Trypsin Treatment Unlocks Barrier for Zoonotic Bat Coronavirus Infection. J. Virol. *94*. https://doi.org/10.1128/JVI.01774-19.

Menachery, V.D., Yount, B.L., Debbink, K., Agnihothram, S., Gralinski, L.E., Plante, J.A., Graham, R.L., Scobey, T., Ge, X.Y., Donaldson, E.F., et al. (2015). A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat. Med. *21*, 1508–1513.

Menachery, V.D., Yount, B.L., Sims, A.C., Debbink, K., Agnihothram, S.S., Gralinski, L.E., Graham, R.L., Scobey, T., Plante, J.A., Royal, S.R., et al. (2016). SARS-like

WIV1-CoV poised for human emergence. Proc. Natl. Acad. Sci. USA *113*, 3048–3053.
Meng, B., Abdullahi, A., Ferreira, I.A.T.M., Goonawardane, N., Saito, A., Kimura, I., Yamasoba, D., Gerber, P.P., Fatihi, S., Rathore, S., et al. (2022). Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. Nature *603*, 706–714.
Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C. & Ferrin, T.E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* 7, 339 (2006).
Mercado, N. B., Zahn, R., Wegmann, F., Loos, C., Chandrashekar, A., Yu, J., Liu, J., Peter, L., McMahan, K., Tostanoski, L. H., He, X., Martinez, D. R., Rutten, L., Bos, R., van Manen, D., Vellinga, J., Custers, J., Langedijk, J. P., Kwaks, T., Bakkers, M. J. G., … Barouch, D. H. (2020). Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature*, *586*(7830), 583–588. https://doi.org/10.1038/s41586-020-2607-z
Merk, A. et al. 1.8 Å resolution structure of β-galactosidase with a 200 kV CRYO ARM electron microscope. *IUCrJ* 7, 639-643 (2020).
Mlcochova, P. et al. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature* **599**, 114–119 (2021).
Millet, J.K., and Whittaker, G.R. (2014). Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proc. Natl. Acad. Sci. U. S. A. *111*, 15214–15219.
Millet, J.K., and Whittaker, G.R. (2015). Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. Virus Res. *202*, 120–134.
Millet, J.K., and Whittaker, G.R. (2016). Murine leukemia virus (MLV)- based coronavirus spike-pseudotyped particle production and infection. Bio Protoc. *6*, 6. https://doi.org/10.21769/BioProtoc.2035.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. *37*, 1530–1534.
Mok, C.K.P., Zhu, A., Zhao, J., Lau, E.H.Y., Wang, J., Chen, Z., Zhuang, Z., Wang, Y., Alshukairi, A.N., Baharoon, S.A., et al. (2020). T-cell responses to MERS coronavirus infection in people with occupational exposure to dromedary camels in Nigeria: an observational cohort study. Lancet Infect. Dis. https://doi.org/10.1016/S1473-3099(20)30599-5.
Momont, C., Dang, H.V., Zatta, F., Hauser, K., Wang, C., di Iulio, J., Minola, A., Czudnochowski, N., De Marco, A., Branch, K., et al. (2023). A pan-influenza antibody inhibiting neuraminidase via receptor mimicry. Nature *618*, 590–597.
Moriarty, N.W., Grosse-Kunstleve, R.W. & Adams, P.D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Cryst.* D65, 1074-1080 (2009).
Mou, H. et al. Mutations derived from horseshoe bat ACE2 orthologs enhance ACE2-Fc neutralization of SARS-CoV-2. *PLoS Pathog.* 17, e1009501 (2021).
Murakami, S. et al. Detection and characterization of bat Sarbecovirus phylogenetically related to SARS-CoV-2, Japan. *Emerg. Infect. Dis.* 26, 3025–3029 (2020).

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography*, *67*(Pt 4), 355–367. https://doi.org/10.1107/S0907444911001314
Mykytyn, A.Z., Fouchier, R.A., and Haagmans, B.L. (2023). Antigenic evolution of SARS coronavirus 2. Curr. Opin. Virol. *62*, 101349.
Nakane, T. et al. Single-particle cryo-EM at atomic resolution. *Nature* 587, 152–156 (2020).
Nakanishi, H. et al. Transport cycle of plasma membrane flippase ATP11C by cryo-EM. *Cell Rep.* 32, 108208 (2020).
Nao, N. *et al.* (2017) Genetic Predisposition to acquire a polybasic cleavage site for highly pathogenic avian influenza virus hemagglutinin. *mBio*. **8(1)**: e002298-16.
Ngere, I., Hunsperger, E.A., Tong, S., Oyugi, J., Jaoko, W., Harcourt, J.L., Thornburg, N.J., Oyas, H., Muturi, M., Osoro, E.M., et al. (2022). Outbreak of Middle East Respiratory Syndrome Coronavirus in Camels and Probable Spillover Infection to Humans in Kenya. Viruses *14*. https://doi.org/10.3390/v14081743.
Nicholls, J. & Chan Chi-wai, M. HKUMed finds Omicron SARS-CoV-2 can infect faster and better than Delta in human bronchus but with less severe infection in lung. www.med.hku.hk/en/news/press/20211215-omicron-sars-cov-2-infection (2021).
Niu, S., Wang, J., Bai, B., Wu, L., Zheng, A., Chen, Q., Du, P., Han, P., Zhang, Y., Jia, Y., et al. (2022). Molecular basis of cross-species ACE2 interactions with SARS-CoV-2-like viruses of pangolin origin. EMBO J. *41*, e109962.
Niu, Z., Zhang, Z., Gao, X., Du, P., Lu, J., Yan, B., Wang, C., Zheng, Y., Huang, H., and Sun, Q. (2021). N501Y mutation imparts cross-species transmission of SARS-CoV-2 to mice by enhancing receptor binding. Signal Transduct Target Ther *6*, 284.
Nutalai, R., Zhou, D., Tuekprakhon, A., Ginn, H.M., Supasa, P., Liu, C., Huo, J., Mentzer, A.J., Duyvesteyn, H.M.E., Dijokaite-Guraliuc, A., et al. (2022). Potent cross-reactive antibodies following Omicron breakthrough in vaccinees. Cell *185*, 2116-2131.e18.
O'Boyle, N.M. et al. Open Babel: An open chemical toolbox. *J. Cheminform.* 3, 33 (2011).
Oldfield, T.J. X-LIGAND: an application for the automated addition of flexible ligands into electron density. *Acta Cryst.* D57, 696-705 (2001).
Otwinowski, J., McCandlish, D. M., & Plotkin, J. B. (2018). Inferring the shape of global epistasis. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(32), E7550–E7558. https://doi.org/10.1073/pnas.1804015115
Ou, X., Liu, Y., Lei, X., Li, P., Mi, D., Ren, L., Guo, L., Guo, R., Chen, T., Hu, J., Xiang, Z., Mu, Z., Chen, X., Chen, J., Hu, K., Jin, Q., Wang, J., & Qian, Z. (2020). Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature communications*, *11*(1), 1620. https://doi.org/10.1038/s41467-020-15562-9
Oude Munnink, B. B., Sikkema, R. S., Nieuwenhuijse, D. F., Molenaar, R. J., Munger, E., Molenkamp, R., van der Spek, A., Tolsma, P., Rietveld, A., Brouwer, M., Bouwmeester-Vincken, N., Harders, F., Hakze-van der Honing, R., Wegdam-

Blans, M. C. A., Bouwstra, R. J., GeurtsvanKessel, C., van der Eijk, A. A., Velkers, F. C., Smit, L. A. M., Stegeman, A., … Koopmans, M. P. G. (2021). Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science (New York, N.Y.)*, *371*(6525), 172–177. https://doi.org/10.1126/science.abe5901

Ozorowski, G., Torres, J.L., Santos-Martins, D., Forli, S. & Ward, A.B. A strain-specific inhibitor of receptor-bound HIV-1 targets a pocket near the fusion peptide. *Cell Rep.* 33, 108428 (2020).

Pallesen, J. *et al.* (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *PNAS*. **114**: E7348–E7357.

Pan, T. et al. Infection of wild-type mice by SARS-CoV-2 B.1.351 variant indicates a possible novel cross-species transmission route. *Sig. Transduct. Target Ther.* **6**, 420 (2021).

Park, H., Zhou, G., Baek, M., Baker, D. & DiMaio, F. Force field optimization guided by small molecule crystal lattice data enables consistent sub-angstrom protein–ligand docking. *J. Chem. Theory Comput*. 17, 2000-2010 (2021).

Park, J.-E., Li, K., Barlan, A., Fehr, A.R., Perlman, S., McCray, P.B., Jr, and Gallagher, T. (2016). Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. Proc. Natl. Acad. Sci. U. S. A. *113*, 12262–12267.

Park, Y.J., Walls, A.C., Wang, Z., Sauer, M.M., Li, W., Tortorici, M.A., Bosch, B.J., DiMaio, F., and Veesler, D. (2019). Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. Nat. Struct. Mol. Biol. *26*, 1151–1157.

Park, Y.-J., Pinto, D., Walls, A.C., Liu, Z., De Marco, A., Benigni, F., Zatta, F., Silacci-Fregni, C., Bassi, J., Sprouse, K.R., et al. (2022). Imprinted antibody responses against SARS-CoV-2 Omicron sublineages. Science, eadc9127.

Park, Y.-J., De Marco, A., Starr, T.N., Liu, Z., Pinto, D., Walls, A.C., Zatta, F., Zepeda, S.K., Bowen, J.E., Sprouse, K.R., et al. (2022). Antibody-mediated broad sarbecovirus neutralization through ACE2 molecular mimicry. Science, eabm8143.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2012). Scikit-learn: Machine Learning in Python. arXiv [cs.LG].

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Piccoli, L., Park, Y.J., Tortorici, M.A., Czudnochowski, N., Walls, A.C., Beltramello, M., Silacci-Fregni, C., Pinto, D., Rosen, L.E., Bowen, J.E., et al. (2020). Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high- resolution serology. Cell *183*, 1024–1042.e21.

Pinto, D., Fenwick, C., Caillat, C., Silacci, C., Guseva, S., Dehez, F., Chipot, C., Barbieri, S., Minola, A., Jarrossay, D., Tomaras, G. D., Shen, X., Riva, A., Tarkowski, M., Schwartz, O., Bruel, T., Dufloo, J., Seaman, M. S., Montefiori, D. C., Lanzavecchia, A., … Weissenhorn, W. (2019). Structural Basis for Broad

HIV-1 Neutralization by the MPER-Specific Human Broadly Neutralizing Antibody LN01. *Cell host & microbe*, *26*(5), 623–637.e8. https://doi.org/10.1016/j.chom.2019.09.016

Pinto, D., Park, Y.J., Beltramello, M., Walls, A.C., Tortorici, M.A., Bianchi, S., Jaconi, S., Culap, K., Zatta, F., De Marco, A., et al. (2020). Cross- neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. Nature *583*, 290–295.

Pinto, D., Sauer, M.M., Czudnochowski, N., Low, J.S., Tortorici, M.A., Housley, M.P., Noack, J., Walls, A.C., Bowen, J.E., Guarino, B., et al. (2021). Broad Betacoronavirus neutralization by a stem helix–specific human antibody. Science *373*, 1109–1116. https://doi.org/10.1126/sci- ence.abj3321.

Plante, J. A., Liu, Y., Liu, J., Xia, H., Johnson, B. A., Lokugamage, K. G., Zhang, X., Muruato, A. E., Zou, J., Fontes-Garfias, C. R., Mirchandani, D., Scharton, D., Bilello, J. P., Ku, Z., An, Z., Kalveram, B., Freiberg, A. N., Menachery, V. D., Xie, X., Plante, K. S., … Shi, P. Y. (2021). Spike mutation D614G alters SARS-CoV-2 fitness. *Nature*, *592*(7852), 116–121. https://doi.org/10.1038/s41586-020-2895-3

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W., Jr, Hammitt, L. L., Türeci, Ö., … C4591001 Clinical Trial Group (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *The New England journal of medicine*, *383*(27), 2603–2615. https://doi.org/10.1056/NEJMoa2034577

Pronker, M.F., Creutznacher, R., Drulyte, I., Hulswit, R.J.G., Li, Z., van Kuppeveld, F.J.M., Snijder, J., Lang, Y., Bosch, B.-J., Boons, G.-J., et al. (2023). Sialoglycan binding triggers spike opening in a human coronavirus. Nature. https://doi.org/10.1038/s41586-023-06599-z.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Punjani, A., Zhang, H., and Fleet, D.J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat. Methods *17*, 1214–1221.

Qu, P., Evans, J.P., Zheng, Y.-M., Carlin, C., Saif, L.J., Oltz, E.M., Xu, K., Gumina, R.J., and Liu, S.-L. (2022). Evasion of neutralizing antibody responses by the SARS-CoV-2 BA.2.75 variant. Cell Host Microbe *30*, 1518-1526.e4.

Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H.W., Müller, M.A., Dijkman, R., Muth, D., Demmers, J.A.A., Zaki, A., Fouchier, R.A.M., et al. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature *495*, 251–254.

Rappazzo, C. G., Tse, L. V., Kaku, C. I., Wrapp, D., Sakharkar, M., Huang, D., Deveau, L. M., Yockachonis, T. J., Herbert, A. S., Battles, M. B., O'Brien, C. M., Brown, M. E., Geoghegan, J. C., Belk, J., Peng, L., Yang, L., Hou, Y., Scobey, T. D., Burton, D. R., Nemazee, D., … Walker, L. M. (2021). Broad and potent activity against SARS-like viruses by an engineered human monoclonal antibody. *Science (New York, N.Y.)*, *371*(6531), 823–829. https://doi.org/10.1126/science.abf4830

Reed, L.J., Muench, H. (1938) A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* ***27****:493-497.* https://doi.org/10.1093/oxfordjournals.aje.a118408
Regeneron Pharmaceuticals Third Quarter 2023 Earnings Call (2023). https://investor.regeneron.com/static-files/27bf1467-a410-414b-9bec-916fa90e7748
Reguera, J., Santiago, C., Mudgal, G., Ordono, D., Enjuanes, L., and Casasnovas, J.M. (2012). Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. PLoS Pathog. *8*, e1002859.
Reguera, J., Mudgal, G., Santiago, C. & Casasnovas, J. M. A structural view of coronavirus-receptor interactions. *Virus Res.* 194, 3–15 (2014).
Ren, W. et al. Difference in receptor usage between severe acute respiratory syndrome (SARS) coronavirus and SARS-like coronavirus of bat origin. *J. Virol.* 82, 1899–1907 (2008).
Robbiani, D.F., Gaebler, C., Muecksch, F., Lorenzi, J.C.C., Wang, Z., Cho, A., Agudelo, M., Barnes, C.O., Gazumyan, A., Finkin, S., et al. (2020). Convergent antibody responses to SARS-CoV-2 in convalescent individuals. Nature *584*, 437–442. https://doi.org/10.1038/s41586-020- 2456-9.
Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. *42*, W320–W324.
Roberts, A., Deming, D., Paddock, C. D., Cheng, A., Yount, B., Vogel, L., Herman, B. D., Sheahan, T., Heise, M., Genrich, G. L., Zaki, S. R., Baric, R., & Subbarao, K. (2007). A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. *PLoS pathogens*, *3*(1), e5. https://doi.org/10.1371/journal.ppat.0030005
Robertson, M.J., van Zundert, G.C.P., Borrelli, K. & Skiniotis, G. GemSpot: A pipeline for robust modeling of ligands into cryo-EM maps. *Structure* 28, 707-716 (2020).
Rockx, B., Corti, D., Donaldson, E., Sheahan, T., Stadler, K., Lanzavecchia, A., & Baric, R. (2008). Structural basis for potent cross-neutralizing human monoclonal antibody protection against lethal human and zoonotic severe acute respiratory syndrome coronavirus challenge. *Journal of virology*, *82*(7), 3220–3235. https://doi.org/10.1128/JVI.02377-07
Roelle, S.M., Shukla, N., Pham, A.T., Bruchez, A.M., and Matreyek, K.A. (2022). Expanded ACE2 dependencies of diverse SARS-like coronavirus receptor binding domains. PLoS Biol. *20*, e3001738.
Roemer, C., Hisner, R., Frohberg, N., Sakaguchi, H. Guelu, F. and Peacock, T., (2022). SARS-CoV-2 evolution, post-Omicron. Virological. https:// virological.org/t/sars-cov-2-evolution-post-omicron/911.
Rogers, T. F., Zhao, F., Huang, D., Beutler, N., Burns, A., He, W. T., Limbo, O., Smith, C., Song, G., Woehl, J., Yang, L., Abbott, R. K., Callaghan, S., Garcia, E., Hurtado, J., Parren, M., Peng, L., Ramirez, S., Ricketts, J., Ricciardi, M. J., … Burton, D. R. (2020). Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science (New York, N.Y.)*, *369*(6506), 956–963. https://doi.org/10.1126/science.abc7520
Rohl, C.A., Strauss, C.E.M., Misura, K.M.S., and Baker, D. (2004). Protein structure prediction using Rosetta. Methods Enzymol. *383*, 66–93.

Ropp, P.J., Kaminsky, J.C., Yablonski, S. & Durrant, J.D. Dimorphite-DL: an open-source program for enumerating the ionization states of drug-like small molecules. *J. Cheminform.* 11, 14 (2019).
Rosenthal, P.B.,and Henderson, R.(2003).Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. *333*, 721–745.
Rozewicki, J., Li, S., Amada, K.M., Standley, D.M., and Katoh, K. (2019). MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. *47*, W5–W10.
Russo, C.J., and Passmore, L.A. (2014). Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. Science *346*, 1377–1380.
Ryu, D. K. et al. Therapeutic effect of CT-P59 against SARS-CoV-2 South African variant. *Biochem. Biophys. Res. Commun.* **566**, 135–140 (2021).
SA55 Injection Phase II Study in the Treatment of Mild/Moderate COVID-19 Patients (2023). Clinicaltrials.gov. www.clinicaltrials.gov/study/NCT06042764.
Sabir, J.S., Lam, T.T., Ahmed, M.M., Li, L., Shen, Y., Abo-Aba, S.E., Qureshi, M.I., Abu-Zeid, M., Zhang, Y., Khiyami, M.A., et al. (2016). Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. Science *351*, 81–84.
Sánchez, C.A., Li, H., Phelps, K.L. *et al.* A strategy to assess spillover risk of bat SARS-related coronaviruses in Southeast Asia. *Nat Commun* **13**, 4380 (2022). https://doi.org/10.1038/s41467-022-31860-w
Sauer, M.M., Tortorici, M.A., Park, Y.J., Walls, A.C., Homad, L., Acton, O.J., Bowen, J.E., Wang, C., Xiong, X., de van der Schueren, W., et al. (2021). Structural basis for broad coronavirus neutralization. Nat. Struct. Mol. Biol. *28*, 478–486.
Saunders, K. O., Lee, E., Parks, R., Martinez, D. R., Li, D., Chen, H., Edwards, R. J., Gobeil, S., Barr, M., Mansouri, K., Alam, S. M., Sutherland, L. L., Cai, F., Sanzone, A. M., Berry, M., Manne, K., Bock, K. W., Minai, M., Nagata, B. M., Kapingidza, A. B., … Haynes, B. F. (2021). Neutralizing antibody vaccine for pandemic and pre-emergent coronaviruses. *Nature*, *594*(7864), 553–559. https://doi.org/10.1038/s41586-021-03594-0.
Saunders, V. A., & Carter, J. (2007). *Virology*. John Wiley.
Schäfer, A., Muecksch, F., Lorenzi, J. C. C., Leist, S. R., Cipolla, M., Bournazos, S., Schmidt, F., Maison, R. M., Gazumyan, A., Martinez, D. R., Baric, R. S., Robbiani, D. F., Hatziioannou, T., Ravetch, J. V., Bieniasz, P. D., Bowen, R. A., Nussenzweig, M. C., & Sheahan, T. P. (2021). Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *The Journal of experimental medicine*, *218*(3), e20201993. https://doi.org/10.1084/jem.20201993
Scheres, S. H., & Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. *Nature methods*, *9*(9), 853–854. https://doi.org/10.1038/nmeth.2115
Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530.
Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Schwegmann-Weßels, C., Glende, J., Ren, X., Qu, X., Deng, H., Enjuanes, L., and Herrler, G. (2009). Comparison of vesicular stomatitis virus pseudotyped with the S proteins from a porcine and a human coronavirus. J. Gen. Virol. *90*, 1724–1729.

Seifert, S., et al. A sarbecovirus found in Russian bats uses human ACE2. bioRXIV. [Preprint]. (2022)

Seifert, S.N., Bai, S., Fawcett, S., Norton, E.B., Zwezdaryk, K.J., Robinson, J., Gunn, B., and Letko, M.C. (2022). An ACE2-dependent Sarbecovirus in Russian bats is resistant to SARS-CoV-2 vaccines. PLoS Pathog. *18*, e1010828. https://doi.org/10.1371/journal.ppat. 1010828.

Seydoux, E., Homad, L. J., MacCamy, A. J., Parks, K. R., Hurlburt, N. K., Jennewein, M. F., Akins, N. R., Stuart, A. B., Wan, Y. H., Feng, J., Nelson, R. E., Singh, S., Cohen, K. W., McElrath, M. J., Englund, J. A., Chu, H. Y., Pancera, M., McGuire, A. T., & Stamatatos, L. (2020). Characterization of neutralizing antibodies from a SARS-CoV-2 infected individual. *bioRxiv : the preprint server for biology*, 2020.05.12.091298. https://doi.org/10.1101/2020.05.12.091298

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., & Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, *581*(7807), 221–224. https://doi.org/10.1038/s41586-020-2179-y

Sheahan, T.P., Sims, A.C., Leist, S.R., Schäfer, A., Won, J., Brown, A.J., Montgomery, S.A., Hogg, A., Babusis, D., Clarke, M.O., et al. (2020). Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. Nat. Commun. *11*, 222.

Sheikh, A. et al. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *Lancet* **397**, 2461–2462 (2021).

Shi, R., Shan, C., Duan, X., Chen, Z., Liu, P., Song, J., Song, T., Bi, X., Han, C., Wu, L., Gao, G., Hu, X., Zhang, Y., Tong, Z., Huang, W., Liu, W. J., Wu, G., Zhang, B., Wang, L., Qi, J., … Yan, J. (2020). A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature*, *584*(7819), 120–124. https://doi.org/10.1038/s41586-020-2381-y

Shirts, M.R., Mobley, D.L., Chodera, J.D., and Pande, V.S. (2007). Accurate and efficient corrections for missing dispersion interactions in molecular simulations. J. Phys. Chem. B *111*, 13052–13063.

Shuai, H., Chan, J. F., Yuen, T. T., Yoon, C., Hu, J. C., Wen, L., Hu, B., Yang, D., Wang, Y., Hou, Y., Huang, X., Chai, Y., Chan, C. C., Poon, V. K., Lu, L., Zhang, R. Q., Chan, W. M., Ip, J. D., Chu, A. W., Hu, Y. F., … Chu, H. (2021). Emerging SARS-CoV-2 variants expand species tropism to murines. *EBioMedicine*, *73*, 103643. https://doi.org/10.1016/j.ebiom.2021.103643

Silva, R.P., Huang, Y., Nguyen, A.W., Hsieh, C.L., Olaluwoye, O.S., Kaoud, T.S., Wilen, R.E., Qerqez, A.N., Park, J.G., Khalil, A.M., et al. (2023). Identification of a conserved S2 epitope present on spike proteins from all highly pathogenic coronaviruses. eLife *12*. https://doi.org/10. 7554/eLife.83710.

Slater, G. S., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC bioinformatics*, *6*, 31. https://doi.org/10.1186/1471-2105-6-31

Snijder, J., Borst, A.J., Dosey, A., Walls, A.C., Burrell, A., Reddy, V.S., Kollman, J.M., and Veesler, D. (2017). Vitrification after multiple rounds of sample application and blotting improves particle density on cryo- electron microscopy grids. J. Struct. Biol. *198*, 38–42.

Sobolevsky, A.I., Rosconi, M.P. & Gouaux, E. X-ray structure of AMPA-subtype glutamate receptor: symmetry and mechanism. *Nature* 462, 745-756 (2009).

Soh, W.T. et al. (2020) The N-terminal domain of spike glycoprotein mediates SARS-CoV-2 infection by associating with L-SIGN and DC-SIGN. bioRxiv

Song, G., He, W. T., Callaghan, S., Anzanello, F., Huang, D., Ricketts, J., Torres, J. L., Beutler, N., Peng, L., Vargas, S., Cassell, J., Parren, M., Yang, L., Ignacio, C., Smith, D. M., Voss, J. E., Nemazee, D., Ward, A. B., Rogers, T., Burton, D. R., … Andrabi, R. (2021). Cross-reactive serum and memory B-cell responses to spike protein in SARS-CoV-2 and endemic coronavirus infection. *Nature communications*, *12*(1), 2938. https://doi.org/10.1038/s41467-021-23074-3

Song, H.-D. et al. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proc. Natl Acad. Sci. USA* 102, 2430–2435 (2005).

Song, W. *et al.* (2018). Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathlog.* **14**: e1007236.

Stalls, V., Lindenberger, J., Gobeil, S.M.-C., Henderson, R., Parks, R., Barr, M., Deyton, M., Martin, M., Janowska, K., Huang, X., et al. (2022). Cryo-EM structures of SARS-CoV-2 Omicron BA.2 spike. Cell Rep. *39*, 111009.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Stamatatos, L., Czartoski, J., Wan, Y.H., Homad, L.J., Rubin, V., Glantz, H., Neradilek, M., Seydoux, E., Jennewein, M.F., MacCamy, A.J., et al. (2021). mRNA vaccination boosts cross-variant neutralizing antibodies elicited by SARS-CoV-2 infection. Science *372*, 1413–1418. https://doi. org/10.1126/science.abg9175.

Stanford HVDB team (2023). Table 1: Virus variants and spike mutations vs monoclonal antibodies. Coronavirus Antiviral & Resistance Database. https://covdb.stanford.edu/susceptibility-data/table-mab-susc/.

Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci.* 25, 1204–1218 (2016).

Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J., Bowen, J.E., Tortorici, M.A., Walls, A.C., et al. (2020). Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. Cell *182*, 1295– 1310.e20.

Starr, T. N., Greaney, A. J., Addetia, A., Hannon, W. W., Choudhary, M. C., Dingens, A. S., Li, J. Z., & Bloom, J. D. (2021). Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science (New York, N.Y.)*, *371*(6531), 850–854. https://doi.org/10.1126/science.abf9302

Starr, T.N., Czudnochowski, N., Liu, Z., Zatta, F., Park, Y.J., Addetia, A., Pinto, D., Beltramello, M., Hernandez, P., Greaney, A.J., et al. (2021). SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. Nature *597*, 97–102. https://doi.org/10.1038/s41586-021- 03807-6.

Starr, T. N., Greaney, A. J., Dingens, A. S., & Bloom, J. D. (2021). Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell reports. Medicine*, *2*(4), 100255. https://doi.org/10.1016/j.xcrm.2021.100255

Starr, T.N., Zepeda, S.K., Walls, A.C., Greaney, A.J., Alkhovsky, S., Veesler, D., and Bloom, J.D. (2022a). ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. Nature *603*, 913–918.

Starr, T.N., Greaney, A.J., Stewart, C.M., Walls, A.C., Hannon, W.W., Veesler, D., and Bloom, J.D. (2022b). Deep mutational scans for ACE2 binding, RBD expression, and antibody escape in the SARS-CoV-2 Omicron BA.1 and BA.2 receptor-binding domains. PLoS Pathog. *18*, e1010951.

Starr, T.N., Greaney, A.J., Hannon, W.W., Loes, A.N., Hauser, K., Dillen, J.R., Ferri, E., Farrell, A.G., Dadonaite, B., McCallum, M., et al. (2022c). Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. Science *377*, 420–424. https://doi.org/ 10.1126/science.abo7896.

Stott, C.J., Sawattrakool, K., Saeng-Chuto, K., Tantituvanont, A., and Nilubol, D. (2022). The phylodynamics of emerging porcine deltacoronavirus in Southeast Asia. Transbound. Emerg. Dis. *69*, 2816–2827.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Leginon system. J. Struct. Biol. *151*, 41–60.

Sun, X., Yi, C., Zhu, Y., et al. (2022). Neutralization mechanism of a human antibody with pan-coronavirus reactivity including SARS-CoV-2. Nat Microbiol *7*, 1063–1074. https://doi.org/10.1038/s41564-022-01155-3.

Suribhatla, R., Starkey, T., Ionescu, M.C., Pagliuca, A., Richter, A., and Lee, L.Y.W. (2023). Systematic review and meta-analysis of the clinical effectiveness of tixagevimab/cilgavimab for prophylaxis of COVID-19 in immunocompromised patients. Br. J. Haematol. *201*, 813–823.

Suryadevara, N., Shrihari, S., Gilchuk, P., VanBlargan, L. A., Binshtein, E., Zost, S. J., Nargi, R. S., Sutton, R. E., Winkler, E. S., Chen, E. C., Fouch, M. E., Davidson, E., Doranz, B. J., Chen, R. E., Shi, P. Y., Carnahan, R. H., Thackray, L. B., Diamond, M. S., & Crowe, J. E., Jr (2021). Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. *Cell*, *184*(9), 2316–2331.e15. https://doi.org/10.1016/j.cell.2021.03.029

Taft, J.M., Weber, C.R., Gao, B., Ehling, R.A., Han, J., Frei, L., Metcalfe, S.W., Overath, M.D., Yermanos, A., Kelton, W., et al. (2022). Deep mutational learning predicts ACE2 binding and antibody escape to combinatorial mutations in the SARS-CoV-2 receptor-binding domain. Cell *185*, 4008-4022.e14.

Takada, A., Robison, C., Goto, H., Sanchez, A., Murti, K. G., Whitt, M. A., & Kawaoka, Y. (1997). A system for functional analysis of Ebola virus glycoprotein. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14764–14769. https://doi.org/10.1073/pnas.94.26.14764

Tan, C.W., Chia, W.N., Zhu, F., Young, B.E., Chantasrisawad, N., Hwa, S.H., Yeoh, A.Y.-Y., Lim, B.L., Yap, W.C., Pada, S.K.M.S., et al. (2022). SARS-CoV-2 Omicron variant emerged under immune selection. Nat. Microbiol. *7*, 1756–1761.

Tan, C.C.S., Trew, J., Peacock, T.P., Mok, K.Y., Hart, C., Lau, K., Ni, D., Orme, C.D.L., Ransome, E., Pearse, W.D., et al. (2023). Genomic screening of 16 UK native bat species through conservationist networks uncovers coronaviruses with zoonotic potential. Nat. Commun. *14*, 3322.

Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., and Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods *14*, 793–796.

Tao, K. et al. The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat. Rev. Genet.* 22, 757–773

Tao, Y., and Tong, S. (2019). Complete genome sequence of a severe acute respiratory syndrome-related coronavirus from Kenyan bats. Microbiol. Resour. Announc. *8*. https://doi.org/10.1128/MRA.00548-19.

Taylor, A.L., and Starr, T.N. (2023). Deep mutational scans of XBB.1.5 and BQ.1.1 reveal ongoing epistatic drift during SARS-CoV-2 evolution. PLoS Pathog. *19*, e1011901.

Taylor, P.C., Adams, A.C., Hufford, M.M., de la Torre, I., Winthrop, K., and Gottlieb, R.L. (2021). Neutralizing monoclonal antibodies for treatment of COVID-19. Nat. Rev. Immunol. *21*, 382–393.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E.J., Msomi, N., et al. (2021). Emergence of a SARS-CoV-2 variant of concern with mutations in spike glycoprotein. Nature *592*, 438–443. https://doi.org/10.1038/s41586- 021-03402-9.

Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E. J., Msomi, N., Mlisana, K., von Gottberg, A., Walaza, S., Allam, M., Ismail, A., Mohale, T., Glass, A. J., Engelbrecht, S., Van Zyl, G., Preiser, W., … de Oliveira, T. (2021). Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*, *592*(7854), 438–443. https://doi.org/10.1038/s41586-021-03402-9

Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. Nat. Methods *16*, 1146–1152.

Telenti, A. et al. After the pandemic: perspectives on the future trajectory of COVID-19. *Nature* **596**, 495–504 (2021).

Temmam, S., Vongphayloth, K., Baquero, E., Munier, S., Bonomi, M., Regnault, B., Douangboubpha, B., Karami, Y., Chre´ tien, D., Sanamxay, D., et al. (2022). Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. Nature *604*, 330–336. https://doi.org/10.1038/s41586- 022-04532-4.

ter Meulen, J., van den Brink, E. N., Poon, L. L., Marissen, W. E., Leung, C. S., Cox, F., Cheung, C. Y., Bakker, A. Q., Bogaards, J. A., van Deventer, E., Preiser, W., Doerr, H. W., Chow, V. T., de Kruif, J., Peiris, J. S., & Goudsmit, J. (2006). Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS medicine*, *3*(7), e237. https://doi.org/10.1371/journal.pmed.0030237

Terashi, G. & Kihara, D. De novo main-chain modeling for EM maps using MAINMAST. *Nat Commun* 9, 1618 (2018).
Terwilliger, T.C., Klei, H., Adams, P.D., Moriarty, N.W. & Cohn, J.D. Automated ligand fitting by core-fragment fitting and extension into density. *Acta Cryst.* D62, 915-922 (2006).
Terwilliger, T.C., Sobolev, O.V., Afonine, P.V. & Adams, P.D. Automated map sharpening by maximization of detail and connectivity. *Acta Cryst.* D74, 545-559. (2018).
Terwilliger, T.C., Adams, P.D., Afonine, P.V. & Sobolev, O.V. Cryo-EM map interpretation and protein model-building using iterative map segmentation. *Protein Science* 29, 87– 99 (2020).
Thépaut, M., Luczkowiak, J., Vivès, C., Labiod, N., Bally, I., Lasala, F., Grimoire, Y., Fenel, D., Sattin, S., Thielens, N., Schoehn, G., Bernardi, A., Delgado, R., & Fieschi, F. (2021). DC/L-SIGN recognition of spike glycoprotein promotes SARS-CoV-2 trans-infection and can be inhibited by a glycomimetic antagonist. *PLoS pathogens*, *17*(5), e1009576. https://doi.org/10.1371/journal.ppat.1009576
Thomson, E. C., Rosen, L. E., Shepherd, J. G., Spreafico, R., da Silva Filipe, A., Wojcechowskyj, J. A., Davis, C., Piccoli, L., Pascall, D. J., Dillen, J., Lytras, S., Czudnochowski, N., Shah, R., Meury, M., Jesudason, N., De Marco, A., Li, K., Bassi, J., O'Toole, A., Pinto, D., … Snell, G. (2021). Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell*, *184*(5), 1171–1187.e20. https://doi.org/10.1016/j.cell.2021.01.037
Thornton, J. W. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat. Rev. Genet.* 5, 366–375 (2004).
Toelzer, C., Gupta, K., Yadav, S.K.N., Borucu, U., Davidson, A.D., Kavanagh Williamson, M., Shoemark, D.K., Garzoni, F., Staufer, O., Milligan, R., et al. (2020). Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. Science *370*, 725–730.
Tolentino, J.E., Lytras, S., Ito, J., and Sato, K. (2024). Recombination analysis on the receptor switching event of MERS-CoV and its close relatives: implications for the emergence of MERS-CoV. Virol. J. *21*, 84.
Tong, S., Conrardy, C., Ruone, S., Kuzmin, I.V., Guo, X., Tao, Y., Niezgoda, M., Haynes, L., Agwanda, B., Breiman, R.F., et al. (2009). Detection of novel SARS-like and other coronaviruses in bats from Kenya. Emerg. Infect. Dis. *15*, 482–485.
Tortorici, M. A., & Veesler, D. (2019). Structural insights into coronavirus entry. *Advances in virus research*, *105*, 93–116. https://doi.org/10.1016/bs.aivir.2019.08.002
Tortorici, M. A. et al. Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* 26, 481–489 (2019).
Tortorici, M. A., Beltramello, M., Lempp, F. A., Pinto, D., Dang, H. V., Rosen, L. E., McCallum, M., Bowen, J., Minola, A., Jaconi, S., Zatta, F., De Marco, A., Guarino, B., Bianchi, S., Lauron, E. J., Tucker, H., Zhou, J., Peter, A., Havenar-Daughton, C., Wojcechowskyj, J. A., … Veesler, D. (2020). Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science (New York, N.Y.)*, *370*(6519), 950–957. https://doi.org/10.1126/science.abe3354

Tortorici, M. A., Czudnochowski, N., Starr, T. N., Marzi, R., Walls, A. C., Zatta, F., Bowen, J. E., Jaconi, S., Di Iulio, J., Wang, Z., De Marco, A., Zepeda, S. K., Pinto, D., Liu, Z., Beltramello, M., Bartha, I., Housley, M. P., Lempp, F. A., Rosen, L. E., Dellota, E., Jr, … Pizzuto, M. S. (2021). Broad sarbecovirus neutralization by a human monoclonal antibody. *Nature*, *597*(7874), 103–108. https://doi.org/10.1038/s41586-021-03817-4.

Tortorici, M.A., Walls, A.C., Joshi, A., Park, Y.-J., Eguia, R.T., Miranda, M.C., Kepl, E., Dosey, A., Stevens-Ayers, T., Boeckh, M.J., et al. (2022). Structure, receptor recognition, and antigenicity of the human coronavirus CCoV-HuPn-2018 spike glycoprotein. Cell *185*, 2279–2291.e17.

Tortorici, M.A., Choi, A., Gibson, C.A., Lee, J., Brown, J.T., Stewart, C., Joshi, A., Harari, S., Willoughby, I., Treichel, C., et al. (2024). Molecular basis of pathogenicity of the recently emerged FCoV-23 coronavirus. bioRxiv, 2024.08.25.607996.

Tortorici, M.A., Addetia, A., Seo, A.J., Brown, J., Sprouse, K., Logue, J., Clark, E., Franko, N., Chu, H., and Veesler, D. (2024). Persistent immune imprinting occurs after vaccination with the COVID-19 XBB.1.5 mRNA booster in humans. Immunity *57*, 904-911.e4.

Tostanoski, L. H., Wegmann, F., Martinot, A. J., Loos, C., McMahan, K., Mercado, N. B., Yu, J., Chan, C. N., Bondoc, S., Starke, C. E., Nekorchuk, M., Busman-Sahay, K., Piedra-Mora, C., Wrijil, L. M., Ducat, S., Custers, J., Atyeo, C., Fischinger, S., Burke, J. S., Feldman, J., … Barouch, D. H. (2020). Ad26 vaccine protects against SARS-CoV-2 severe clinical disease in hamsters. *Nature medicine*, *26*(11), 1694–1700. https://doi.org/10.1038/s41591-020-1070-6.

Tse, L.V., Hou, Y.J., McFadden, E., Lee, R.E., Scobey, T.D., Leist, S.R., Martinez, D.R., Meganck, R.M., Schäfer, A., Yount, B.L., et al. (2023). A MERS-CoV antibody neutralizes a pre-emerging group 2c bat coronavirus. Sci. Transl. Med. *15*, eadg5567.

U.S. Department of Health and Human Services COVID-19 Therapeutics Thresholds, Orders, and Replenishment. https://aspr.hhs.gov/COVID-19/Therapeutics/Orders/.

VanBlargan, L. A., Adams, L. J., Liu, Z., Chen, R. E., Gilchuk, P., Raju, S., Smith, B. K., Zhao, H., Case, J. B., Winkler, E. S., Whitener, B. M., Droit, L., Aziati, I. D., Bricker, T. L., Joshi, A., Shi, P. Y., Creanga, A., Pegu, A., Handley, S. A., Wang, D., … Diamond, M. S. (2021). A potently neutralizing SARS-CoV-2 antibody inhibits variants of concern by utilizing unique binding residues in a highly conserved epitope. *Immunity*, *54*(10), 2399–2416.e6. https://doi.org/10.1016/j.immuni.2021.08.016

VanBlargan, L. A. et al. An infectious SARS-CoV-2 B.1.1.529 Omicron virus escapes neutralization by therapeutic monoclonal antibodies. *Nat. Med*. https://doi.org/10.1038/s41591-021-01678-y (2022).

Vant, J.W. et al. Flexible fitting of small-molecules into electron microscopy maps using molecular dynamics simulations with neural network potentials. *J. Chem. Inf. Model* 60, 2591-2604 (2020).

van Regenmortel, M. H., & Mahy, B. W. (2004). Emerging issues in virus taxonomy. *Emerging infectious diseases*, *10*(1), 8–13. https://doi.org/10.3201/eid1001.030279

van Zundert, G.C.P., Moriarty, N.W., Sobolev, O.V., Adams, P.D. & Borrelli, K.W. Macromolecular refinement of X-ray and cryoelectron microscopy structures with Phenix/OPLS3e for improved structure and ligand quality. *Structure* 29, 913-921 (2021).

Viana, R., Moyo, S., Amoako, D. G., Tegally, H., Scheepers, C., Althaus, C. L., Anyaneji, U. J., Bester, P. A., Boni, M. F., Chand, M., Choga, W. T., Colquhoun, R., Davids, M., Deforche, K., Doolabh, D., du Plessis, L., Engelbrecht, S., Everatt, J., Giandhari, J., Giovanetti, M., … de Oliveira, T. (2022). Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*, *603*(7902), 679–686. https://doi.org/10.1038/s41586-022-04411-y

Vilas, J.L. et al. MonoRes: Automatic and accurate estimation of local resolution for electron microscopy maps. Structure 26, 337-344 (2018).

Voss, W. N., Hou, Y. J., Johnson, N. V., Delidakis, G., Kim, J. E., Javanmardi, K., Horton, A. P., Bartzoka, F., Paresi, C. J., Tanno, Y., Chou, C. W., Abbasi, S. A., Pickens, W., George, K., Boutz, D. R., Towers, D. M., McDaniel, J. R., Billick, D., Goike, J., Rowe, L., … Ippolito, G. C. (2021). Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. *Science (New York, N.Y.)*, *372*(6546), 1108–1112. https://doi.org/10.1126/science.abg5268

Wacharapluesadee, S., Tan, C.W., Maneeorn, P., Duengkae, P., Zhu, F., Joyjinda, Y., Kaewpom, T., Chia, W.N., Ampoot, W., Lim, B.L., et al. (2021). Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. Nat. Commun. *12*. https://doi. org/10.1038/s41467-021-21240-1.

Walls, A. *et al*. (2016) Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol,* **23**: 899-905.

Walls, A. C., Tortorici, M. A., Bosch, B. J., Frenz, B., Rottier, P. J. M., DiMaio, F., Rey, F. A., & Veesler, D. (2016). Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature*, *531*(7592), 114–117. https://doi.org/10.1038/nature16988

Walls, A.C., Tortorici, M. A., Snijder, J., Xiong, X., Bosch, B. J., Rey, F. A., Veesler, D. (2017) Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11157–11162.

Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al. (2019). Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. Cell *176*, 1026–1039.e15

Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020). Structure, function, and antigenicity of the SARS- CoV-2 spike glycoprotein. Cell *181*, 281.

Walls, A. C., Fiala, B., Schäfer, A., Wrenn, S., Pham, M. N., Murphy, M., Tse, L. V., Shehata, L., O'Connor, M. A., Chen, C., Navarro, M. J., Miranda, M. C., Pettie, D., Ravichandran, R., Kraft, J. C., Ogohara, C., Palser, A., Chalk, S., Lee, E. C., Guerriero, K., … King, N. P. (2020). Elicitation of Potent Neutralizing Antibody

Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell*, *183*(5), 1367–1382.e17. https://doi.org/10.1016/j.cell.2020.10.043

Walls, A.C., Miranda, M.C., Schäfer, A, Pham, M.N., Greaney, A., Arunachalam, P.S., Navarro, M.J., Tortorici, M.A., Rogers, K., O'Connor, M.A., et al. (2021). Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines. Cell *184*, 5432–5447.e16. https://doi.org/10.1016/j.cell.2021.09.015.

Walker, L., Kaku, C.I., Schmidt, P., Dillinger, L., and Garrett RAPPAZZO, C. (2024). Sars-cov2 antibodies and uses thereof. World Patent.

Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.* 94, e00127–20 (2020).

Wang, C., Li, W., Drabek, D., Okba, N. M. A., van Haperen, R., Osterhaus, A. D. M. E., van Kuppeveld, F. J. M., Haagmans, B. L., Grosveld, F., & Bosch, B. J. (2020). A human monoclonal antibody blocking SARS-CoV-2 infection. *Nature communications*, *11*(1), 2251. https://doi.org/10.1038/s41467-020-16256-y

Wang, C., van Haperen, R., Gutiérrez-Álvarez, J., Li, W., Okba, N. M. A., Albulescu, I., Widjaja, I., van Dieren, B., Fernandez-Delgado, R., Sola, I., Hurdiss, D. L., Daramola, O., Grosveld, F., van Kuppeveld, F. J. M., Haagmans, B. L., Enjuanes, L., Drabek, D., & Bosch, B. J. (2021). A conserved immunogenic and vulnerable site on the coronavirus spike protein delineated by cross-reactive monoclonal antibodies. *Nature communications*, *12*(1), 1715. https://doi.org/10.1038/s41467-021-21968-w

Wang, L., Wu, D., Robinson, C.V., Wu, H. & Fu, T.M. Structures of a complete human V-ATPase reveal mechanisms of its assembly. *Mol. Cell* 80, 501-511 (2020).

Wang, J., Wang, W., Kollman P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling* 25, 247260 (2006).

Wang, J., Pan, Y.-F., Yang, L.-F., Yang, W.-H., Luo, C.-M., Wang, J., Kuang, G.-P., Wu, W.-C., Gou, Q.-Y., Xin, G.-Y., et al. (2022). Individual bat viromes reveal the co-infection, spillover and emergence risk of potential zoonotic viruses. bioRxiv, 2022.11.23.517609. 10.1101/2022.11.23.517609.

Wang, L., Shi, W., Chappell, J. D., Joyce, M. G., Zhang, Y., Kanekiyo, M., Becker, M. M., van Doremalen, N., Fischer, R., Wang, N., Corbett, K. S., Choe, M., Mason, R. D., Van Galen, J. G., Zhou, T., Saunders, K. O., Tatti, K. M., Haynes, L. M., Kwong, P. D., Modjarrad, K., … Graham, B. S. (2018). Importance of Neutralizing Monoclonal Antibodies Targeting Multiple Antigenic Sites on the Middle East Respiratory Syndrome Coronavirus Spike Glycoprotein To Avoid Neutralization Escape. *Journal of virology*, *92*(10), e02002-17. https://doi.org/10.1128/JVI.02002-17.

Wang, M., Yan, M., Xu, H., Liang, W., Kan, B., Zheng, B., Chen, H., Zheng, H., Xu, Y., Zhang, E., et al. (2005). SARS-CoV infection in a restaurant from palm civet. Emerg. Infect. Dis. *11*, 1860–1865.

Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., et al. (2013). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. Cell Res. *23*, 986–993.

Wang, N., Li, S.Y., Yang, X.L., Huang, H.M., Zhang, Y.J., Guo, H., Luo, C.M., Miller, M., Zhu, G., Chmura, A.A., et al. (2018). Serological evidence of bat SARS-related coronavirus infection in humans, China. Virol. Sin. *33*, 104–107.

Wang, N., Rosen, O., Wang, L., Turner, H. L., Stevens, L. J., Corbett, K. S., Bowman, C. A., Pallesen, J., Shi, W., Zhang, Y., Leung, K., Kirchdoerfer, R. N., Becker, M. M., Denison, M. R., Chappell, J. D., Ward, A. B., Graham, B. S., & McLellan, J. S. (2019). Structural Definition of a Neutralization-Sensitive Epitope on the MERS-CoV S1-NTD. *Cell reports*, *28*(13), 3395–3405.e6. https://doi.org/10.1016/j.celrep.2019.08.052.

Wang, P., Nair, M. S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P. D., Graham, B. S., Mascola, J. R., Chang, J. Y., Yin, M. T., Sobieszczyk, M., Kyratsous, C. A., Shapiro, L., Sheng, Z., Huang, Y., & Ho, D. D. (2021). Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature*, *593*(7857), 130–135. https://doi.org/10.1038/s41586-021-03398-2.

Wang, Q., Qi, J., Yuan, Y., Xuan, Y., Han, P., Wan, Y., Ji, W., Li, Y., Wu, Y., Wang, J., et al. (2014). Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. Cell Host Microbe *16*, 328–337.

Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K. Y., Wang, Q., Zhou, H., Yan, J., & Qi, J. (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*, *181*(4), 894–904.e9. https://doi.org/10.1016/j.cell.2020.03.045.

Wang, Q., Guo, Y., Zhang, R.M., Ho, J., Mohri, H., Valdez, R., Manthei, D.M., Gordon, A., Liu, L., and Ho, D.D. (2023). Antibody neutralisation of emerging SARS-CoV-2 subvariants: EG.5.1 and XBC.1.6. Lancet Infect. Dis. *23*, e397–e398.

Wang, Q., Guo, Y., Liu, L., Schwanz, L.T., Li, Z., Nair, M.S., Ho, J., Zhang, R.M., Iketani, S., Yu, J., et al. (2023). Antigenicity and receptor affinity of SARS-CoV-2 BA.2.86 spike. Nature. https://doi.org/10.1038/s41586-023-06750-w.

Wang, R.Y.R. et al. De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nat Methods* 12, 335–338 (2015).

Wang, R.Y., Song, Y., Barad, B.A., Cheng, Y., Fraser, J.S., and DiMaio, F. (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. eLife *5*. https://doi.org/10.7554/ eLife.17219.

Wang X, Zhou Y, Liu L, Ma J, Wu H, Zhao L, Chen W, Li SC, Ran Q, Li Z. Coronavirus disease 2019 (COVID-19): diagnosis and prognosis. Blood Genom. 4, 96-107. (2020).

Wang, Z., Schmidt, F., Weisblum, Y., Muecksch, F., Barnes, C. O., Finkin, S., Schaefer-Babajew, D., Cipolla, M., Gaebler, C., Lieberman, J. A., Oliveira, T. Y., Yang, Z., Abernathy, M. E., Huey-Tubman, K. E., Hurley, A., Turroja, M., West, K. A., Gordon, K., Millard, K. G., Ramos, V., … Nussenzweig, M. C. (2021). mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature*, *592*(7855), 616–622. https://doi.org/10.1038/s41586-021-03324-6

Wang, Z., Muecksch, F., Cho, A., Gaebler, C., Hoffmann, H.-H., Ramos, V., Zong, S., Cipolla, M., Johnson, B., Schmidt, F., et al. (2022). Analysis of memory B cells

identifies conserved neutralizing epitopes on the N-terminal domain of variant SARS-Cov-2 spike proteins. Immunity *55*, 998–1012.e8.
Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., & Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science (New York, N.Y.)*, *369*(6501), 330–333. https://doi.org/10.1126/science.abb9983
Wec, A. Z., Wrapp, D., Herbert, A. S., Maurer, D. P., Haslwanter, D., Sakharkar, M., Jangra, R. K., Dieterle, M. E., Lilov, A., Huang, D., Tse, L. V., Johnson, N. V., Hsieh, C. L., Wang, N., Nett, J. H., Champney, E., Burnina, I., Brown, M., Lin, S., Sinclair, M., … Walker, L. M. (2020). Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science (New York, N.Y.)*, *369*(6504), 731–736. https://doi.org/10.1126/science.abc7424
Whelan, S.P., Ball, L.A., Barr, J.N., and Wertz, G.T. (1995). Efficient recovery of infectious vesicular stomatitis virus entirely from cDNA clones. Proc. Natl. Acad. Sci. U. S. A. *92*, 8388–8392.
Weinreich, D. M. et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with Covid-19. *N. Engl. J. Med.***384**, 238–251 (2021).
Weisblum, Y., Schmidt, F., Zhang, F., DaSilva, J., Poston, D., Lorenzi, J. C., Muecksch, F., Rutkowska, M., Hoffmann, H. H., Michailidis, E., Gaebler, C., Agudelo, M., Cho, A., Wang, Z., Gazumyan, A., Cipolla, M., Luchsinger, L., Hillyer, C. D., Caskey, M., Robbiani, D. F., … Bieniasz, P. D. (2020). Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife*, *9*, e61312. https://doi.org/10.7554/eLife.61312
Wells, H.L., Letko, M., Lasso, G., Ssebide, B., Nziza, J., Byarugaba, D.K., Navarrete-Macias, I., Liang, E., Cranfield, M., Han, B.A., et al. (2021). The evolutionary history of ACE2 usage within the coronavirus subgenus Sarbecovirus. Virus Evol. *7*, veab007.
Wentz, A.E., and Shusta, E.V. (2007). A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins. Appl. Environ. Microbiol. *73*, 1189–1198.
Westendorf, K., Zentelis, S., Wang, L., Foster, D., Vaillancourt, P., Wiggin, M., Lovett, E., van der Lee, R., Hendle, J., Pustilnik, A., et al. (2022). LY-CoV1404 (bebtelovimab) potently neutralizes SARS-CoV-2 variants. Cell Rep. *39*, 110812.
Wibmer, C. K., Ayres, F., Hermanus, T., Madzivhandila, M., Kgagudi, P., Oosthuysen, B., Lambson, B. E., de Oliveira, T., Vermeulen, M., van der Berg, K., Rossouw, T., Boswell, M., Ueckermann, V., Meiring, S., von Gottberg, A., Cohen, C., Morris, L., Bhiman, J. N., & Moore, P. L. (2021). SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nature medicine*, *27*(4), 622–625. https://doi.org/10.1038/s41591-021-01285-x
Wickham, H. ggplot2: Elegant graphics for data analysis. *Springer-Verlag New York* (2016). https://ggplot2.tidyverse.org
Wilhelm, A. et al. Reduced neutralization of SARS-CoV-2 Omicron variant by vaccine sera and monoclonal antibodies. Preprint at https://doi.org/10.1101/2021.12.07.21267432 (2021).
Willett, B.J., Grove, J., MacLean, O.A., Wilkie, C., De Lorenzo, G., Furnon, W., Cantoni, D., Scott, S., Logan, N., Ashraf, S., et al. (2022). SARS-CoV-2 Omicron is an

immune escape variant with an altered cell entry pathway. Nat. Microbiol. *7*, 1161–1179.
Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., and Nagarajan, N. (2012). LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. *40*, 11189–11201.
Winkler,E.S.,Gilchuk,P.,Yu,J.,Bailey,A.L.,Chen,R.E.,Chong,Z.,Zost, S.J., Jang, H., Huang, Y., Allen, J.D., et al. (2021). Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. Cell *184*, 1804–1820.e16.
Wong, A.H., and Rini, J.M. (2017). Crystal structure of the human coronavirus 229E spike protein receptor binding domain in complex with human aminopeptidase N. Preprint at Worldwide Protein Data Bank, https://doi.org/10.2210/pdb6atk/pdb https://doi.org/10.2210/pdb6atk/pdb.
Wong et al (2017) Receptor-binding loops in alphacoronavirus adaptation and evolution. *Nat Commun* **8:** 1735. https://doi.org/10.1038/s41467-017-01706-x
Woo, P.C., Lau, S.K., Li, K.S., Poon, R.W., Wong, B.H., Tsoi, H.W., Yip, B.C., Huang, Y., Chan, K.H., and Yuen, K.Y. (2006). Molecular diversity of coronaviruses in bats. Virology *351*, 180–187.
World Health Organization (2004) Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. https://www.who.int/csr/sars/country/table2004_04_21/en/.
World Health Organization (2019) Middle East respiratory syndrome coronavirus (MERS-CoV), https://www.who.int/emergencies/mers-cov/en/.
World Health Organization (2020) Coronavirus disease (COVID-19) pandemic.
Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367, 1260–1263 (2020).
Wrobel, A. G., Benton, D. J., Hussain, S., Harvey, R., Martin, S. R., Roustan, C., Rosenthal, P. B., Skehel, J. J., & Gamblin, S. J. (2020). Antibody-mediated disruption of the SARS-CoV-2 spike glycoprotein. *Nature communications*, *11*(1), 5337. https://doi.org/10.1038/s41467-020-19146-5
Wrobel, A. G., Benton, D. J., Xu, P., Roustan, C., Martin, S. R., Rosenthal, P. B., Skehel, J. J., & Gamblin, S. J. (2020). SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nature structural & molecular biology*, *27*(8), 763–767. https://doi.org/10.1038/s41594-020-0468-7
Wrobel, A.G., Benton, D.J., Xu, P., Calder, L.J., Borg, A., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2021). Structure and binding properties of Pangolin-CoV spike glycoprotein inform the evolution of SARS-CoV-2. Nat. Commun. *12*, 837.
Wu, K., Li, W., Peng, G., & Li, F. (2009). Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(47), 19970–19974. https://doi.org/10.1073/pnas.0908837106
Wu, K., Werner, A. P., Koch, M., Choi, A., Narayanan, E., Stewart-Jones, G. B. E., Colpitts, T., Bennett, H., Boyoglu-Barnum, S., Shi, W., Moliva, J. I., Sullivan, N.

J., Graham, B. S., Carfi, A., Corbett, K. S., Seder, R. A., & Edwards, D. K. (2021). Serum Neutralizing Activity Elicited by mRNA-1273 Vaccine. *The New England journal of medicine*, *384*(15), 1468–1470. https://doi.org/10.1056/NEJMc2102179

Wu, Y., Wang, F., Shen, C., Peng, W., Li, D., Zhao, C., Li, Z., Li, S., Bi, Y., Yang, Y., Gong, Y., Xiao, H., Fan, Z., Tan, S., Wu, G., Tan, W., Lu, X., Fan, C., Wang, Q., Liu, Y., … Liu, L. (2020). A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science (New York, N.Y.)*, *368*(6496), 1274–1278. https://doi.org/10.1126/science.abc2241

Xia, S., Yan, L., Xu, W., Agrawal, A.S., Algaissi, A., Tseng, C.-T.K., Wang, Q., Du, L., Tan, W., Wilson, I.A., et al. (2019). A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. Sci Adv *5*, eaav4580.

Xia, S., Liu, M., Wang, C., Xu, W., Lan, Q., Feng, S., Qi, F., Bao, L., Du, L., Liu, S., et al. (2020). Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. Cell Res. *30*, 343–355.

Xiao, K. et al. (2020) Isolation of SARS-CoV-2-related Coronavirus from Malayan Pangolins. Nature. doi: 10.1038/s41586-020-2313-x.

Xie, X., Muruato, A. E., Zhang, X., Lokugamage, K. G., Fontes-Garfias, C. R., Zou, J., Liu, J., Ren, P., Balakrishnan, M., Cihlar, T., Tseng, C. K., Makino, S., Menachery, V. D., Bilello, J. P., & Shi, P. Y. (2020). A nanoluciferase SARS-CoV-2 for rapid neutralization testing and screening of anti-infective drugs for COVID-19. *Nature communications*, *11*(1), 5214. https://doi.org/10.1038/s41467-020-19055-7

Xiao, P. et al. Ligand recognition and allosteric regulation of DRD1-Gs signaling complexes. *Cell* 184, 943-956 (2021).

Xiong, Q., Cao, L., Ma, C., Tortorici, M.A., Liu, C., Si, J., Liu, P., Gu, M., Walls, A.C., Wang, C., et al. (2022). Close relatives of MERS-CoV in bats use ACE2 as their functional receptors. Nature *612*, 748–757.

Xiong, X., Corti, D., Liu, J., Pinna, D., Foglierini, M., Calder, L. J., Martin, S. R., Lin, Y. P., Walker, P. A., Collins, P. J., Monne, I., Suguitan, A. L., Jr, Santos, C., Temperton, N. J., Subbarao, K., Lanzavecchia, A., Gamblin, S. J., & Skehel, J. J. (2015). Structures of complexes formed by H5 influenza hemagglutinin with a potent broadly neutralizing human monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(30), 9430–9435. https://doi.org/10.1073/pnas.1510816112

Xiong, X., Qu, K., Ciazynska, K.A., Hosmillo, M., Carter, A.P., Ebrahimi, S., Ke, Z., Scheres, S.H.W., Bergamaschi, L., Grice, G.L., et al. (2020). A thermostable, closed SARS-CoV-2 spike protein trimer. Nat. Struct. Mol. Biol. https://doi.org/10.1038/s41594-020-0478-5.

Yamin, R., Jones, A. T., Hoffmann, H. H., Schäfer, A., Kao, K. S., Francis, R. L., Sheahan, T. P., Baric, R. S., Rice, C. M., Ravetch, J. V., & Bournazos, S. (2021). Fc-engineered antibody therapeutics with improved anti-SARS-CoV-2 efficacy. *Nature*, *599*(7885), 465–470. https://doi.org/10.1038/s41586-021-04017-w

Yan, H., Jiao, H., Liu, Q., Zhang, Z., Xiong, Q., Wang, B.-J., Wang, X., Guo, M., Wang, L.-F., Lan, K., et al. (2021). ACE2 receptor usage reveals variation in susceptibility to SARS-CoV and SARS-CoV-2 infection among bat species. Nat Ecol Evol *5*, 600–608.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science (New York, N.Y.)*, *367*(6485), 1444–1448. https://doi.org/10.1126/science.abb2762.

Yang, M., Li, T., Wang, Y., Tran, C., Zhao, S., and Ao, G. (2022). Monoclonal antibody therapy improves severity and mortality of COVID-19 in organ transplant recipients: A meta-analysis. J. Infect. *85*, 436–480.

Yang, S., Yu, Y., Jian, F., Song, W., Yisimayi, A., Chen, X., Xu, Y., Wang, P., Wang, J., Yu, L., et al. (2023). Antigenicity and infectivity characterisation of SARS-CoV-2 BA.2.86. Lancet Infect. Dis. *23*, e457–e459.

Yang, S., Yu, Y., Xu, Y., Jian, F., Song, W., Yisimayi, A., Wang, P., Wang, J., Liu, J., Yu, L., et al. (2023). Fast evolution of SARS-CoV-2 BA.2.86 to JN.1 under heavy immune pressure. Lancet Infect. Dis. 10.1016/s1473-3099(23)00744-2.

Yang, X.L., Hu, B., Wang, B., Wang, M.N., Zhang, Q., Zhang, W., Wu, L.J., Ge, X.Y., Zhang, Y.Z., Daszak, P., et al. (2015). Isolation and characterization of a novel bat coronavirus closely related to the direct progenitor of severe acute respiratory syndrome coronavirus. J. Virol. *90*, 3253–3256.

Yang, Y., Liu, C., Du, L., Jiang, S., Shi, Z., Baric, R.S., and Li, F. (2015). Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. J. Virol. *89*, 9119–9123.

Yariv, B., Yariv, E., Kessel, A., Masrati, G., Chorin, A.B., Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. (2023). Using evolutionary data to make sense of macromolecules with a “face-lifted” ConSurf. Protein Sci. *32*, e4582.

Ye, X., Chen, Y., Zhu, X., Guo, J., Xie, D., Hou, Z., Xu, S., Zhou, J., Fang, L., Wang, D., et al. (2020). Cross-species transmission of deltacoronavirus and the origin of porcine deltacoronavirus. Evol. Appl. *13*, 2246–2253.

Yin, Y. et al. Structural basis of cooling agent and lipid sensing by the cold-activated TRPM8 channel. *Science* 363, eaav9334 (2019).

Yin, Y. et al. Structural basis for aggregate dissolution and refolding by the Mycobacterium tuberculosis ClpB-DnaK bi-chaperone system. *Cell Rep.* 35, 109166 (2021).

Yip, K.M., Fischer, N., Paknia, E., Chari, A. & Holger, S. Atomic-resolution protein structure determination by cryo-EM. *Nature* 587, 157–161 (2020).

Yu, J., Tostanoski, L. H., Peter, L., Mercado, N. B., McMahan, K., Mahrokhian, S. H., Nkolola, J. P., Liu, J., Li, Z., Chandrashekar, A., Martinez, D. R., Loos, C., Atyeo, C., Fischinger, S., Burke, J. S., Slein, M. D., Chen, Y., Zuiani, A., Lelis, F. J. N., Travers, M., … Barouch, D. H. (2020). DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science (New York, N.Y.)*, *369*(6505), 806–811. https://doi.org/10.1126/science.abc6284.

Yu, J. et al. Hippocampal AMPA receptor assemblies and mechanism of allosteric inhibition. *Nature* 594, 448-453 (2021).

Yu P, Hu B, Shi ZL, and Cui J. (2019).zGeographical structure of bat SARS-related coronaviruses. *Infect Genet Evol*. 69:224-229.

Yue, C., Song, W., Wang, L., Jian, F., Chen, X., Gao, F., Shen, Z., Wang, Y., Wang, X., and Cao, Y. (2023). ACE2 binding and antibody evasion in enhanced transmissibility of XBB.1.5. Lancet Infect. Dis. https://doi.org/10.1016/S1473-3099(23)00010-5.

Yuan, M., Wu, N. C., Zhu, X., Lee, C. D., So, R. T. Y., Lv, H., Mok, C. K. P., & Wilson, I. A. (2020). A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science (New York, N.Y.)*, *368*(6491), 630–633. https://doi.org/10.1126/science.abb7269

Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., et al. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nat. Commun. *8*, 15092.

Yurkovetskiy, L., Wang, X., Pascal, K. E., Tomkins-Tinch, C., Nyalile, T. P., Wang, Y., Baum, A., Diehl, W. E., Dauphin, A., Carbone, C., Veinotte, K., Egri, S. B., Schaffner, S. F., Lemieux, J. E., Munro, J. B., Rafique, A., Barve, A., Sabeti, P. C., Kyratsous, C. A., Dudkina, N. V., … Luban, J. (2020). Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell*, *183*(3), 739–751.e8. https://doi.org/10.1016/j.cell.2020.09.032

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., and Fouchier, R.A. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med. *367*, 1814–1820.

Zang, R., Gomez Castro, M. F., McCune, B. T., Zeng, Q., Rothlauf, P. W., Sonnek, N. M., Liu, Z., Brulois, K. F., Wang, X., Greenberg, H. B., Diamond, M. S., Ciorba, M. A., Whelan, S. P. J., & Ding, S. (2020). TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Science immunology*, *5*(47), eabc3582.

Zhang, D., Watson, J.F., Matthews, P.M., Cais, O. & Greger, I.H. Gating and modulation of a hetero-octameric AMPA glutamate receptor. *Nature* 594, 454-458 (2021).

Zhang, L., Cao, L., Gao, X. S., Zheng, B. Y., Deng, Y. Q., Li, J. X., Feng, R., Bian, Q., Guo, X. L., Wang, N., Qiu, H. Y., Wang, L., Cui, Z., Ye, Q., Chen, G., Lu, K. K., Chen, Y., Chen, Y. T., Pan, H. X., Yu, J., … Zhu, F. C. (2021). A proof of concept for neutralizing antibody-guided vaccine design against SARS-CoV-2. *National science review*, *8*(8), nwab053. https://doi.org/10.1093/nsr/nwab053.

Zhang, S., Liang, Q., He, X., Zhao, C., Ren, W., Yang, Z., Wang, Z., Ding, Q., Deng, H., Wang, T., et al. (2022). Loss of Spike N370 glycosylation as an important evolutionary event for the enhanced infectivity of SARS- CoV-2. Cell Res. *32*, 315–318.

Zhao, J., Wan, W., Yu, K., Lemey, P., Pettersson, J.H.-O., Bi, Y., Lu, M., Li, X., Chen, Z., Zheng, M., et al. (2024). Farmed fur animals harbour viruses with zoonotic spillover potential. Nature. https://doi.org/10.1038/s41586-024-07901-3.

Zhou, D., Ren, J., Fry, E.E., and Stuart, D.I. (2023). Broadly neutralizing antibodies against COVID-19. Curr. Opin. Virol. *61*, 101332.

Zhou, H., Chen, Y., Zhang, S., Niu, P., Qin, K., Jia, W., Huang, B., Zhang, S., Lan, J., Zhang, L., Tan, W., & Wang, X. (2019). Structural definition of a neutralization epitope on the N-terminal domain of MERS-CoV spike glycoprotein. *Nature communications*, *10*(1), 3068. https://doi.org/10.1038/s41467-019-10897-4

Zhou, H., *et al.* (2020) A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr Biol.* **30(11)**: 2196-2203.

Zhou, H., Ji, J., Chen, X., Bi, Y., Li, J., Wang, Q., Hu, T., Song, H., Zhao, R., Chen, Y., et al. (2021). Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. Cell *184*, 4380–4391.e14.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature *579*, 270–273. https://doi.org/10.1038/s41586-020-2012-7.

Zhou, P., Yuan, M., Song, G., Beutler, N., Shaabani, N., Huang, D., He, W. T., Zhu, X., Callaghan, S., Yong, P., Anzanello, F., Peng, L., Ricketts, J., Parren, M., Garcia, E., Rawlings, S. A., Smith, D. M., Nemazee, D., Teijaro, J. R., Rogers, T. F., … Andrabi, R. (2022). A human antibody reveals a conserved site on beta-coronavirus spike proteins and confers protection against SARS-CoV-2 infection. *Science translational medicine*, *14*(637), eabi9215. https://doi.org/10.1126/scitranslmed.abi9215.

Zhou, P., Song, G., Liu, H., Yuan, M., He, W.-T., Beutler, N., Zhu, X., Tse, L.V., Martinez, D.R., Schäfer, A., et al. (2023). Broadly neutralizing anti-S2 antibodies protect against all three human betacoronaviruses that cause deadly disease. Immunity *56*, 669–686.e7.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. *382*, 727–733. https://doi. org/10.1056/NEJMoa2001017.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*. https://doi. org/10.7554/eLife.42166.

Zivanov, J., Nakane, T., and Scheres, S.H.W. (2019). A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. IUCrJ *6*, 5–17.

Zost, S. J., Gilchuk, P., Case, J. B., Binshtein, E., Chen, R. E., Nkolola, J. P., Schäfer, A., Reidy, J. X., Trivette, A., Nargi, R. S., Sutton, R. E., Suryadevara, N., Martinez, D. R., Williamson, L. E., Chen, E. C., Jones, T., Day, S., Myers, L., Hassan, A. O., Kafai, N. M., … Crowe, J. E., Jr (2020). Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature*, *584*(7821), 443–449. https://doi.org/10.1038/s41586-020-2548-6

Zwart, P.H., Langer, G.G. & Lamzin, V.S. Modelling bound ligands in protein crystal structures. *Acta Cryst.* D60, 2230-2239 (2004).