Structural Dynamic Insights into Coronavirus Spike Conformational States, Receptor Activation, and Vaccines

Chengbo Chen

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Kelly Lee, Chair

Abhinav Nath

Miklos Guttman

Program Authorized to Offer Degree:

Medicinal Chemistry

University of Washington

Abstract

Structural Dynamic Insights into Coronavirus Spike Conformational States, Receptor Activation, and Vaccines

Chengbo Chen

Chair of the Supervisory Committee:

Kelly Lee

Medicinal Chemistry

Viruses in the *Sarbecovirus* subgenus have given rise to two highly transmissible coronaviruses in recent human history: severe acute respiratory syndrome coronavirus (SARS-CoV) and SARS-CoV-2. These viruses enter human cells through the binding between viral spike (S) glycoprotein and a common human angiotensin-converting enzyme 2 (hACE2) receptor. However, they exhibit differences in interactions with hACE2 as well as in proteolytic processing of S that trigger the fusion machinery. Identifying the molecular basis of how these differences impact S activation as well as the effects of mutations found in novel SARS-CoV-2 variants of concern (VOCs) is key to understand S function and viral pathogenesis phenotypes.

The hypothesis I am testing is that differences in structural and conformational dynamics in SARS-CoV and SARS-CoV-2 spike trimers influence their ability to bind and be activated by the hACE2 receptor. To probe the structural and dynamic differences among SARS-CoV, SARS-CoV-2 and VOCs that exhibit different transmissibility, we perform hydrogen/deuterium-

exchange mass spectrometry (HDX-MS), which measures protein dynamics under native conditions. HDX-MS reveals differences in spike dynamics at various levels, which will be discussed in three chapters with specific focus.

In Chapter 2, HDX-MS reveals differences in dynamics of unbound S, featuring the D614G mutation-induced S conformational switch to open states and S stability. This open conformation, involving the receptor-binding domain (RBD) in the up conformation, is impaired when its N-glycosylation at position 343 is knocked down, indicating that RBD dynamics are influenced by glycan-facilitated neighboring N-terminal domain (NTD)-RBD crosstalk. In Chapter 3, we discover that hACE2 binding leads to more prominent dynamic behaviors reflecting hACE2-induced S activation. Notable differences in transduction of allosteric changes are observed, extending from the RBD to regions proximal to proteolytic cleavage sites, suggesting that the highly dynamic fusion peptide region in SARS-CoV-2 S can confer an advantage in fusion. In Chapter 4, we investigate both S conformational dynamics and local structural ordering with a focus on mosaic spike heterotrimers mimicking possible antigenic assemblies from bivalent mRNA vaccination. Both trimer stability and antigenicity are well-conserved in the mosaic trimer formation we study. The mosaic trimer co-expressed form Omicron and Hu-1, resembling the S sequences used in mRNA vaccines, also shows prominent dynamic changes in the fusion peptide proximal region.

These results provide mechanistic insights into receptor-induced S activation. In such a highly dynamic Class I fusion machine, critical variations in amino acid sequences or post-translational modifications can significantly trigger allosteric effects through dynamic motions and interactions between domains, further impacting their transmission phenotypes and viral fitness.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

ACKNOWLEDGEMENTS

This five-year Ph.D. journey has been the most valuable and treasured time of my life. For nearly two years, this thesis project was conducted under the challenging conditions of the COVID-19 pandemic, perhaps the hardest time our generation has ever experienced. Beyond intellectual inputs, the most significant contributions to this journey have been the unwavering support, encouragement, and love from many people around me.

First and foremost, I would like to express my gratitude to my advisor, Dr. Kelly Lee, who has made this project possible. Beyond his guidance and expertise, the most valuable things I have learned from him are the importance of critical thinking and the positive attitude. I am also grateful to his efforts in recommending me for conferences and organizations, as well as introducing me to numerous collaborative projects that have greatly enriched my experience.

I am also thankful to my thesis committee members: Dr. Abhinav Nath, Dr. Miklos Guttman, Dr. Gaurav Bhardwaj and Dr. Neil King, for their constructive feedback and support. Specifically, I have worked in many collaborative projects closely with Dr. Neil King and Dr. Miklos Guttman and gained valuable experiences from them. Dr. Abhinav Nath and Dr. Gaurav Bhardwaj also helped and supported me in the course teaching and departmental events.

I am deeply appreciative of my family, especially my father, Feng Chen; my mother, Hui Wang; and my girlfriend, Vivian Huang. This Ph.D. training is a long journey and the COVID-19 made it more difficult when travel restrictions kept us apart. Besides the financial support from my parents, their mental and emotional support throughout these four years when we could not meet in person, is invaluable. Vivian has been providing constant love and companionship to me.

She is the kindest person I have ever met. She has shared in my successes and witnessed many of my milestone moments, making this journey more meaningful.

I would like to extend my thanks to the University of Washington and my lab. First, I want to particularly thank Edgar Hodge, not only for his contributions in this thesis project, but also for his warm-hearted and patient help outside the lab during the time when we were roommates. His friendship has been a true source of comfort. I also want to thank Richard Zhu, Adam Nguyen, Klaus Lovendahl, Sally Kephart, Jake Croft and Vada Becker, who have also greatly contributed to this project. Special thanks go to Mint Laohajaratsang, the best manager in our lab. She has expertly organized lab events and maintained a great environment, being a model of high emotional intelligence from whom I should learn. Outside of work, playing with her dog, Gunter, has always brought joy and relaxation. I would also like to acknowledge other lab members: Nancy Hom, Alex Mileant, Mark Benhaim, James Williams, Vidya Prasad, Mason Saunders, Meghan McGrath, Kiran Awatramani, Nastassia Parker and Elizabeth Martinez. Each of you has made our research lab a supportive and lovely place to work.

I am also appreciative of the people from my departments, programs, and other research labs who have supported me. Special thanks to the Munro lab at the University of Massachusetts, the Guttman lab, and the King lab for their collaborations. I am grateful to Dale Whittington and Scott Edgar from the Mass Spectrometry Center, as well as Priska von Haller from the University of Washington's Proteomics Resource, for providing excellent research facilities and expert instruction. The people in the Department of Medicinal Chemistry and the Biological Physics, Structure and Design program have been incredibly helpful and kind to me, making my life at the University of Washington over the past five years both enriching and enjoyable.

Last but not least, I would like to acknowledge my friends, both in China and the US. Jessie Yu and I have known each other since we were very young, and we have maintained the longest friendship till now. Xiaoyi Liu, my high school roommate, has shared many important moments with me, and provided useful experience for living in the US. Yue Ma has always motivated me to pursue science and truth, both in life and research. Chang Su kindly crafted and sent me my favorite and luckiest gift, which stayed with me during my most challenging times. I also feel fortunate to have befriended Zhaoqian Wang, Wenxue Zhang, Mingxuan Shi, Lei Gao and Wendy Sun, who have been truly supportive of my life in Seattle.

My sincere appreciation goes to everyone who has appeared in my life, motivating and supporting me to reach my current progress and milestones. Your presence has been invaluable on this journey.

DEDICATION

In memory of my dear grandfather, who took care of me and taught me math when I was young, and encouraged and supported my academic career when I studied abroad, but could not witness this milestone after a two-year battle with cancer.

Wish for no more suffering from diseases and world peace.

Chapter 1. Introduction

1.1 Overview

At the time this thesis project was shaped, both the human world and the biomedical field were experiencing significant transformations. The sudden COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has impacted millions of lives. In response to this unprecedented pandemic, breakthroughs in vaccine development have significantly protected people from this deadly virus. At the same time, progress in scientific research has also led to more comprehensive understanding of the virus from various aspects.

In the field of structural biology, an increasing number of SARS-CoV-2 spike (S) glycoprotein structures have been illustrated by cryo-electron microscopy (Cryo-EM) studies, providing valuable insights into its structural organization and protein interactions. However, static images of protein have inherent limitations in capturing protein dynamics, which is a key feature of such Class I fusion protein.

Studying protein dynamics is crucial for understanding the mechanisms underlying protein function because proteins are not static entities. They undergo various dynamic motions in a wide range, which are essential for their biological roles. This is particularly highlighted in Class I fusion proteins, which are characterized by significantly large-scale conformational changes from pre-fusion to post-fusion states during viral fusion events (Figure 1). Understanding the dynamic behavior of the SARS-CoV-2 S protein not only aids in interpreting its biological function, but also extends to other Class I fusion systems, providing a comprehensive view of how differences in protein dynamics can affect viral infectious phenotypes.

Therefore, the scope of this thesis project lies in utilizing hydrogen/deuterium-exchange mass spectrometry (HDX-MS) to understand the structural dynamic impacts on the emerging S fusion protein and to interpret the biological relevance. Through this project, we aim to compare the scientific research findings summarized later in this chapter and contribute to bridging the gaps between structures and biological functions from the dynamic perspective. Additionally, we seek to provide insights into the structure-based vaccine and antiviral development which can further protect individuals from suffering diseases.

In the following chapters, I will introduce the background of the SARS-CoV-2 S glycoprotein, discuss the achievements and gaps in understanding S protein dynamics, and review the applications of structural mass spectrometry (MS) in studying protein dynamics in Chapter 1. Subsequent chapters will address, investigate, and discuss the biological questions related to the effects of sequence variations and glycan compositions on S conformational preferences (Chapter 2), S activation models based on receptor-binding-induced S dynamic changes (Chapter 3), and the structural dynamic impacts of co-assembled mosaic S heterotrimers (Chapter 4).

1.2 *Betacoronavirus* and COVID-19

Betacoronavirus, a genus in the *Coronaviridae* family, contains a group of viruses that can infect mammals and birds. *Betacoronavirus* has drawn significant public attention due to its association with three major outbreaks and pandemics in the past two decades: the severe acute respiratory syndrome coronavirus (SARS-CoV) outbreak in 2002; the Middle East respiratory syndrome coronavirus (MERS-CoV) outbreak in 2012; and the COVID-19 pandemic (SARS-CoV-2) that began in late 2019.[1]

Viruses in the *Betacoronavirus* genus are enveloped, positive-sense single-stranded RNA viruses. They possess one of the largest genomes among RNA viruses, approximately 26 to 32 kilobases in length.[2, 3] The genome encodes viral structural proteins including the nucleocapsid (N), membrane (M), envelop (E) proteins and S glycoprotein, as well as various non-structural proteins (NSPs) such as the replicase protein and RNA-dependent RNA polymerase (RdRp), essential for viral stability and proliferation.[4]

As the name 'Corona' reveals, the crown-like morphology of the virus is characterized by the S glycoproteins on the virus surface, which functionally facilitate host cell entry by binding to specific receptors on the host cell membrane. For instance, MERS-CoV S recognizes the dipeptidyl peptidase-4 (DPP4) receptor, while both SARS-CoV and SARS-CoV-2 S target the angiotensin-converting enzyme 2 (ACE2) receptor.[5-8]

SARS-CoV and SARS-CoV-2 belong to the *Sarbecovirus* subgenus and share 79% genome sequence similarity.[2] While the SARS-CoV outbreak in 2002 was contained two years after the spillover event, the SARS-CoV-2 virus, responsible for COVID-19, has not yet been contained.[9] Genomic studies have shown high genome sequence similarity with viruses found in bats and pangolins, suggesting possible origins and intermediate hosts of the virus before spilling over into human society.[3, 10] The monopartite genome (genome consists of a single molecule of nucleic acid) and proofreading function of exoribonuclease (ExoN) during SARS-CoV-2 genome replication, have limited its mutation space and rate, compared to many other infectious pathogens.[11, 12] However, once SARS-CoV-2 became mature and circulated in the human population, novel mutations and variants of concern (VOCs) emerged, enhanced the fusogenicity and infectivity of the virus, and resulted in significant morbidity and mortality worldwide.[13-15]

In response to the pandemic, both pharmaceutical and non-pharmaceutical interventions were implemented. Non-pharmaceutical measures such as mask use, social distancing, indoor restrictions, and travel limitations effectively reduced virus spread before vaccines were developed and widely administrated.[16, 17] Many structure-based vaccine designs and technologies had their debut during the COVID-19 pandemic. Among them, mRNA vaccines, Pfizer-BioNTech (BNT162b2) and Moderna (mRNA-1273), have demonstrated high safety and efficacy in protecting populations from virus infections and severe diseases.[18, 19]

1.3 Spike Structure and Function

The S glycoprotein, encoded by the viral genome open reading frame 2 (ORF2), is a protein homotrimer comprised of three ~1300-amino-acid-long protomers.[4] S is displayed on the viral membrane surface and has crucial function for viral entry into host cells, similarly to other trimeric Class I fusion proteins. Structural studies using cryo-electron tomography (cryo-ET) and cryo-electron microscopy (cryo-EM) have elucidated structures of both full-length S on the viral surface and the soluble S ectodomain, which differs from a C-terminal transmembrane domain (TM) that anchors the S trimer into the viral membrane.[20-22]

In the ectodomain (Figure 1), the S protein consists of two structural subunits, the receptor-binding subunit (S1) and the fusion subunit (S2), connected by a furin proteolytic cleavage site (S1/S2). The S1 subunit includes the receptor-binding domain (RBD), locating at the apex of the S trimer, which is primarily responsible for recognizing the host cell ACE2 receptor. Its conformational positions and essential function make RBD the primary target for neutralizing antibodies (nAbs) generated by the immune system.[23-26] The N-terminal domain (NTD) in S1 packs the space between neighboring RBDs and is also targeted by some nAbs, as identified by

structural and immunological studies.[27] S2 subunits constitute the base and core of the spike trimer in its pre-fusion state. A key region, the fusion peptide (FP), is located downstream of the S1/S2 site at a second proteolytic cleavage site (S2') and targeted primarily by TMPRSS2 and cathepsin. Domains including heptad repeat 1 (HR1) and central helix (CH) assemble the central helical bundle in the pre-fusion spike core and participate in post-fusion helix elongation, driving the membrane fusion event (Figure 1, Figure 2).

S proteins are synthesized in the rough endoplasmic reticulum (ER), and cleaved into S1 and S2 subunits by furin in the ER-Golgi intermediate complex (ERGIC), where they are incorporated into nascent virions that bud from the ERGIC membranes.[28] Successful binding between ACE2 and RBD in the open conformation activates S, allowing subsequent cleavage at S2' by a second protease, either TMPRSS2 at the cell surface or cathepsin if the virion is endocytosed.[29, 30] Protease processing of the S2 subunit releases its N-terminal fusion peptide, which facilitates membrane fusion (Figure 2).

Within the *Sarbecovirus* subgenus where SARS-CoV-2 belongs, spike proteins generally share similar sequences and structural organizations. However, their biologically relevant features, such as conformational preferences in the pre-fusion state, and binding affinity to the human ACE2 (hACE2) receptor, vary significantly. For example, a *Betacoronavirus* found in bat, RaTG13, sharing 96% whole-genome level identity with SARS-CoV-2, is widely believed to be associated with the spillover event into the human population.[3] Despite the 90% sequence identity in S glycoproteins and RBD domains, RaTG13 S trimer exhibits over 400-fold weaker binding to hACE2 than SARS-CoV-2 S trimer.[31] The other virus outside the SARS-CoV-2 clade in *Sarbecovirus* subgenus is SARS-CoV, the virus caused 2002-2003 outbreak. Although SARS-CoV and SARS-CoV-2 viruses share less than 80% similarity in their S sequences and even lower

similarity in the RBD sequences, both spike proteins exhibit strong binding affinity to the hACE2 receptor, at lower nanomolar levels.[22, 32] However, the sequence difference in the S1/S2 furin cleavage site results in significantly enhanced activity in the SARS-CoV-2 S, facilitated by a polybasic sequence motif (RRAR) compared to a single arginine in the SARS-CoV S.[33, 34]

RNA viruses typically exhibit high genetic variation, which aids in their evolution and adaptation.[35] Although they have proofreading mechanisms, a number of VOCs have emerged over the course of the pandemic. Just two months after original Wuhan-Hu-1 SARS-CoV-2 S reference sequence was identified, a variant with the point mutation D614G on spike appeared.[13] This variant was soon reported as highly transmissible and became dominant in infections since June 2020.[36] Cryo-EM structures revealed that the D614G mutation bolstered contacts between S1 and S2 subunits. Additionally, studies found that the proportion of the open conformation in S G614 increased, with two major populations exhibiting one or two RBDs in up, receptor-binding-capable conformations.[37, 38]

The majority of neutralizing antibodies elicited after natural infection or vaccine immunization target RBD and NTD of the spike protein.[24, 25] Subsequent mutations in the RBD, where neutralizing antibodies compete with host ACE2 receptor binding, have further enhanced viral fitness. Notable hotspot mutations in the RBD, particularly within the receptor-binding motif (RBM), such as N501Y, E484K/A, and K417N/T, have shown an enhanced ability in antibody escape. As the virus has matured with other mutations on RBD and NTD compensating for hACE2 binding and conformational stability, the virus has evolved to highly infectious VOCs, such as Alpha (B.1.1.7), Delta (B.1.617.2), and Omicron (B.1.1.529), leading to peaks in infection waves from 2020 to 2023.[14, 15, 39]

1.4 Protein Dynamics and Techniques

Fusion proteins are highly dynamic machines that undergo conformational changes essential for their biological functions. As a typical Class I fusion protein, SARS-CoV-2 S undergoes conformational changes from pre-fusion state to post-fusion state with the elongation of the central helical bundle during membrane fusion.[20] The pre-fusion S also switches between closed and open states, with one or more RBDs shifting to receptor-binding-competent conformations (Figure 1).[21, 22] These metastable structures, trapped in local energy minima, can be captured and resolved using advanced cryo-ET and cryo-EM techniques.

However, protein dynamics occur constantly at multiple levels. Both local fluctuations, involving rapid, small-scale movements of atoms in the residues, and large segment movements, such as loops and entire domains shifting between conformational states, can happen.[40] Single particle studies from cryo-EM exhibit superb ability to image structures in the stable endpoint conformations, but they have limited resolution in the intrinsically disordered regions (IDRs) and transient conformations. Therefore, although cryo-EM results have provided valuable insights into S structural organization and conformational samplings, comprehending mechanisms of protein biological functions without information from dynamic studies, may be speculative.

Several experimental and computational techniques are employed to study the dynamics of the spike protein. Single-molecule Förster resonance energy transfer (smFRET) technique monitors real-time energy transfer from a donor fluorophore to an acceptor fluorophore based on their spatial orientation and distance.[41] One smFRET study on the S protein incorporated one fluorophore-labeled tag at the RBD and the other one on the S1 subunit hinge region, tracking the distance changes when the RBD shifts conformations.[42] The smFRET study showed an RBD state with the intermediate distance leading to moderate FRET between the RBD being positioned up

(long distance, low FRET) and down (short distance, high FRET). This RBD state was captured when the neighboring RBD was hACE2-bound, indicating a possible coordination process in receptor recognition.

All-atom molecular dynamics (MD) simulations model the dynamic movements of protein molecules at the atomic level computationally.[43] Microsecond-long MD can predict *in silico* protein dynamics over a wide range. Therefore, dynamic effects from peptide side chains and modifications could be easily captured by MD simulations. For example, all-atom MD studies on S dynamics suggested N-linked glycosylation at N165, N234 and N343 influenced the RBD movement process.[44, 45] This finding highlighted the impact of glycosylation on S conformational states, which static structural studies are likely to overlook due to the poor resolutions on refinements of glycan structures and conformations.

Protein dynamics are integral to many biological processes, a fact that is particularly emphasized in Class I fusion proteins like coronavirus S. While techniques like smFRET and MD simulations are informative, they have inherent limitations in capturing dynamics across the entire protein simultaneously due to the need for fluorophore labeling and required computing power, respectively. To overcome these challenges, hydrogen/deuterium-exchange mass spectrometry (HDX-MS) becomes a common choice for projects focusing on understanding comprehensive protein dynamics and protein functions. The biophysics of HDX-MS and its applications in spike protein dynamics will be discussed in the following sections.

1.5 The Biophysics of Hydrogen/Deuterium-Exchange Mass Spectrometry

Proteins are highly dynamic molecules that exhibit a range of dynamic behaviors, from rapid, small-scale local fluctuations to slow, large-scale domain motions, which can lead to

significant conformational changes.[40] The relatively labile nature of hydrogen on the protein polypeptide backbone amide enables its exchange with water hydrogen, which is further facilitated by the protein dynamics.[46, 47]

HDX-MS technique is developed to sense these innate dynamic motions across the peptide backbone for picturing structural and conformational features of proteins under native solution conditions by probing the accessibility of backbone amide hydrogen to exchange with solvent deuterium.[48-50] The intrinsic rate of the amide hydrogen exchange reaction is dependent on temperature, pH, and specific amino acid type, necessitating rigorous control of these factors during experimental design and performance.[51]

Under physiological and well-controlled conditions, the deuterium exchange reaction rate is influenced by the local structural order and solvent accessibility of the amide hydrogens. Lower exchange rates are observed in peptides stabilized by hydrogen bonding and secondary structures, such as α-helix and β-sheet. Peptides in more flexible and unstructured regions, where the amide groups are more exposed, lead to faster deuterium exchanges.[52]

HDX-MS is a particularly powerful approach for monitoring protein dynamics and conformational changes across various levels and aspects. In continuous labeling HDX experiments, deuterium exchanges with solvent-accessible labile amide hydrogen over different time scales, ranging orders of magnitude from milliseconds to hours.[53] This allows labeling under equilibrium conditions for amide hydrogen in various environments, differentiating a range of dynamic motions. HDX reactions can be both acid- and base-catalyzed and most proteins' physiological pHs are under the base-catalyzed range. A typical HDX-MS study involves HDX labeling at 3-4 orders of magnitude durations under physiological pH or the pH where protein is stably and natively behaved (Figure 3a).[54, 55] The HDX reaction is commonly quenched at pH 2.5

and near 0 °C to halt the reaction with the lowest rate a typical experimental setup can reach. The sample is inline digested with proteases capable of functioning at such pH and temperature conditions, such as pepsin, and analyzed by MS. Each deuterium exchanged results in a +1 Dalton in peptide mass, producing an overall mass distribution in the peptide mass spectra.[54, 56]

HDX mass spectra can monitor dynamic samplings of subpopulation states if they result in amide hydrogens exhibiting different labeling rates (Figure 3b). In reality, most protein structural transitions between subpopulation states occur on a faster time scale than labeling rate, termed EX2 kinetics in HDX-MS. Rapid switching between exposed and protected environments allows only the exposed amide hydrogens to be captured, leading to gradual deuterium exchange. Macroscopically, integration of these events results in the mass spectra envelopes shifting to higher mass over time in a unimodal manner.[50, 57]

HDX-MS is also capable of characterizing both conformational states during large-scale conformational switching. If structural transitions are slower than the deuterium labeling rate, referred to as EX1 kinetics, and peptides in these two conformational states exhibit distinct deuterium exchange rates, the mass spectra will reflect the summation of these two deuterium uptake profiles, displaying bimodal mass envelopes.[50, 58-60] Since conformational states are often functionally distinct, such as activated or inactivated, and receptor-binding-competent or incompetent, the kinetics of state switching can reveal important information about the protein's behavior and functional activity. The bimodal spectra can be a strong indicator of relatively slow conformational sampling between states, or it can report on structural heterogeneity in the sample.

Accurate analysis of HDX-MS data is crucial for meaningful interpretations of protein structural dynamics. Considering whether an amide hydrogen undergoes exchange or not is a binomial event, HDX mass spectra envelopes of the peptide in a particular state with similar amide

hydrogen exchange rates follow a binomial distribution. The expectation (mean) of this binomial distribution is calculated using centroid fitting, which reflects the averaged deuterium uptake level of the peptide. While centroid fitting is commonly used for EX2 kinetics when the spectra is unimodal, it may not reflect sufficient structural information from EX1 kinetics.[61] For EX1 kinetics where bimodal mass spectra appeared, resolving two models and separately calculating the uptake levels of two populations are necessary to characterize distinct conformational states.[59] The ratio and effect size between these populations provide detailed structural information on the conformational samplings.

In summary, HDX-MS is a powerful tool for investigating protein dynamic motions and conformational changes, providing insights into protein behavior and function. Through rigorous design and analysis of HDX-MS experiments under nearly native conditions, we can dissect the structural dynamics information, enhancing our understanding of protein interaction and biological roles in addition to other structural methods.

1.6 Structural Mass Spectrometry on Protein Dynamics

Compared to other dynamic techniques, HDX-MS is conducted with a label-free system under conditions that closely mimic the native environments of the proteins. Its versatility allows HDX-MS to study the dynamics of various proteins, including engineered and designed proteins. Due to the close association between protein dynamics and biological function, HDX-MS is commonly used to study surface functional proteins, such as receptors, transporters and fusion proteins. HDX-MS's powerful advantages highlight its ability to capture transient conformational states and sense rapid local structural order fluctuations.

HDX-MS has demonstrated exceptional performance in studies of Class I fusion proteins, monitoring dynamic processes during pre-fusion metastable state switches and structural reorganization during activation and membrane fusion. For example, influenza hemagglutinin (HA), a typical Class I fusion protein, consists of the receptor-binding subunit (HA1) and fusion subunit (HA2). It was long time assumed that HA activation followed a spring-loaded mechanism, where at activation pH, HA1 shed first, triggering the release of the fusion peptide in HA2, from studies limited in capturing transient transitions in the activation process. Recent HDX-MS studies revealed that during the activation of intact X-31 influenza virion (H3 subtype) HA, both peptides at the HA1/HA2 interface and the fusion peptide (FP) showed dynamic changes different from their pre- and post-fusion states. This indicated a transient intermediate state with concurrent HA interface reorganization and FP release.[62] However, this intermediate state was not observed when using the soluble bromelain-released HA (BHA) ectodomain. The activation modes were also shown to be strain-dependent, with H1 subtype HA favoring priming with HA1 uncaging but H3 subtype HA favoring FP exposure.[63] Strain-specific dynamics over local structural orders were observed and resolved by HDX-MS studies on the HIV envelope proteins (Env), detecting rapid dynamic fluctuations at common epitope peptide sequences.[64] These HDX-MS studies suggested the differences across strains in their strategies against broadly neutralizing antibodies.

The label-free feature of HDX-MS enables it to overview dynamics across all regions of the intact protein simultaneously and is often applied to detect allosteric effects. Recent HDX-MS studies on SARS-CoV-2 S variants observed concurrent dynamic changes in the NTD and the distant trimer C-terminal stalk region in the Wuhan-Hu-1 G614, Alpha, and Delta variants.[65] This observation was partially supported by another study on spikes of VOCs, which found increased dynamics in the NTD and S2 subunit across all emerging variants.[66] However, the dynamics of the

RBD did not increase along the evolutionary timeline of variant emergences, indicating that the Omicron variant possessed less open state conformations despite increased dynamics in the NTD and S2 subunits.

To understand the interaction between S and hACE2, besides the observations of RBD conformational shifts and direct receptor interactions, which could also be characterized by other structural techniques, HDX-MS provided insights into allosteric effects, enhancing our understanding of the biological mechanisms. Separated studies observed allosteric effects with increased dynamics in the fusion peptide proximal region (FPPR)[67] and the S1/S2 site[68], respectively, under hACE2 saturated and unsaturated conditions.

The sensitivity of HDX-MS in characterizing rapid local dynamic fluctuations and peptide structural orders is also widely used in examining antigenicity. Structure-based vaccine candidates, including engineered stabilized protein forms and designed proteins, require structural orders resembling the native antigens. The incorporation of HDX-MS technique on the respiratory syncytial virus (RSV) stabilized F glycoprotein DS-Cav-1[69] and SARS-CoV-2 RBD[70, 71] antigen-displayed nanoparticles, have provided valuable insights into antigenicity alongside binding experiments, aiding in vaccine development.

1.7 Summary

In this chapter, we briefly introduce the background on the biological system and the protein of interest, SARS-CoV-2 S glycoprotein, which is the focus of this thesis study. Despite the relative novelty of the SARS-CoV-2 virus, advancements in scientific technology have elucidated the structures of the spike protein and facilitated the development of highly effective vaccines within a very short period. However, the spike protein is intrinsically highly dynamic,

and understanding the connections between its structural features and biological impacts requires further studies on protein dynamics. HDX-MS, as a powerful technique for investigating protein dynamics, has been extensively utilized in research on various viral surface proteins, which are known for their highly dynamic nature, including the SARS-CoV-2 S.

In this thesis, HDX-MS is used to comprehensively investigate the dynamics of the SARS-CoV-2 S, testing the central hypothesis that the phenotypical features of the SARS-CoV-2 virus are significantly influenced by the dynamic behaviors of S variants during receptor binding and activation processes. Three specific questions on dynamic features are studied and discussed in the following chapters:

1. How do amino acid sequence variations and glycosylation modifications drive S into preferred conformational states?

2. How does hACE2 receptor binding induce allosteric effects, which can prime and activate S?

3. How do protein dynamics and antigenicity change when protomers with distinct dynamics are co-assembled, potentially resulting from bivalent mRNA antigen displays.

These questions are crucial for understanding the structural dynamics and functional implications of the spike protein, providing insights into both biological knowledge on SARS-CoV-2 S and strategies for vaccine development.

Figures

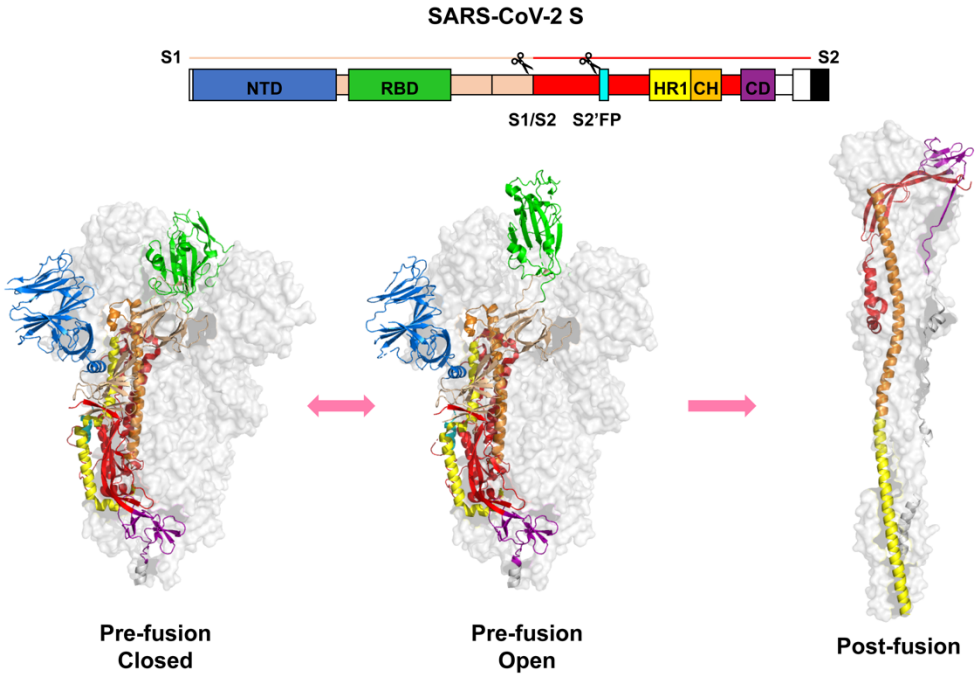

Figure 1. **Structural organization of SARS-CoV-2 S.** Color-coded S domains in sequence displays (Upper) and cartoon structure displays (Lower) from PDB structures describe spatial location of S domains during conformational changes. PDB: pre-fusion closed state: 6VXX; pre-fusion open state: 6VYB; post-fusion state: 6XRA.

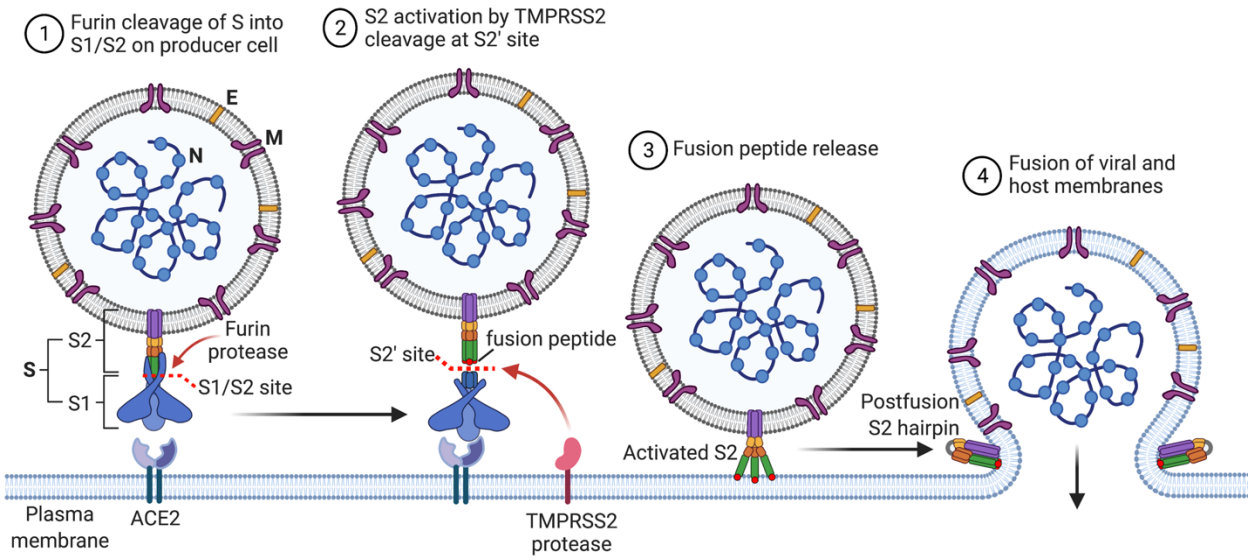

Figure 2. **Cartoon diagram on S-facilitated SARS-CoV-2 host cell entry pathway.** Viral surface displaying S in open state is recognized by host ACE2 receptor. Successful binding between S and ACE2 receptor recruits TMPRSS2 to cleave at S2' site. Released fusion peptides insert into host membrane, resulting in S conformational changes into post-fusion form, which drives the membrane fusion. Figure is created and edited from the BioRender templates.

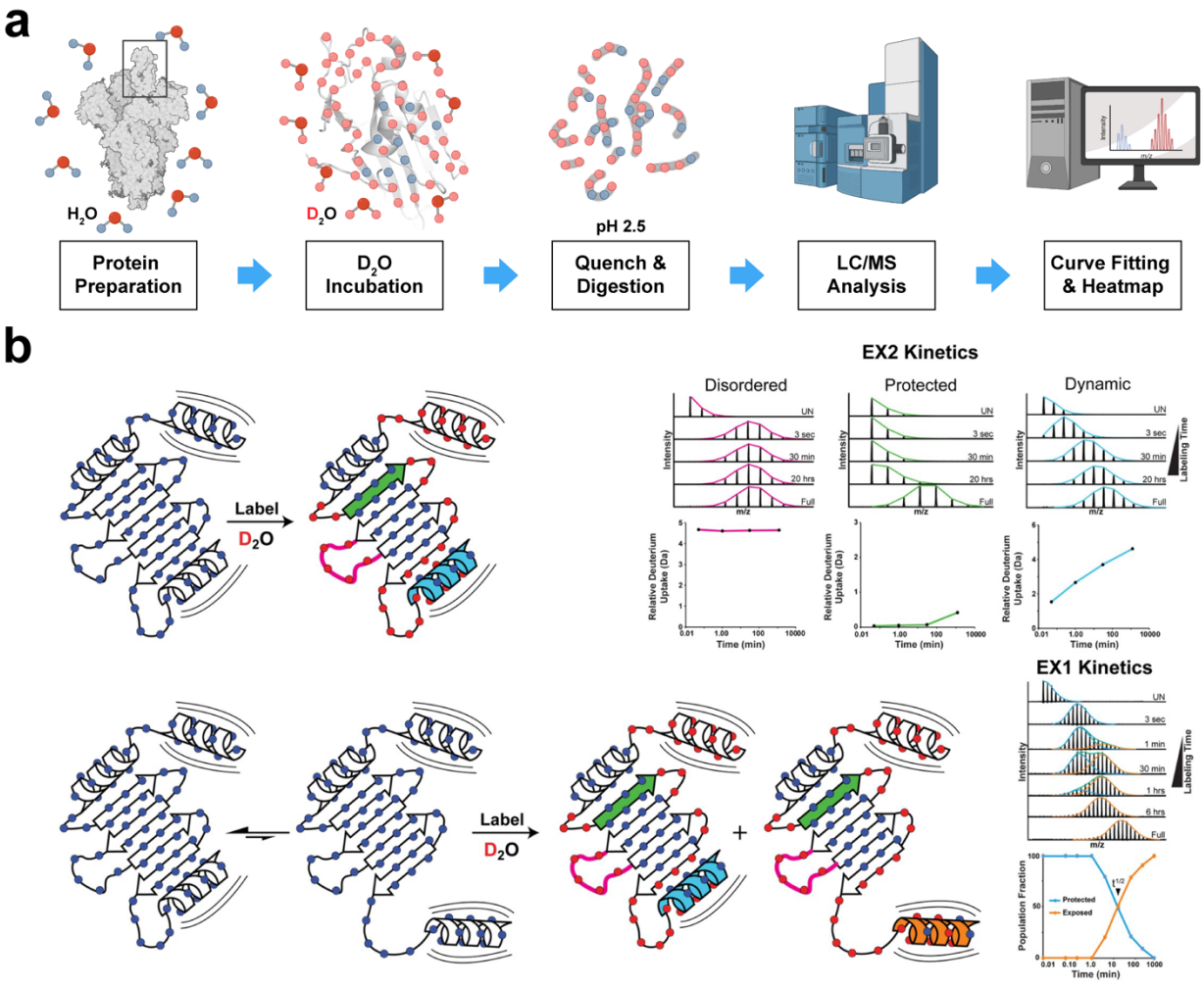

Figure 3. **Schematic illustration and dynamics interpretation by HDX-MS. (a)** General experimental process of continuous labeling HDX-MS on SARS-CoV-2 S. **(b)** Dynamic motions interpreted from the apparent deuterium uptake plots. In EX2 kinetics, only one state is HDX captured, and uptake plots reveal the corresponding environments of different peptides. In EX1 kinetics, peptide in both states can be detected, resulting in bimodal mass envelopes from deuterium uptake in two states. Figure is adapted and edited from Reference 50, with permission.

References

(1) Gorbalenya, A. E.; Baker, S. C.; Baric, R. S.; de Groot, R. J.; Drosten, C.; Gulyaeva, A. A.; Haagmans, B. L.; Lauber, C.; Leontovich, A. M.; Neuman, B. W.; et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology* **2020**, *5* (4), 536-544. DOI: 10.1038/s41564-020-0695-z.

(2) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* **2020**, *395* (10224), 565-574. DOI: 10.1016/S0140-6736(20)30251-8 (acccessed 2024/06/26).

(3) 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. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579* (7798), 270-273. DOI: 10.1038/s41586-020-2012-7.

(4) Kim, D.; Lee, J.-Y.; Yang, J.-S.; Kim, J. W.; Kim, V. N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181* (4), 914-921.e910. DOI: 10.1016/j.cell.2020.04.011 (acccessed 2024/06/26).

(5) Wang, N.; Shi, X.; Jiang, L.; Zhang, S.; Wang, D.; Tong, P.; Guo, D.; Fu, L.; Cui, Y.; Liu, X.; et al. Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Research* **2013**, *23* (8), 986-993. DOI: 10.1038/cr.2013.92.

(6) 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. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **2003**, *426* (6965), 450-454. DOI: 10.1038/nature02145.

(7) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367* (6485), 1444-1448. DOI: 10.1126/science.abb2762.

(8) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581* (7807), 215-220. DOI: 10.1038/s41586-020-2180-5.

(9) Zhong, N. S.; Zheng, B. J.; Li, Y. M.; Poon; Xie, Z. H.; Chan, K. H.; Li, P. H.; Tan, S. Y.; Chang, Q.; Xie, J. P.; et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* **2003**, *362* (9393), 1353-1358. DOI: 10.1016/s0140-6736(03)14630-2.

(10) Zhang, T.; Wu, Q.; Zhang, Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. *Current Biology* **2020**, *30* (7), 1346-1351.e1342. DOI: https://doi.org/10.1016/j.cub.2020.03.022.

(11) Gussow, A. B.; Auslander, N.; Faure, G.; Wolf, Y. I.; Zhang, F.; Koonin, E. V. Genomic determinants of pathogenicity in SARS-CoV-2 and other human coronaviruses. *Proceedings of the National Academy of Sciences* **2020**, *117* (26), 15193-15199. DOI: doi:10.1073/pnas.2008176117.

(12) Moeller, N. H.; Shi, K.; Demir, Ö.; Belica, C.; Banerjee, S.; Yin, L.; Durfee, C.; Amaro, R. E.; Aihara, H. Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN. *Proceedings of the National Academy of Sciences* **2022**, *119* (9), e2106379119. DOI: doi:10.1073/pnas.2106379119.

(13) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; et al. Tracking Changes in SARS-CoV-2 Spike:

Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020**, *182* (4), 812-827.e819. DOI: 10.1016/j.cell.2020.06.043.

(14) Khan, A.; Zia, T.; Suleman, M.; Khan, T.; Ali, S. S.; Abbasi, A. A.; Mohammad, A.; Wei, D. Q. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *J Cell Physiol* **2021**, *236* (10), 7045-7057. DOI: 10.1002/jcp.30367.

(15) Volz, E.; Mishra, S.; Chand, M.; Barrett, J. C.; Johnson, R.; Geidelberg, L.; Hinsley, W. R.; Laydon, D. J.; Dabrera, G.; O'Toole, Á.; et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* **2021**, *593* (7858), 266-269. DOI: 10.1038/s41586-021-03470-x.

(16) Lai, S.; Ruktanonchai, N. W.; Zhou, L.; Prosper, O.; Luo, W.; Floyd, J. R.; Wesolowski, A.; Santillana, M.; Zhang, C.; Du, X.; et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **2020**, *585* (7825), 410-413. DOI: 10.1038/s41586-020-2293-x.

(17) Lison, A.; Banholzer, N.; Sharma, M.; Mindermann, S.; Unwin, H. J. T.; Mishra, S.; Stadler, T.; Bhatt, S.; Ferguson, N. M.; Brauner, J.; et al. Effectiveness assessment of non-pharmaceutical interventions: lessons learned from the COVID-19 pandemic. *The Lancet Public Health* **2023**, *8* (4), e311-e317. DOI: 10.1016/S2468-2667(23)00046-4 (acccessed 2024/06/26).

(18) Polack, F. P.; Thomas, S. J.; Kitchin, N.; Absalon, J.; Gurtman, A.; Lockhart, S.; Perez, J. L.; Marc, G. P.; Moreira, E. D.; Zerbini, C.; et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **2020**, *383* (27), 2603-2615. DOI: doi:10.1056/NEJMoa2034577.

(19) Baden, L. R.; Sahly, H. M. E.; Essink, B.; Kotloff, K.; Frey, S.; Novak, R.; Diemert, D.; Spector, S. A.; Rouphael, N.; Creech, C. B.; et al. Efficacy and Safety of the mRNA-1273 SARS-

CoV-2 Vaccine. *New England Journal of Medicine* **2021**, *384* (5), 403-416. DOI: doi:10.1056/NEJMoa2035389.

(20) Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; et al. Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020**, *183* (3), 730-738.e713. DOI: 10.1016/j.cell.2020.09.018.

(21) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367* (6483), 1260-1263. DOI: 10.1126/science.abb2507.

(22) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e286. DOI: 10.1016/j.cell.2020.02.058.

(23) Liu, L.; Wang, P.; Nair, M. S.; Yu, J.; Rapp, M.; Wang, Q.; Luo, Y.; Chan, J. F.; Sahi, V.; Figueroa, A.; et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **2020**, *584* (7821), 450-456. DOI: 10.1038/s41586-020-2571-7.

(24) 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.; et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **2020**, *588* (7839), 682-687. DOI: 10.1038/s41586-020-2852-1.

(25) 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. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **2020**, *584* (7821), 437-442. DOI: 10.1038/s41586-020-2456-9.

(26) Guenthoer, J.; Lilly, M.; Starr, T. N.; Dadonaite, B.; Lovendahl, K. N.; Croft, J. T.; Stoddard, C. I.; Chohan, V.; Ding, S.; Ruiz, F.; et al. Identification of broad, potent antibodies to functionally constrained regions of SARS-CoV-2 spike following a breakthrough infection. *Proceedings of the National Academy of Sciences* **2023**, *120* (23), e2220948120. DOI: doi:10.1073/pnas.2220948120.

(27) Chi, X.; Yan, R.; Zhang, J.; Zhang, G.; Zhang, Y.; Hao, M.; Zhang, Z.; Fan, P.; Dong, Y.; Yang, Y.; et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **2020**, *369* (6504), 650-655. DOI: 10.1126/science.abc6952.

(28) Broer, R.; Boson, B.; Spaan, W.; Cosset, F. L.; Corver, J. Important role for the transmembrane domain of severe acute respiratory syndrome coronavirus spike protein during entry. *J Virol* **2006**, *80* (3), 1302-1310. DOI: 10.1128/jvi.80.3.1302-1310.2006.

(29) Glowacka, I.; Bertram, S.; Müller, M. A.; Allen, P.; Soilleux, E.; Pfefferle, S.; Steffen, I.; Tsegaye, T. S.; He, Y.; Gnirss, K.; et al. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J Virol* **2011**, *85* (9), 4122-4134. DOI: 10.1128/jvi.02232-10.

(30) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181* (2), 271-280.e278. DOI: 10.1016/j.cell.2020.02.052.

(31) Wrobel, A. G.; Benton, D. J.; Xu, P.; Roustan, C.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* **2020**, *27* (8), 763-767. DOI: 10.1038/s41594-020-0468-7.

(32) Wu, K.; Peng, G.; Wilken, M.; Geraghty, R. J.; Li, F. Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. *J Biol Chem* **2012**, *287* (12), 8904-8911. DOI: 10.1074/jbc.M111.325803.

(33) Sasaki, M.; Uemura, K.; Sato, A.; Toba, S.; Sanaki, T.; Maenaka, K.; Hall, W. W.; Orba, Y.; Sawa, H. SARS-CoV-2 variants with mutations at the S1/S2 cleavage site are generated in vitro during propagation in TMPRSS2-deficient cells. *PLoS Pathog* **2021**, *17* (1), e1009233. DOI: 10.1371/journal.ppat.1009233.

(34) Johnson, B. A.; Xie, X.; Bailey, A. L.; Kalveram, B.; Lokugamage, K. G.; Muruato, A.; Zou, J.; Zhang, X.; Juelich, T.; Smith, J. K.; et al. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* **2021**, *591* (7849), 293-299. DOI: 10.1038/s41586-021-03237-4.

(35) Domingo, E.; Holland, J. J. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* **1997**, *51*, 151-178. DOI: 10.1146/annurev.micro.51.1.151.

(36) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **2020**, *183* (3), 739-751.e738. DOI: 10.1016/j.cell.2020.09.032.

(37) Gobeil, S. M.; Janowska, K.; McDowell, S.; Mansouri, K.; Parks, R.; Manne, K.; Stalls, V.; Kopp, M. F.; Henderson, R.; Edwards, R. J.; et al. D614G Mutation Alters SARS-CoV-2 Spike Conformation and Enhances Protease Cleavage at the S1/S2 Junction. *Cell Rep* **2021**, *34* (2), 108630. DOI: 10.1016/j.celrep.2020.108630.

(38) Benton, D. J.; Wrobel, A. G.; Roustan, C.; Borg, A.; Xu, P.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. The effect of the D614G substitution on the structure of the spike

glycoprotein of SARS-CoV-2. *Proc Natl Acad Sci U S A* **2021**, *118* (9). DOI: 10.1073/pnas.2022586118.

(39) Harvey, W. T.; Carabelli, A. M.; Jackson, B.; Gupta, R. K.; Thomson, E. C.; Harrison, E. M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S. J.; et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology* **2021**, *19* (7), 409-424. DOI: 10.1038/s41579-021-00573-0.

(40) Bu, Z.; Callaway, D. J. E. Chapter 5 - Proteins MOVE! Protein dynamics and long-range allostery in cell signaling. In *Advances in Protein Chemistry and Structural Biology*, Donev, R. Ed.; Vol. 83; Academic Press, 2011; pp 163-221.

(41) Roy, R.; Hohng, S.; Ha, T. A practical guide to single-molecule FRET. *Nature Methods* **2008**, *5* (6), 507-516. DOI: 10.1038/nmeth.1208.

(42) Lu, M.; Uchil, P. D.; Li, W.; Zheng, D.; Terry, D. S.; Gorman, J.; Shi, W.; Zhang, B.; Zhou, T.; Ding, S.; et al. Real-Time Conformational Dynamics of SARS-CoV-2 Spikes on Virus Particles. *Cell Host Microbe* **2020**, *28* (6), 880-891.e888. DOI: 10.1016/j.chom.2020.11.001.

(43) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99* (6), 1129-1143. DOI: 10.1016/j.neuron.2018.08.011.

(44) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; et al. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent Sci* **2020**, *6* (10), 1722-1734. DOI: 10.1021/acscentsci.0c01056.

(45) Sztain, T.; Ahn, S. H.; Bogetti, A. T.; Casalino, L.; Goldsmith, J. A.; Seitz, E.; McCool, R. S.; Kearns, F. L.; Acosta-Reyes, F.; Maji, S.; et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat Chem* **2021**, *13* (10), 963-968. DOI: 10.1038/s41557-021-00758-3.

(46) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. In *Advances in Protein Chemistry*, Anfinsen, C. B., Anson, M. L., Edsall, J. T., Richards, F. M. Eds.; Vol. 21; Academic Press, 1966; pp 287-386.

(47) Englander, S. W.; Mayne, L.; Bai, Y.; Sosnick, T. R. Hydrogen exchange: the modern legacy of Linderstrøm-Lang. *Protein Sci* **1997**, *6* (5), 1101-1109. DOI: 10.1002/pro.5560060517.

(48) Marcsisin, S. R.; Engen, J. R. Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Anal Bioanal Chem* **2010**, *397* (3), 967-972. DOI: 10.1007/s00216-010-3556-4.

(49) Iacob, R. E.; Engen, J. R. Hydrogen exchange mass spectrometry: are we out of the quicksand? *J Am Soc Mass Spectrom* **2012**, *23* (6), 1003-1010. DOI: 10.1007/s13361-012-0377-z.

(50) Hodge, E. A.; Benhaim, M. A.; Lee, K. K. Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci* **2020**, *29* (4), 843-855. DOI: 10.1002/pro.3790.

(51) Persson, F.; Halle, B. How amide hydrogens exchange in native proteins. *Proceedings of the National Academy of Sciences* **2015**, *112* (33), 10383-10388. DOI: doi:10.1073/pnas.1506079112.

(52) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins* **1993**, *17* (1), 75-86. DOI: 10.1002/prot.340170110.

(53) Deng, Y.; Zhang, Z.; Smith, D. L. Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics. *J Am Soc Mass Spectrom* **1999**, *10* (8), 675-684. DOI: 10.1016/s1044-0305(99)00038-0.

(54) Narang, D.; Lento, C.; D, J. W. HDX-MS: An Analytical Tool to Capture Protein Motion in Action. *Biomedicines* **2020**, *8* (7). DOI: 10.3390/biomedicines8070224.

(55) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; et al. Recommendations for performing, interpreting

and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods* **2019**, *16* (7), 595-602. DOI: 10.1038/s41592-019-0459-y.

(56) Wang, L.; Pan, H.; Smith, D. L. Hydrogen Exchange-Mass Spectrometry: Optimization of Digestion Conditions*. *Molecular & Cellular Proteomics* **2002**, *1* (2), 132-138. DOI: https://doi.org/10.1074/mcp.M100009-MCP200.

(57) Ferraro, D. M.; Lazo, N.; Robertson, A. D. EX1 hydrogen exchange and protein folding. *Biochemistry* **2004**, *43* (3), 587-594. DOI: 10.1021/bi035943y.

(58) Low, J. S.; Jerak, J.; Tortorici, M. A.; McCallum, M.; Pinto, D.; Cassotta, A.; Foglierini, M.; Mele, F.; Abdelnabi, R.; Weynand, B.; et al. ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies. *Science* **2022**, *377* (6607), 735-742. DOI: 10.1126/science.abq2679.

(59) Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *J Am Soc Mass Spectrom* **2013**, *24* (12), 1906-1912. DOI: 10.1007/s13361-013-0727-5.

(60) Sivaraman, T.; Robertson, A. D. Kinetics of conformational fluctuations by EX1 hydrogen exchange in native proteins. *Methods Mol Biol* **2001**, *168*, 193-214. DOI: 10.1385/1-59259-193-0:193.

(61) Zhang, J.; Ramachandran, P.; Kumar, R.; Gross, M. L. H/D exchange centroid monitoring is insufficient to show differences in the behavior of protein states. *J Am Soc Mass Spectrom* **2013**, *24* (3), 450-453. DOI: 10.1007/s13361-012-0555-z.

(62) Benhaim, M. A.; Mangala Prasad, V.; Garcia, N. K.; Guttman, M.; Lee, K. K. Structural monitoring of a transient intermediate in the hemagglutinin fusion machinery on influenza virions. *Sci Adv* **2020**, *6* (18), eaaz8822. DOI: 10.1126/sciadv.aaz8822.

(63) Garcia, N. K.; Kephart, S. M.; Benhaim, M. A.; Matsui, T.; Mileant, A.; Guttman, M.; Lee, K. K. Structural dynamics reveal subtype-specific activation and inhibition of influenza virus hemagglutinin. *J Biol Chem* **2023**, *299* (6), 104765. DOI: 10.1016/j.jbc.2023.104765.

(64) Hodge, E. A.; Naika, G. S.; Kephart, S. M.; Nguyen, A.; Zhu, R.; Benhaim, M. A.; Guo, W.; Moore, J. P.; Hu, S.-L.; Sanders, R. W.; et al. Structural dynamics reveal isolate-specific differences at neutralization epitopes on HIV Env. *iScience* **2022**, *25* (6), 104449. DOI: https://doi.org/10.1016/j.isci.2022.104449.

(65) Braet, S. M.; Buckley, T. S. C.; Venkatakrishnan, V.; Dam, K. A.; Bjorkman, P. J.; Anand, G. S. Timeline of changes in spike conformational dynamics in emergent SARS-CoV-2 variants reveal progressive stabilization of trimer stalk with altered NTD dynamics. *Elife* **2023**, *12*. DOI: 10.7554/eLife.82584.

(66) Calvaresi, V.; Wrobel, A. G.; Toporowska, J.; Hammerschmid, D.; Doores, K. J.; Bradshaw, R. T.; Parsons, R. B.; Benton, D. J.; Roustan, C.; Reading, E.; et al. Structural dynamics in the evolution of SARS-CoV-2 spike glycoprotein. *Nat Commun* **2023**, *14* (1), 1421. DOI: 10.1038/s41467-023-36745-0.

(67) Chen, C.; Zhu, R.; Hodge, E. A.; Díaz-Salinas, M. A.; Nguyen, A.; Munro, J. B.; Lee, K. K. hACE2-Induced Allosteric Activation in SARS-CoV versus SARS-CoV-2 Spike Assemblies Revealed by Structural Dynamics. *ACS Infectious Diseases* **2023**, *9* (6), 1180-1189. DOI: 10.1021/acsinfecdis.3c00010.

(68) Raghuvamsi, P. V.; Tulsian, N. K.; Samsudin, F.; Qian, X.; Purushotorman, K.; Yue, G.; Kozma, M. M.; Hwa, W. Y.; Lescar, J.; Bond, P. J.; et al. SARS-CoV-2 S protein:ACE2 interaction reveals novel allosteric targets. *Elife* **2021**, *10*. DOI: 10.7554/eLife.63646.

(69) Marcandalli, J.; Fiala, B.; Ols, S.; Perotti, M.; de van der Schueren, W.; Snijder, J.; Hodge, E.; Benhaim, M.; Ravichandran, R.; Carter, L.; et al. Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **2019**, *176* (6), 1420-1431.e1417. DOI: 10.1016/j.cell.2019.01.046.

(70) 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.; et al. Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell* **2020**, *183* (5), 1367-1382.e1317. DOI: 10.1016/j.cell.2020.10.043.

(71) Ellis, D.; Brunette, N.; Crawford, K. H. D.; Walls, A. C.; Pham, M. N.; Chen, C.; Herpoldt, K. L.; Fiala, B.; Murphy, M.; Pettie, D.; et al. Stabilization of the SARS-CoV-2 Spike Receptor-Binding Domain Using Deep Mutational Scanning and Structure-Based Design. *Front Immunol* **2021**, *12*, 710263. DOI: 10.3389/fimmu.2021.710263.

Chapter 2. Strain Specific Spike Structural Dynamics and Glycan Composition Impacts on Spike Stability and Conformational States

The content in this chapter is reproduced in part with permission from: Chen, C.; Zhu, R.; Hodge, E. A.; Díaz-Salinas, M. A.; Nguyen, A.; Munro, J. B.; Lee, K. K. hACE2-Induced Allosteric Activation in SARS-CoV versus SARS-CoV-2 Spike Assemblies Revealed by Structural Dynamics. *ACS Infectious Diseases* **2023**, *9* (6), 1180-1189. DOI: 10.1021/acsinfecdis.3c00010.

2.1 Introduction

Since the first documented case of SARS-CoV-2 infection[1], the COVID-19 pandemic, has caused over 770 million documented infections and 7 million deaths as of May 2024 (WHO). In response, dedicated efforts have been made to develop effective antivirals and vaccines, alongside numerous studies investigating various aspects of this novel virus, including its origin, evolution, and structural properties.

Coronavirus SARS-CoV-2 utilizes the spike (S) glycoprotein to recognize target receptor, human angiotensin-converting enzyme 2 (hACE2) for entering host cells.[2, 3] The soluble S pre-fusion ectodomain cryogenic electron microscopy (cryo-EM) structure of SARS-CoV-2 S was determined soon after genome reference sequence of S was published.[4-6] A soluble S ectodomain that maintained a prefusion conformation was engineered by furin cleavage site silencing, di-proline mutating at the apex of the heptad repeat 1 (HR1) domain and the central helix (CH) domain (to prevent transition to post-fusion state) and addition of a C-terminal trimerization foldon domain.[5-7] This consecutive di-proline mutated S trimer is referred to as S-2P. A hyper-stabilized form with four additional proline mutations was later developed and is referred to as S-6P or

'Hexapro S'.[8] Both S-2P and S-6P have comparable hACE2 binding affinities, reported at 11.3 nM and 13.3 nM, respectively.[8] Structures of full-length, detergent solubilized S trimers have also been reported[9, 10], revealing that the S-2P and S-6P stabilized ectodomain forms of S closely mimic the full-length form without engineered changes but they display rotated N-terminal domains (NTD) and are with a reduced propensity to adopt a receptor-binding domain (RBD)-down closed conformation.

The proline-stabilized S ectodomain constructs have facilitated numerous structural studies including investigations of the structural and conformational properties of S such as monitoring RBD up and down conformations and measuring conformational changes induced by receptor or antibody binding. Since the hACE2 receptor recognizes the receptor binding motif (RBM), the amino acid sequence on the RBD surface, only when the RBD is in its up positioning, the ratio of up versus down conformers significantly impacts the likelihood of S receptor binding and activation as takes place to initiate cell entry.[11, 12] Cryo-EM structures of SARS-CoV-2 S-2P showed an approximately equal distribution of all-RBD-down (closed) conformation and one-RBD-up (open) conformation.[6] S-6P yielded majority of open conformations, with roughly 30% two-RBD-up conformation.[8] Cryo-electron tomography (cryo-ET) studies on the intact virions fixed with paraformaldehyde indicated conformational distributions of ~40% each of closed and one-RBD-up conformation and 13% particles showing two-RBD-up conformation.[13] It has been suggested that S states bearing more RBDs in the up conformations may be energetically unfavorable, and that hACE2 binding stabilized the higher energy activated states, priming the spike for subsequent conformational changes associated with proteolytic cleavage of S1/S2 subunits and membrane fusion.[14] As the virus evolved over the course of the pandemic, mutations that altered its stability, conformational profile as well as antigenicity arose. The first new variant

that emerged in June 2020, just months after the pandemic began, featured a single point mutation, D614G, in the spike protein, which conferred adaptive advantages such that D614G S quickly rose to become the dominant strain in circulation.[15, 16] Cryo-EM studies revealed that D614G mutation increased the proportion of S in the open conformations, with two major populations exhibiting one- or two-RBD-up conformations.[16-18] It is believed that this enhanced virus infectivity by increasing receptor engagement.

While static structures reveal the S structural organization and conformational states, understanding S function and mutational effects requires investigating this fusion protein's structural and conformational dynamics as well as transitions between known conformational states. Single-molecule Förster resonance energy transfer (smFRET) study on S trimer from virus-like particle (VLP) with one fluorophore on the RBD and a second one on S1 hinge region, has been used to monitor RBD conformational changes.[14] This smFRET study revealed the existence of an intermediate state between closed and one-RBD-up states reported by a moderate intensity FRET signal, which was interpreted as an RBD being positioned between up (long distance, low FRET) and down (short distance, high FRET). Later smFRET studies also revealed that D614G mutation led to S more frequently sampling in the RBD-up states.[19] Furthermore, all-atom molecular dynamics (MD) simulations suggested the involvement of N343 glycan in modulating the RBD opening process.[20, 21] These findings highlighted the impacts that glycosylation can also have on S conformational states.

In this chapter, I will describe the use of hydrogen/deuterium-exchange mass spectrometry (HDX-MS), a powerful technique for monitoring protein dynamics under native solution conditions, to investigate S conformational dynamics and the effects of the D614G mutation and N343 glycan on S stability, conformational switching, and structural dynamics.

2.2 Materials and Methods

Plasmid construction

For experiments comparing SARS-CoV-2 S D614 and G614: the gene sequences encoding SARS-CoV-2 S ectodomain (residues 14-1211) with di-proline mutations (S-2P: K986P and V987P), SGAG substitution at S1/S2 site (RRAR) and C-terminal T4 fibritin foldon with His-tag were cloned into pcDNA3.1(-) vector. The D614G mutation was generated from Wuhan-Hu-1 SARS-CoV-2 S wild type sequence with a Q5 Site-Directed Mutagenesis Kit (NEB).

For experiments knocking down SARS-CoV-2 S N343 glycosylation: the gene sequences encoding SARS-CoV-2 S ectodomain with hexa-proline mutations (S-6P: F817P, A892P, A899P, A942P, K986P and V987P), GSAS substitution at S1/S2 site (RRAR) and C-terminal T4 fibritin foldon with His-tag were cloned into pCMV plasmid. The wild-type sequence in this experiment refers to S with glycine at residue 614 (G614). Both N343A and N343Q mutations were generated from S G614 sequence with a Q5 Site-Directed Mutagenesis Kit (NEB). Disulfide mutated closed-state S was constructed by introducing intra-molecular disulfide bond from S383C and D985C mutations, which were also generated by the Q5 Site-Directed Mutagenesis Kit (NEB).

Transient transfection

Suspension Expi293F cell cultures were used for S-2P and S-6P protein expression. For each liter of cell culture at confluency (3×10^6 cells/mL), 1 mg plasmid DNA stated above and 3 mL filtered polyethylenimine (PEI) solution at 1 mg/mL were pre-incubated in 25 mL Gibco Opti-MEM (Thermo Fisher Scientific) for 5 minutes before mixing. The mixture was incubated for 15 minutes and then added dropwise into the cell culture. The total transfection volume of each

construct depends on the protein yield and experimental need. The transfected Expi293F cultures were incubated at 37 °C and 8% CO_2 with constantly shaking at 125 RPM.

Protein purification

On the fifth day post-transfection, transfected Expi293F cultures were harvested by centrifugation at 2000 RCF for 20 minutes. Since all the protein constructs used in this chapter had C-terminal His-tag, the His-tag affinity purification protocols used for all the constructs were similar. Supernatant was vacuum-filtered through Nalgene™ 0.45-μm αPES filters (Thermo Fisher Scientific) and supplemented with tris hydrochloride (Tris-HCl, pH 8.0), arginine hydrochloride (ArgCl, pH 6.5), sodium chloride (NaCl) and imidazole to a final concentration of 10 mM Tris-HCl, 50 mM ArgCl, 150 mM NaCl and 5 mM imidazole.[22] For each liter of the supernatant, 5 mL of Ni-NTA resin was washed and equilibrated with wash buffer (50 mM Tris-HCl, 100 mM ArgCl, 150 mM NaCl and 5 mM imidazole, pH 8.0), followed by adding into the supernatant for a 2-hour batch binding at room temperature. The mixture was then flowed through a gravity column (Bio-Rad), followed by three times 15 mL (3 column volumes, CV) wash with wash buffer. The bound protein was eluted from the resin with three times 5 mL (1 CV) elution buffer (50 mM Tris-HCl, 100 mM ArgCl, 150 mM NaCl and 500 mM imidazole, pH 8.0). Elution fractions were SDS-PAGE checked, and then combined. Pooled elution was buffer-exchanged with size-exclusion chromatography (SEC) running buffers: S-2P SEC running buffer (50 mM Tris-HCl, 100 mM ArgCl, 150 mM NaCl and 0.02% sodium azide (NaN_3), pH 8.0) and S-6P SEC running buffer (10 mM HEPES, 200 mM NaCl and 0.02% NaN_3, pH 7.5) to reduce the imidazole level below 1 mM and concentrated to ~1 mg/mL protein concentration with an Amicon Ultra-15-mL-centrifugal

filter (30 K, Millipore). Concentrated spike samples were flash frozen in liquid nitrogen and stored in -80 °C freezer.

All the spike samples were finally SEC purified before major experiments. 500 μL sample was thawed and spun at 15000 RCF for at least 10 minutes to remove possible protein aggregation pellet before injecting to Superose 6 10/300 GL column (Cytiva) on an AKTA Pure system (GE). With the SEC running buffer flowing at 0.5 mL/min rate, spike trimer was eluted at around 13 mL. The SEC purified spike trimers were concentrated to desired concentrations according to the experiment need.

Dynamic light scattering and thermal analysis

Dynamic light scattering (DLS) and thermal transition temperature (Tm) measurements were performed on a DynaPro NanoStar II (Wyatt) to characterize spike trimer integrity, homogeneity and thermal stability. SEC purified spike samples at 1 mg/mL were spun down at 15000 RCF for 20 minutes to remove aggregation pellet. 5 μL of the sample was then injected into a 2-μL quartz cuvette (Wyatt). Each DLS run was measured with twenty 10-second acquisitions at 25 °C by LASER auto-attenuation. Thermal stability determinations began DLS measurements at 30 °C and ended at 90 °C at a 3 °C/min rate.

Glycan analysis

Comprehensive glycol-profiling of the stabilized S-2P construct was conducted using 1.5 mg/mL spike protein. Four portions of samples were prepared for four digestion conditions. The samples were incubated in the denaturing buffer (25 mM Tris-HCl, 7 M guanidinium chloride (GdnHCl), and 50 mM dithiothreitol (DTT), pH 8.0) at 90 °C for 30 minutes. Denatured protein

was incubated with iodoacetamide (IAA) to a 100 mM final concentration at room temperature in the dark for alkylating the reduced cysteines. After 1-hour incubation, excess IAA was quenched with 50 mM DTT. Resulting solution was diluted ~11-fold with 10 mM Tris-HCl (pH 8.0) buffer containing 2 mM calcium chloride to reduce GndHCl concentration to 0.6 M. One half (275 µL) of current solution was separated and treated with 10 units of recombinant Peptide N-glycanase F (PNGase F) and incubated at 37 °C for 1 hour.[23] This enzyme could convert glycosylated Asn to deglycosylated Asp for later determining glycan occupancy.

Both PNGase F-treated and non-treated samples underwent protease digestion in parallel. Under four protease digestion conditions: Lys-C & Glu-C, trypsin, chymotrypsin, and alpha-lytic protease (Promega), respectively, protease was added at a ratio of 1:30 (w/w) to S-2P samples. All protease incubations were at 37 °C for 8 hours while Lys-C & Glu-C condition was performed with adding Lys-C for the first 4 hours before mixing with Glu-C. Digestion reactions were stopped with 0.02% formic acid (FA, Optima, Fisher) quenching.

Glycoform determination of S-2P was analyzed by nano LC-MS using an Orbitrap Fusion™ mass spectrometer (Thermo Fisher Scientific). Digested samples were desalted using Sep-Pak C18 cartridges (Waters) according to the manufacturer's protocol. A 2 cm trapping column and a 35 cm analytical column were freshly prepared in fused silica (100 µm) with 5 µM ReproSil-Pur C18 AQ beads (Dr. Maisch). 8 µL sample injection was run on a 60-minute linear gradient from 2% to 30% acetonitrile (ACN) with 0.1% FA, followed by 80% ACN solution for 10 minutes. The EThcD method was optimized with the following parameters: ion source at 2.1 kV for positive mode; ion transfer tube temperature at 350 °C; resolution: MS^1 = 120000, MS^2 = 30000; AGC target: $MS^1 = 2e^5$, $MS^2 = 1e^5$; and injection time: MS^1 = 50 ms, MS^2 = 60 ms.

Raw MS data were processed and visualized by Byonic™ and Byologic™ (Version 3.8, Protein Metrics Inc.) with a precursor mass tolerance at 6 ppm and fragment mass tolerance at 10 ppm. The glycopeptides were searched and determined using the N-glycan 309 mammalian database in Protein Metrics PMI-Suite as well as the assignments of correct c- and z- fragment ions. Highly confident entities were also manually validated by the presence of glycan oxonium ions m/z = 204 (HexNAc ions) and 366 (HexNAcHex ions). The relative abundance of each glycoform was determined by analyzing peak area in Byologic™. Glycoforms were classified in three major types: Oligo (Oligomannose), Hybrid, and Complex. Detailed classification including subtypes in Complex form, described in the previous study[24], was listed as follow: HexNAc(2)Hex(9-5) is M(annose)9 to M5; HexNAc(3)Hex(5-6) is Hybrid; HexNAc(3)Hex(3-4)X is A1 subtype; HexNAc(4)X is A2/A1B; HexNAc(5)X is A3/A2B and HexNAc(6)X is A4/A3B subtype. In addition, fucosylated Hybrid and Complex forms were separately listed as FHybrid and FComplex, respectively. Lastly, Glycan occupancy at each N-glycosylation site was determined in parallel, by computing the peak area ratio of the non-glycosylated (Asn) to the deglycosylated (Asp) glycopeptide.

Hydrogen/Deuterium-exchange mass spectrometry

For the hydrogen/deuterium-exchange (HDX) experiments in this chapter, 10 μL each of spike trimer (including D614, G614 S-2P and WT, N343A, N343Q and disulfide S-6P) at 1 mg/mL (pre-mixed with 1% internal standards Pro-Pro-Pro-Phe (PPPF)[25]) was incubated in 90 μL deuteration buffer (S-2P experiments: 10 mM Tris, pH* 8.0, 85% D_2O, Cambridge Isotope Laboratories, Inc.; S-6P experiments: 10 mM HEPES, pH* 7.5, 85% D_2O) at 23 °C for different timepoints. In the S-2P experiments featuring D614G mutation, 3-, 60-, 1800- and 72000-second

(three replicates for 3 seconds and two replicates for the rest) timepoints were applied, while in the S-6P experiments featuring N343 mutations, 3-, 60- and 900-second timepoints were applied. Exchanged sample (100 μL) was immediately mixed with 100 μL ice-cold quench buffer containing 8 M urea, 200 mM tris(2-chloroethyl) phosphate (TCEP) and 0.2% formic acid (FA) to pH 2.5, followed by flash frozen in liquid nitrogen.

All samples were thawed on ice for 5 minutes, loaded to the LC-MS system on a sample injection platform and analyzed by a Synapt G2 mass spectrometer (Waters) with the settings described in our previous study. Samples were in-line digested into peptides by an immobilized pepsin column (2.1 × 50 mm) and loaded onto CSH C18 trap cartridge (Waters) with loading buffer (2% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA)) at 200 μL/min. Peptides were then separated by a UPLC CSH C18 column (1.0 × 100 mm, 1.7 μm, Waters) with a 20-min linear gradient from 3% to 40% buffer B (buffer A: 2% ACN, 0.1% FA, 0.025% TFA; buffer B: 99.9% ACN, 0.1% FA) at a 40 μL/min flow rate. A series of wash steps with various of solvents were performed to remove carryover and regenerate pepsin column for next run.[26] A fully deuteration control was prepared by collecting digested samples from pepsin column, drying by speed-vaccum, resuspending in 10 μL SEC buffer and incubating in 90 μL deuteration buffer for 1 hour at 85 °C.

Peptides identification and mapping were performed by running digested samples on an Orbitrap Fusion™ mass spectrometer (Thermo Fisher Scientific) with the same settings mentioned in method glycan analysis. Peptide mass spectra were further confirmed on DriftScope (Waters) and identified with specific retention time and drift time. HDExaminer (Sierra Analytics) was used to process the mass spectrometer datasets and calculate deuterium uptake of each peptide identified by the specific retention time and drift time. In the cases of bimodal spectra, the mass envelopes were resolved by HX-Express v2 applying bimodal deconvolution and binomial fitting.[27]

SDS-PAGE and Native-PAGE

For each SDS-PAGE gel using Coomassie Blue staining, 12 μL spike sample was mixed with 4 μL 4× NuPAGE™ LDS loading dye (Invitrogen) before loading to 4-12% NuPAGE™ Bis-Tris gel (Invitrogen) well. SDS-PAGE gels were run at 120 Volts for 10 minutes and 150 Volts for 60 minutes in MES SDS running buffer (Invitrogen), followed by gel collection and PageBlue™ (Thermo Scientific) staining.

For each Native-PAGE gel, 15 μL spike sample was mixed with 5 μL 4× NativePAGE™ loading dye (Invitrogen) before loading to 4-12% NativePAGE™ Bis-Tris gel (Invitrogen) well. The gel was set up with mixed 1× Native-PAGE cathode buffer (Invitrogen) and 1× Native-PAGE running buffer (Invitrogen) inside the gel chamber and 1× Native-PAGE running buffer outside the gel chamber and run on ice at 150 Volts for 2 hours.

Both SDS-PAGE and Native-PAGE gels were de-stained in wash buffer (10% methanol and 7% acetic acid) and rinsed with Mili-Q (Millipore) water before imaging. Gels were imaged on the Odyssey® M-XS imaging system (LI-COR Biosciences) using 700 nm channel and optimized exposure intensity.

2.3 Results

D614G mutation increases spike stability and recombinant protein yields

The D614G variant that arose from the ancestral Wuhan SARS-CoV-2 strain (WT, Hu-1), was the first remarkable mutation that emerged as the virus adapted to enhance its transmission through the human population.[15] To investigate how D614G mutation affects spike (S) stability and helps in virus proliferation, we purified the soluble S ectodomain, tethered by a trimerization domain called foldon for both Wuhan-Hu-1 (D614) and D614G mutant (G614) (Figure 1a). S-2P

D614 and G614 constructs included a di-proline substitution at the apex of HR1 and CH domains and replaced the native polybasic RRAR sequence at the S1/S2 cleavage site with SGAG sequence to maintain the S trimers close to their native, prefusion conformations (Figure 1a). Both constructs were transfected and expressed in mammalian Expi293F cells to maintain native glycosylation and proper folding of trimer assemblies.

From all protein purification trials in this chapter, recombinant S-2P G614 protein consistently exhibited 5-to-10-fold higher expression levels and trimer yields compared to S-2P D614 from parallel purifications (Table S1). SDS-PAGE gels with each His-tag affinity purification step and pooled spike elution (Figure 1c, band above 160 kDa) from a 1 L Expi293F culture, revealed similar levels of common Expi293F secreted protein in the supernatant and flow-through (FT) but a higher level of S-2P G614 than D614 (Figure 1c). Although denaturing gels are limited in reflecting yields of well-folded trimers, a native-PAGE gel run alongside showed that both purified spikes D614 and G614 in the pooled elution were predominantly trimeric (Figure 1d, band above 720 kDa).

Size-exclusion chromatography (SEC) also indicated that the trimeric spike protein peak at ~13 mL represented over 90% of the sample inputs for both S-2P D614 and G614 purification (Figure 1b). The best post-SEC trimer yields were 1.2 mg for S-2P G614 and 0.2 mg for D614 from 1 L Expi293F culture (Table S1). The SEC chromatogram also suggested that the size of S-2P D614 trimer was slightly bulkier than G614 trimer, as the peak eluted slightly earlier in the SEC running buffer at 4 °C.

The difference in trimer sizes was also confirmed by dynamic light scattering (DLS) measurements. Under the same purification process and buffer condition at 25 °C, S-2P D614 trimer showed an averaged hydration radius of 9.6 nm, while the G614 trimer's averaged hydration

radius was 8.8 nm (Table S1). Furthermore, by monitoring the DLS signals over the course of thermal denaturation, the global protein stability could be measured and compared (Figure 1e). The D614 trimer began thermal denaturation, resulting in protein aggregation and increased turbidity, around 50 ˚C. In contrast, the G614 trimer remained stable as its radius of hydration remained constant until over 60 ˚C before it began to show very modest increase in its hydration radius by DLS (Figure 1e).[28]

These findings reveal that the D614G mutation can significantly increase the trimer stability and boost the soluble S-2P production several-fold.

D614G mutation dynamically stabilizes spike open state

Residue 614 is centrally positioned at the interface of the spike receptor-binding subunit (S1) and fusion subunit (S2) (Figure 1a, Figure 2c). Previous cryo-EM structures suggested that the D614G mutation could help strengthen the interaction between these two subunits, which could help explain the increased thermal stability we observed for S-2P G614.[29] To comprehensively understand how this single point mutation dynamically impacts on spike trimer stability and conformational states, we have performed HDX-MS analysis on both D614 and G614 S-2P trimers. Because of the nearly identical sequence alignment between these two constructs, the HDX peptide sequence coverages could reach 85% with 11 N-glycosylation sites covered for both D614 and G614 S-2P trimers (Figure S2), and pairwise comparison between homologous peptides could be easily conducted (Figure 2, Figure S3).

The HDX butterfly plots (Figure 2a) and color-coded heatmaps (Figure S1) provided the detailed visualization of integral spike dynamics, as well as at the domain level, reflecting by deuterium uptake levels over time. Differential uptake plots comparing D614 and G614 S-2P,

subtracted from the corresponding butterfly plots, revealed significant differences caused by the single D614G substitution (Figure 2b). Two regions exhibited substantial dynamic differences: the C-terminal RBD hinge region, where the RBD pivots relative to the rest of S1 subunit (Figure 2c, peptide #1, Figure S3); and most of the S2 subunit located at the bottom half of the trimer (Figure 2b, HR1-CH, Figure 2c, peptide #2-4, Figure S3).

Increased deuteration across the RBD C-terminal peptide (Figure 2c, peptide #1) in the G614 mutant indicated greater solvent accessibility in this region, is consistent with the RBD populating the up conformation at higher levels on average than D614. This supported cryo-EM data suggesting the D614G mutation could shift the dynamic equilibrium towards RBD-up conformations.[16, 18] Regions associated with RBD-up conformations, such as the apex of central helical bundles, also showed slightly more dynamic behavior in G614 over D614, supporting that S G614 possessed more open state trimers (Figure S3).

Additionally, the D614G mutation resulted in greater ordering at the S2 base regions, spatially below the interfaces around the mutation (Figure 2c, Figure S3).[18,29] Notably, peptides at the bottom of the HR1 and CH domains including many secondary structure regions in the D614 construct, exhibited greater levels of deuterium uptake (10% or higher difference) along the time course (Figure 2c, peptide #2-4, Figure S3). The weaker structural ordering behavior in such secondary structure at bottom region and central helices suggested an overall de-stabilization at trimer basis in D614 construct.

Besides the peptides discussed above, some peptides in the S2 subunit also indicated the weaker structural ordering in S-2P D614, although the deuterium uptake differences were not over 10% in most time points (Figure 3). Instead, these peptides in D614 construct exhibited bimodal mass spectral envelopes, with one distinguishable slow-uptake population resembled the dynamic

behaviors of peptides in G614 construct (Figure 3, purple population) and one fast-uptake population (green population). Some fast-uptake populations reached the total deuteration level in 3 seconds to 1 minute, suggesting highly exposed, possibly unstructured states these peptide regions could adopt. Peptides at intrinsically more dynamic regions, such as hinge (residues 542-568) and helical apex (residues 978-990), fast-uptake populations were merged with slow-uptake populations in between 3 seconds and 1 minute, while at intrinsically less dynamic regions (residues 892-902, 1175-1188), the bimodal behaviors could stay over 1 minute.

From our HDX dynamic analysis, the local structural ordering in S2 subunit is significantly increased by this single point mutation D614G. The differences observed in local dynamics are consistent with global structural stability and homogeneity variations, reflected by the thermal stability and size differences reported in the previous section. The stable structural ordering in S2 subunit from D614G mutation also reduces the energy cost of spike adopting energetically unfavorable RBD-up conformations, promoting G614 trimer to dynamically adopt open states.

N343 glycosylation aids RBD conformational switch

Given the glycoprotein nature of S, the effects of glycans on S structure and dynamics are also important to consider. We performed glycan analysis on the purified S-2P to understand the N-glycosylation patterns. 16 out of 22 N-glycosylation sites were analyzed with high confidence. This analysis demonstrated that 4 sites predominantly featured high mannose forms, 11 sites predominantly featured complex forms and 1 site showed a mixture (Figure S4). Most of the sites we analyzed showed an agreement in glycoforms and subtypes with previously published results.[24] In addition, the predominance of complex forms at the 2 N-glycosylation sites on the RBD (N331

and N343), was highly conserved across various constructs reported previously, including the RBD monomer, RBD trimer, S-2P trimer and native S trimer.[7]

The N343 glycan, located near RBM on RBD, is functionally important to S. It was suggested by molecular dynamics (MD) simulations to modulate RBD movement.[20, 21] To further examine this statement in real experiments and investigate how N343 glycan affects entire S structural stability and drives RBD motion, we performed HDX on S-6P possessing N343 glycan knock-down mutations N343A and N343Q with S-6P G614 as WT control and disulfide mutated (SS) S-6P as all-RBD-down control.

In the regions sensing the most structural environment changes from conformational switches, both C-terminal RBD hinge (Figure 4, peptide #3) and apex of HR1-CH helical bundles (peptide #6-8) exhibited substantial dynamic differences resulting from modification of the N343 residue. N343A and N343Q mutants showed identical deuterium uptake levels, intermediate between those from the WT and SS. Compared to the dynamics in WT construct, which samples equilibration between open and closed conformations, the N343 knock-down mutants showed limited to moderate dynamic motions of RBD. However, the solvent exposure resulting from RBD dynamic motions in these mutants was still significantly higher than in the all-RBD-down disulfide mutant (Figure 4, peptide #3, #6, and #8).

N343 mutants showed notable dynamic effects at the S1 hinge region (Figure 4, peptide #4, and #5). At the N-terminal loop supporting the RBD, both N343 mutants exhibited similar deuterium uptake profiles to the WT, reflecting a flexible and exposed environment as in the WT (Figure 4, peptide #4). However, at another loop that closely interacts with the NTD of the neighboring protomer, both N343 mutants showed overlapping uptake plots with the SS mutant (Figure 4, peptide #5). Considering that the NTD could also be packed differently in parallel with

RBD conformations, this phenomenon suggested that in the N343 glycan knock-down mutants, the NTD remained in the closed conformations regardless of RBD movements.

The engagement of NTD in RBD movements was also supported by the divergence in NTD dynamics between the N343A and N343Q mutants. At two β-loop-β sites on the NTD, N343Q behaved similarly to the WT construct with normal RBD movements, while N343A behaved more like the SS mutant with limited RBD movements (Figure 4, peptide #1 and #2). The glutamine at position 343 could better mimic the native molecular interactions of asparagine than an alanine would. The dynamic differences in the NTD region suggested that the original N343 glycosylation impacts the RBD conformational switch through crosstalk between the RBD and the neighboring NTD.

2.4 Discussion

Class I fusion proteins, such as S glycoprotein, are highly dynamic in nature. This is likely due in part to their central function in undergoing significant conformational changes from pre-fusion to post-fusion forms during the viral entry process.[30, 31] Dynamic sampling among various conformations of metastable pre-fusion forms has been anticipated and often occurred.[31, 32]

The phenotype of SARS-CoV-2 virus acquiring S D614G mutation was seen to be highly infectious from the facts that it became prevalent in two months with increasing R_0 (WHO). The virus gained evolutionary advantages from this mutation, facilitating higher percentage of S in open states, which was observed in several cryo-EM studies.[16-18] At the same time, the increased S density displayed on the viral membrane was also observed from pseudotyped virus and VLP models.[33]

Through our protein expression and HDX experiments, we observed and investigated the effects of the D614G mutation on both boosting the yield and favoring more open conformational states of the spike trimer. Our experiments captured the entire protein dynamics and compared the spatial dynamic differences at peptide resolution. Group of peptides directly featuring RBD conformational changes were identified, while other peptides exhibited dynamic differences over time or switched to highly disordered states in the D614 variant, were found primarily at the S2 base. In contrast, the G614 variant maintained more ordered and protected peptide backbones in these regions. This structural stabilization of the S2 subunit lowered the free energy cost of adopting metastable open conformations. In addition, the well-ordered S2 subunit also improved structural homogeneity, suggesting higher trimer production with reduced misfolding and aggregation risks.

While virus gains evolutionary advantages from mutations, conserved elements can also have significant impacts. The occupancy and composition of many N-glycosylation sites on S were well-conserved, across various protein constructs, expression batches, and conditions, including the native condition.[7, 13, 24] The N343 glycosylation site, in particular, is highly conserved. MD simulations predicted that knocking down N343 glycosylation could impair the RBDs' ability to adopt open conformation.[21]

We examined this hypothesis in HDX experiments. By constructing positive control (WT) and negative control (SS) of spike open states, we have confirmed the group of peptides identified in D614G experiments had the abilities in reflecting RBD conformational switching and used for our future analysis. At the same time, differential dynamic behaviors occurred at NTDs, indicated their varying positions and dynamics in open and closed conformations, which were also observed in other dynamic studies. From our HDX experiment, a more comprehensive model on S

conformational states was established from the dynamics perspective. A long-range allosteric effect between RBD conformational switches and well-ordered key functional regions in the S2 was facilitated by NTD-RBD crosstalk, aided by N-glycosylation.

Temperature effects on S conformational states were not examined in this study but have been observed by other groups.[34, 35] Designing HDX experiments on testing protein under different temperature conditions or varying exchange temperatures could be challenging. However, our HDX experiments were comprehensively designed and rigorously controlled to minimize potential confounding factors.

In conclusion, in this chapter, we have used HDX-MS to investigate structural dynamics of pre-fusion S-2P assemblies under various conditions. Two notable dynamic impacts affecting trimer conformational switches were discovered and summarized. The spike trimer could favor more towards RBD-up open states from the prevalent mutation D614G, which lowered energy cost from the stabilized S2 subunit; and through the packed RBD-NTD crosstalk achieved from conserved N343 glycosylation on RBD.

These findings revealed that in the highly dynamic Class I fusion protein machinery, allosteric effects through dynamic motions and interactions between domains can be ubiquitous and biologically relevant. Despite high sequence similarity, tiny amino acid mutations or post-translational modification differences could significantly alter protein structural conformations and impact biological function through allosteric dynamic effects along the entire protein structure. This behavior necessitates a comprehensive perspective to understand structure- and dynamics-related biological questions, utilizing various structural techniques that have expertized and complemented each other.

Figures

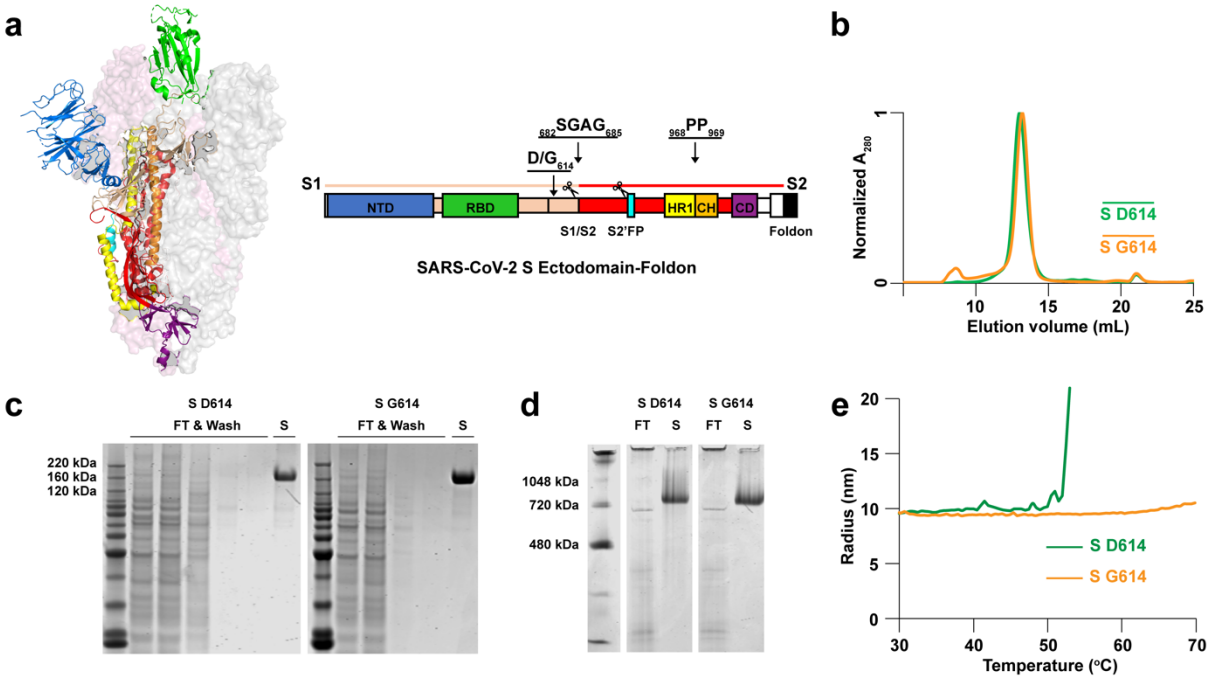

Figure 1. **Structural organization and characterization of SARS-CoV-2 S-2P D614 and G614.** **(a)** Color-coded spike domains on cartoon diagram (PDB: 6VSB) and sequence displays. S1: receptor-binding subunit; S2: fusion subunit; NTD: N-terminal domain; RBD: receptor-binding domain; FP: fusion peptide; HR1: heptad repeat 1; CH: central helix; CD: connector domain. **(b)** SEC chromatograms of S-2P D614 and G614 purification. Major peak eluting at ~13 mL reveals the trimer integrity and homogeneity. **(c)** SDS-PAGE gels imaging S-2P D614 and G614 His-tag purification fractions. Spike bands locate slightly above 160 kDa. **(d)** Native-PAGE gels imaging S-2P D614 and G614 trimers. Trimeric spike bands locate above 720 kDa. **(e)** Thermal denaturing curves of S-2P D614 and G614 from DLS. S-2P D614 denatures 10 ˚C lower than G614.

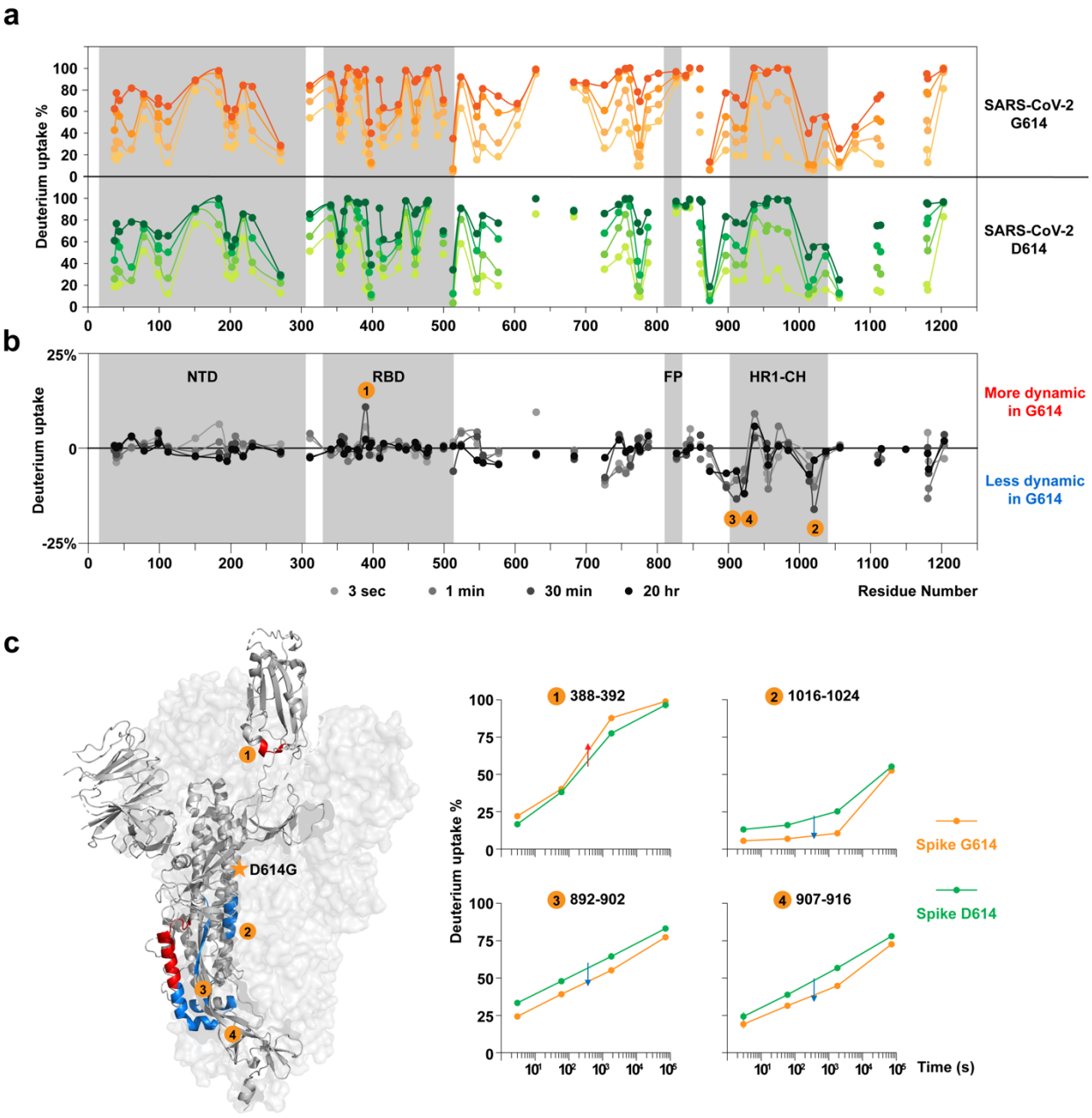

Figure 2. **Dynamic differences between S-2P D614 and G614 trimers. (a)** Butterfly plots and **(b)** differential plots of S-2P D614 versus G614 reveal the similarities and differences in dynamic behavior across the spike sequences. Grey shadings highlight the RBD, fusion peptide and HR1-CH region. **(c)** Differential heatmaps and uptake plots for selected peptides showing prominent differences in G614 compared to D614 spikes. In the cartoon diagram heatmap, red indicates more exposure and blue indicates more protection under D614G mutation. G614 construct samples more

open conformations with exposed RBD C-terminal peptide #1 and exhibits a more stabilized S2 subunit than D614 construct reflected by protections on peptides #2-4.

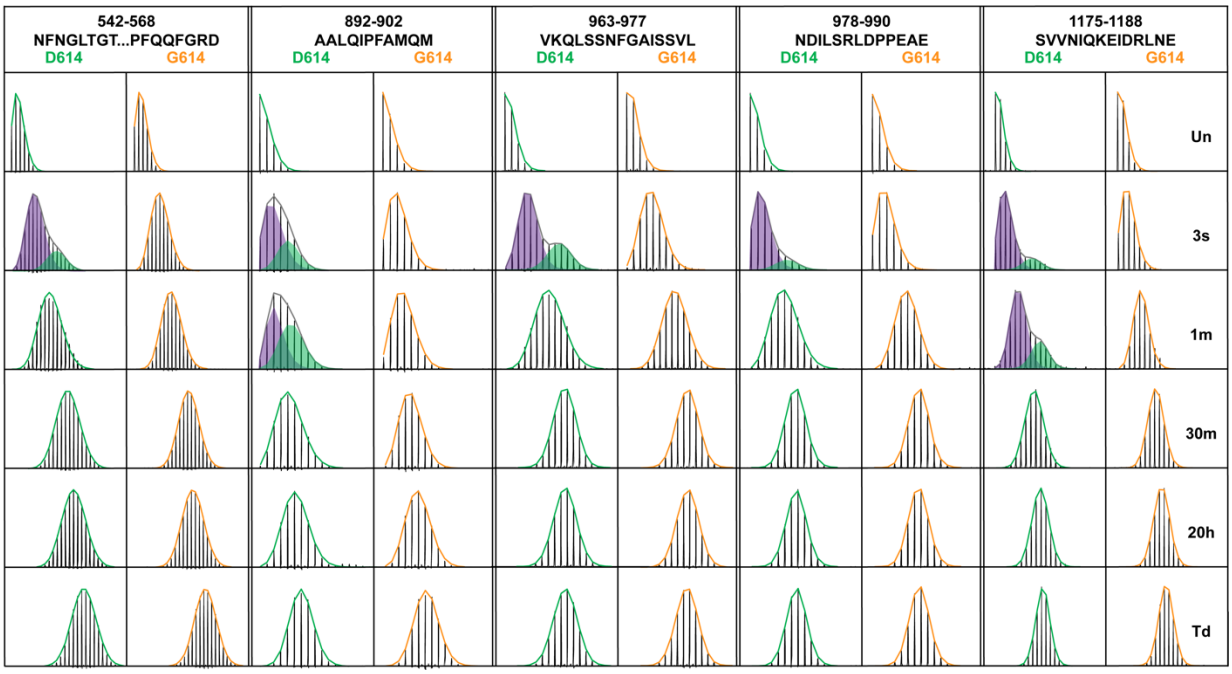

Figure 3. **Differences in structural ordering revealed by bimodal uptake profile.** Bimodal m/z spectra show the different dynamic samplings for selected peptides in D614 construct compared to the G614 construct. Un: undeuterated control; Td: totally deuterated control. Broadened mass envelopes are deconvoluted into bimodal. Regression testing with 95% confidence level and Ashman's D statistics > 2 are used to evaluate bimodal distributions. Slow-uptake populations are highlighted in purple and fast-uptake populations are highlighted in green.

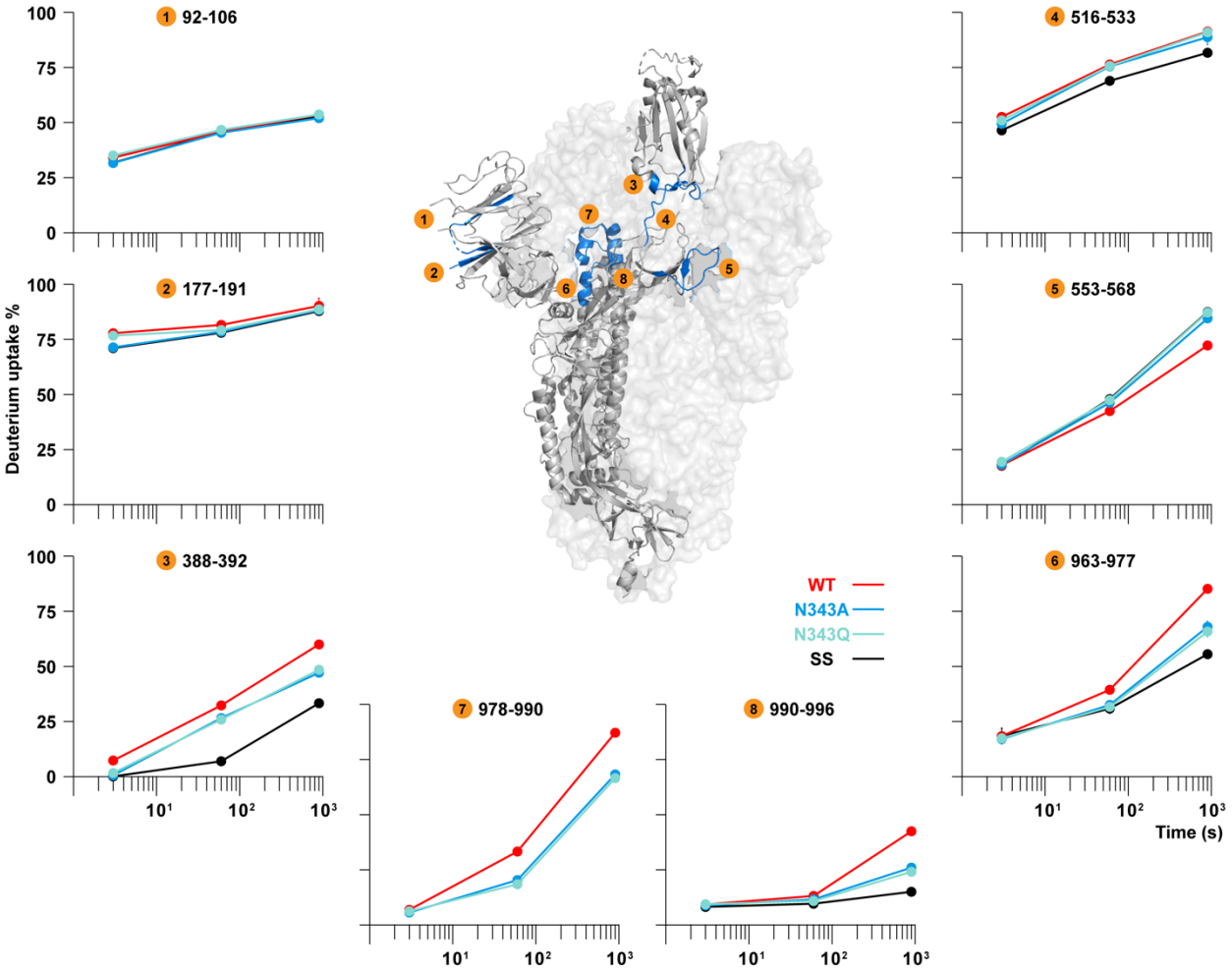

Figure 4. **Dynamic impacts from N343 glycosylation on S conformational states.** Deuterium uptake plots for selected peptides show prominent differences in N343 glycan knock-down mutants. Peptides #1-2 are featuring NTD. Peptide #3 is on the C-terminal of the RBD. Peptide #4-5 are in S1 hinge region. Peptides #6-8 are around the apex of central helical bundle. Disulfide mutated SS construct's uptake plot is not included in peptide #7 due to D985C mutation in SS.

Supplemental Information

Table S1. DLS measurements and per liter culture protein yields after SEC purification of S-2P D614 and G614 trimers.

Spike	Radius (nm)	PD (%)	Yield (mg)
D614	9.6±0.2	7.9±2.0	0.1-0.2*
G614	8.8±0.2	7.2±3.8	0.7-1.2

* Due to the protein loss from SEC purification, pre-SEC samples below 0.2 mg are pooled with other preparations for higher yield after SEC.

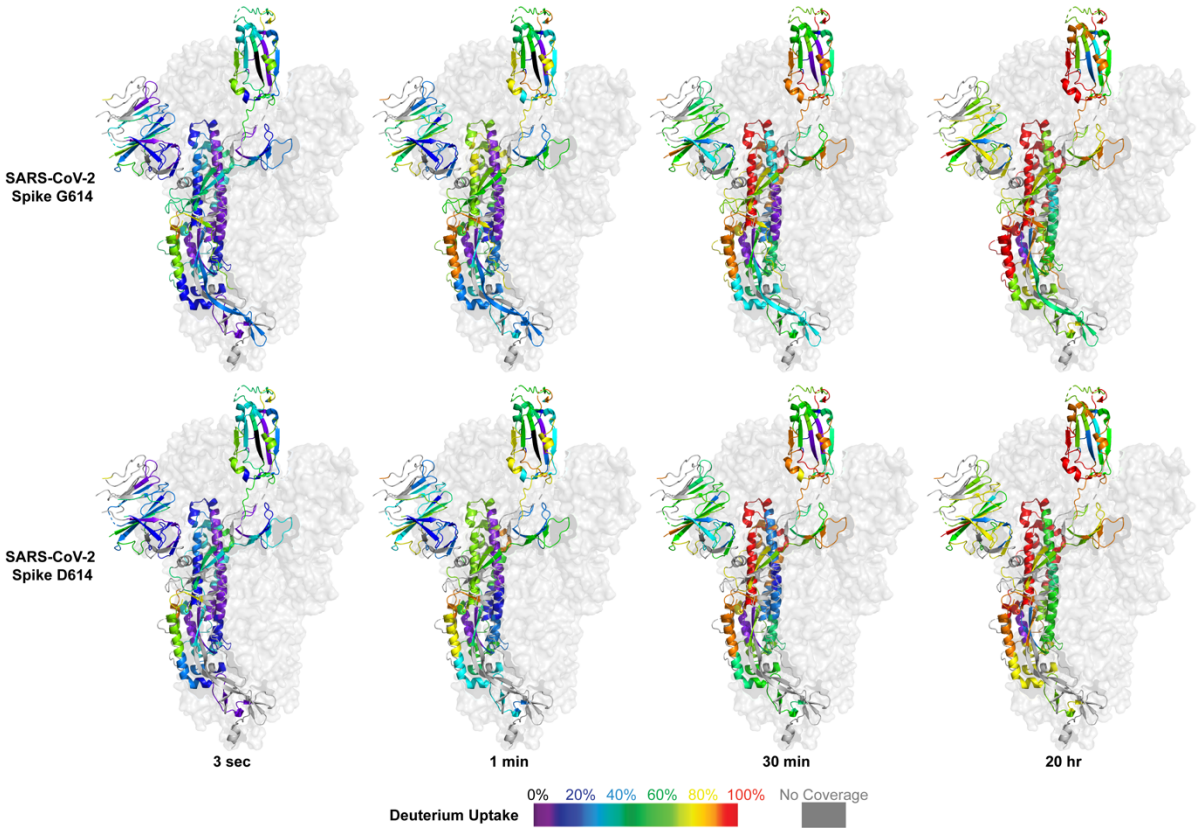

Figure S1. **HDX heatmap of SARS-CoV-2 S-2P D614 and G614.** Color-coded D614 and G614 deuterium uptake heatmaps are shown in peptide level resolution at 3 sec, 1 min, 30 min, and 20 hr labeling timepoints.

Figure S2. **HDX coverage map of SARS-CoV-2 S-2P D614 and G614.** Peptide coverage map is from pepsin digestion. Blue shade represents RBD sequence; Orange shade represents engineered mutations at S1/S2 and HR1-CH apex sites; Pink shade represents fusion peptide.

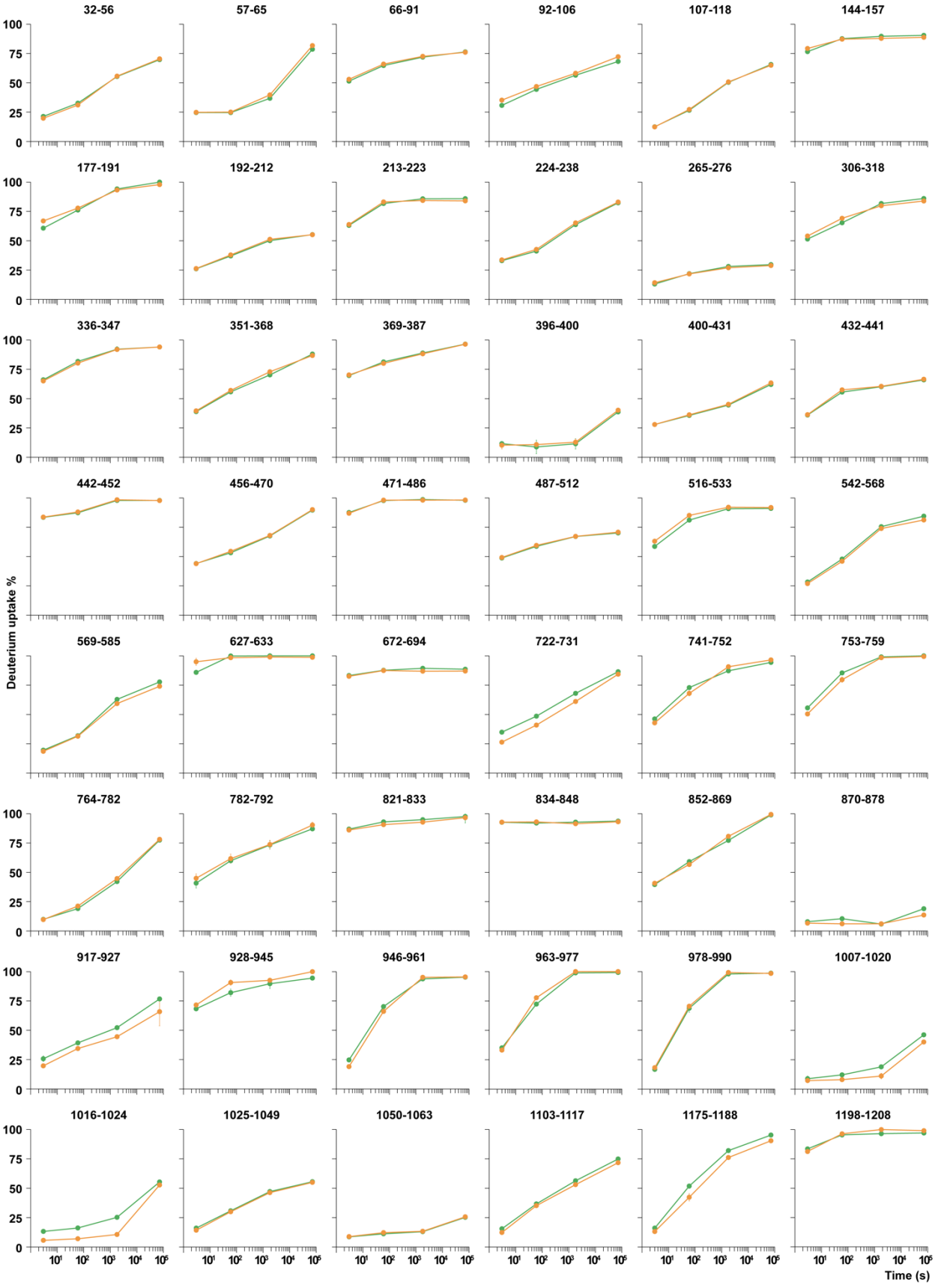
32-56
57-65
66-91
92-106
107-118
144-157
177-191
192-212
213-223
224-238
265-276
306-318
336-347
351-368
369-387
396-400
400-431
432-441
442-452
456-470
471-486
487-512
516-533
542-568
569-585
627-633
672-694
722-731
741-752
753-759
764-782
782-792
821-833
834-848
852-869
870-878
917-927
928-945
946-961
963-977
978-990
1007-1020
1016-1024
1025-1049
1050-1063
1103-1117
1175-1188
1198-1208
Deuterium uptake %
Time (s)

Figure S3. **Deuterium uptake plots of SARS-CoV-2 S-2P D614 and G614.** This figure is related to Figure 2.

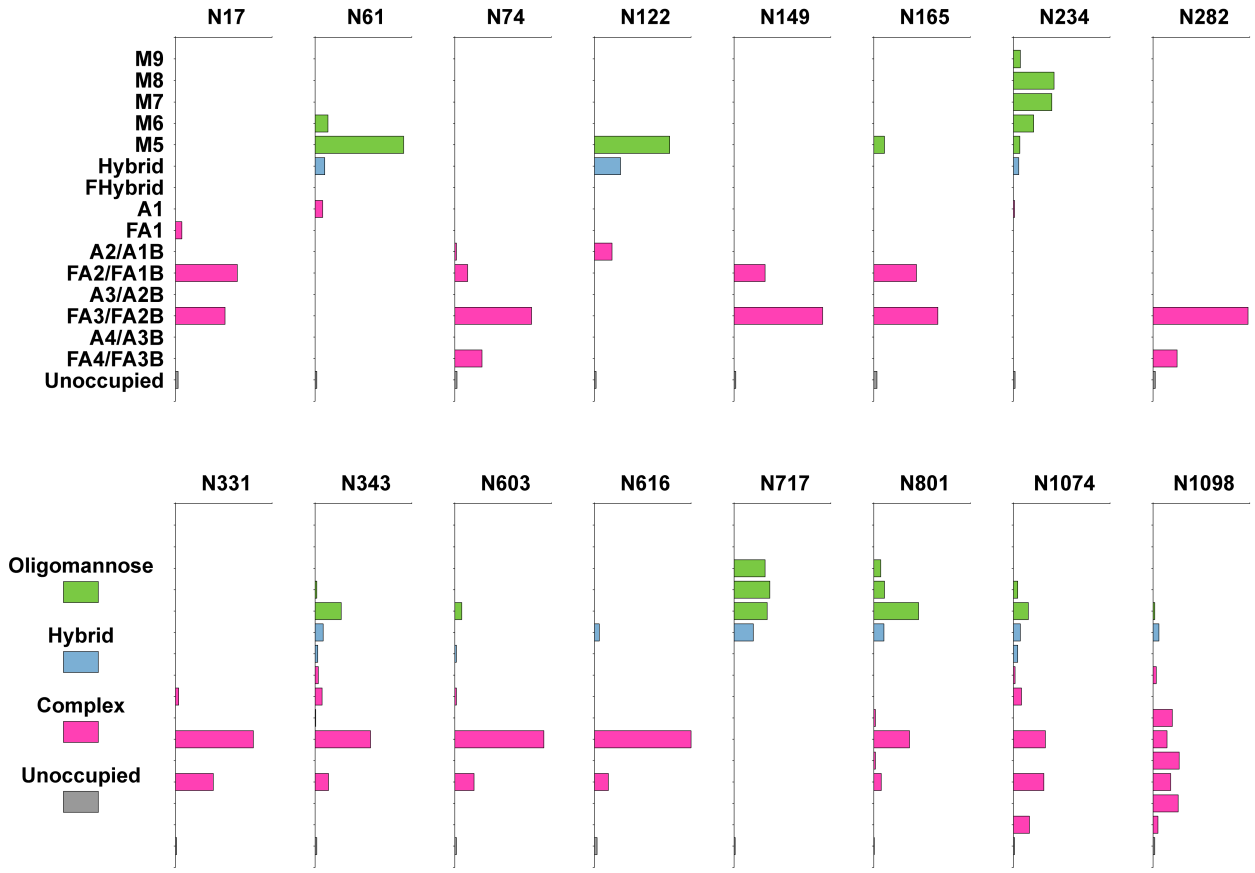

Figure S4. **Glycan compositions of each N-glycosylation site on SARS-CoV-2 S-2P.** The definition of each glycan type and subtype is described in the glycan analysis method section. N657, N709, N1134, N1158, N1173 and N1194 are not characterized. All N-glycosylation sites characterized show high glycan occupancy.

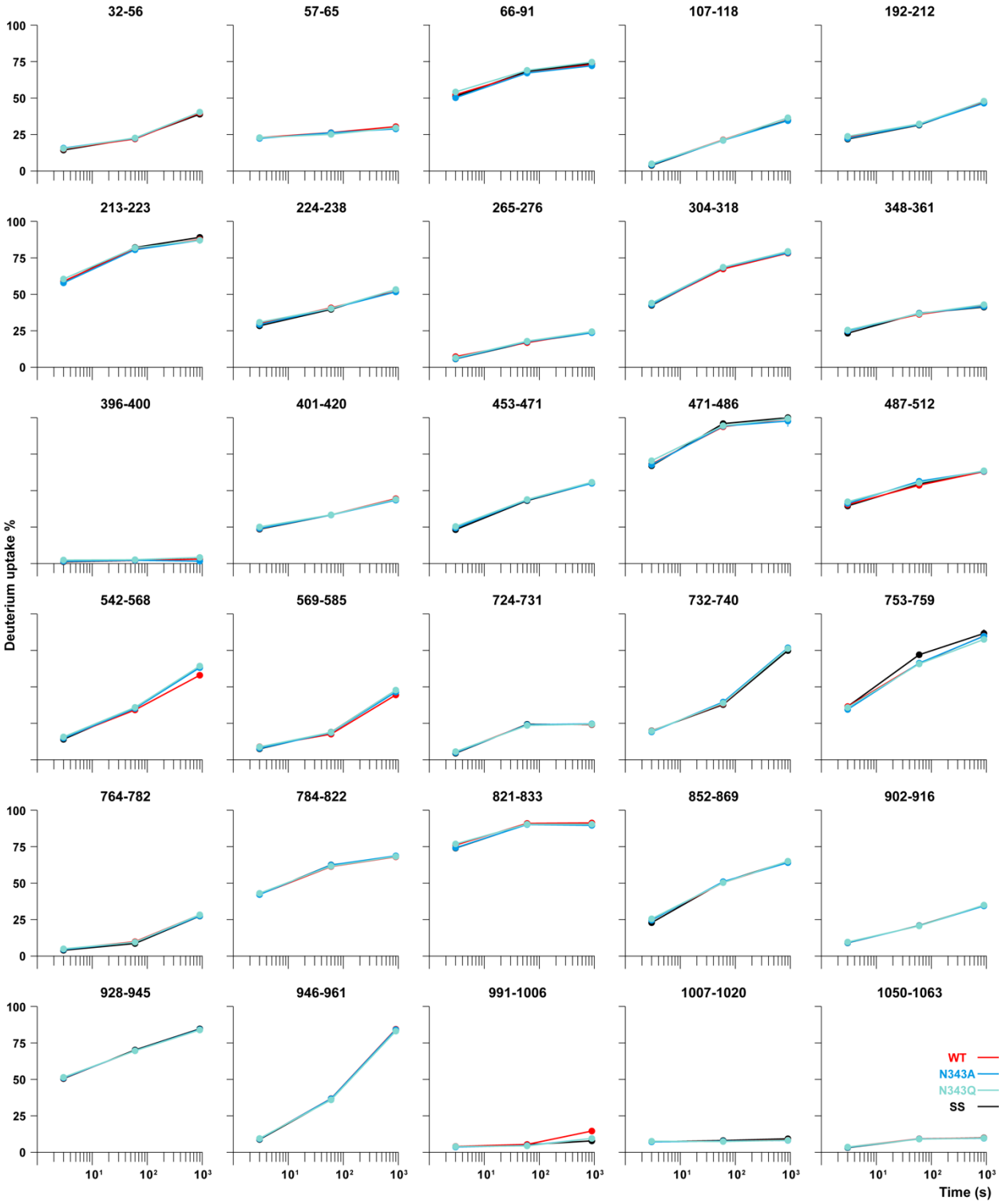

Figure S5. **Deuterium uptake plots of SARS-CoV-2 S-6P N343 mutants**. This figure is related to Figure 4.

References

(1) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **2020**, *382* (8), 727-733. DOI: 10.1056/NEJMoa2001017.

(2) 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. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **2013**, *503* (7477), 535-538. DOI: 10.1038/nature12711.

(3) Tortorici, M. A.; Veesler, D. Structural insights into coronavirus entry. *Adv Virus Res* **2019**, *105*, 93-116. DOI: 10.1016/bs.aivir.2019.08.002.

(4) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579* (7798), 265-269. DOI: 10.1038/s41586-020-2008-3.

(5) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367* (6483), 1260-1263. DOI: 10.1126/science.abb2507.

(6) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e286. DOI: 10.1016/j.cell.2020.02.058.

(7) 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.; et al. Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell* **2020**, *183* (5), 1367-1382.e1317. DOI: 10.1016/j.cell.2020.10.043.

(8) 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. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **2020**, *369* (6510), 1501-1505. DOI: doi:10.1126/science.abd0826.

(9) Bangaru, S.; Ozorowski, G.; Turner, H. L.; Antanasijevic, A.; Huang, D.; Wang, X.; Torres, J. L.; Diedrich, J. K.; Tian, J. H.; Portnoff, A. D.; et al. Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science* **2020**, *370* (6520), 1089-1094. DOI: 10.1126/science.abe1502.

(10) Cai, Y.; Zhang, J.; Xiao, T.; Peng, H.; Sterling, S. M.; Walsh, R. M., Jr.; Rawson, S.; Rits-Volloch, S.; Chen, B. Distinct conformational states of SARS-CoV-2 spike protein. *Science* **2020**, *369* (6511), 1586-1592. DOI: 10.1126/science.abd4251.

(11) Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* **2017**, *27* (1), 119-129. DOI: 10.1038/cr.2016.152.

(12) 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. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2019**, *176* (5), 1026-1039.e1015. DOI: 10.1016/j.cell.2018.12.028.

(13) Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; et al. Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020**, *183* (3), 730-738.e713. DOI: 10.1016/j.cell.2020.09.018.

(14) Lu, M.; Uchil, P. D.; Li, W.; Zheng, D.; Terry, D. S.; Gorman, J.; Shi, W.; Zhang, B.; Zhou, T.; Ding, S.; et al. Real-Time Conformational Dynamics of SARS-CoV-2 Spikes on Virus Particles. *Cell Host Microbe* **2020**, *28* (6), 880-891.e888. DOI: 10.1016/j.chom.2020.11.001.

(15) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020**, *182* (4), 812-827.e819. DOI: 10.1016/j.cell.2020.06.043.

(16) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **2020**, *183* (3), 739-751.e738. DOI: 10.1016/j.cell.2020.09.032.

(17) Gobeil, S. M.; Janowska, K.; McDowell, S.; Mansouri, K.; Parks, R.; Manne, K.; Stalls, V.; Kopp, M. F.; Henderson, R.; Edwards, R. J.; et al. D614G Mutation Alters SARS-CoV-2 Spike Conformation and Enhances Protease Cleavage at the S1/S2 Junction. *Cell Rep* **2021**, *34* (2), 108630. DOI: 10.1016/j.celrep.2020.108630.

(18) Benton, D. J.; Wrobel, A. G.; Roustan, C.; Borg, A.; Xu, P.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. *Proc Natl Acad Sci U S A* **2021**, *118* (9). DOI: 10.1073/pnas.2022586118.

(19) Yang, Z.; Han, Y.; Ding, S.; Shi, W.; Zhou, T.; Finzi, A.; Kwong, P. D.; Mothes, W.; Lu, M. SARS-CoV-2 Variants Increase Kinetic Stability of Open Spike Conformations as an Evolutionary Strategy. *mBio* **2022**, *13* (1), e0322721..

(20) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; et al. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent Sci* **2020**, *6* (10), 1722-1734. DOI: 10.1021/acscentsci.0c01056.

(21) Sztain, T.; Ahn, S. H.; Bogetti, A. T.; Casalino, L.; Goldsmith, J. A.; Seitz, E.; McCool, R. S.; Kearns, F. L.; Acosta-Reyes, F.; Maji, S.; et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat Chem* **2021**, *13* (10), 963-968. DOI: 10.1038/s41557-021-00758-3.

(22) Herrera, N. G.; Morano, N. C.; Celikgil, A.; Georgiev, G. I.; Malonis, R. J.; Lee, J. H.; Tong, K.; Vergnolle, O.; Massimi, A. B.; Yen, L. Y.; et al. Characterization of the SARS-CoV-2 S Protein: Biophysical, Biochemical, Structural, and Antigenic Analysis. *ACS Omega* **2021**, *6* (1), 85-102. DOI: 10.1021/acsomega.0c03512.

(23) Krenkova, J.; Szekrenyes, A.; Keresztessy, Z.; Foret, F.; Guttman, A. Oriented immobilization of peptide-N-glycosidase F on a monolithic support for glycosylation analysis. *J Chromatogr A* **2013**, *1322*, 54-61. DOI: 10.1016/j.chroma.2013.10.087.

(24) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, *369* (6501), 330-333. DOI: 10.1126/science.abb9983.

(25) Zhang, Z.; Zhang, A.; Xiao, G. Improved protein hydrogen/deuterium exchange mass spectrometry platform with fully automated data processing. *Anal Chem* **2012**, *84* (11), 4942-4949. DOI: 10.1021/ac300535r.

(26) Verkerke, H. P.; Williams, J. A.; Guttman, M.; Simonich, C. A.; Liang, Y.; Filipavicius, M.; Hu, S. L.; Overbaugh, J.; Lee, K. K. Epitope-Independent Purification of Native-Like Envelope Trimers from Diverse HIV-1 Isolates. *J Virol* **2016**, *90* (20), 9471-9482. DOI: 10.1128/jvi.01351-16.

(27) Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *J Am Soc Mass Spectrom* **2013**, *24* (12), 1906-1912. DOI: 10.1007/s13361-013-0727-5.

(28) Juraszek, J.; Rutten, L.; Blokland, S.; Bouchier, P.; Voorzaat, R.; Ritschel, T.; Bakkers, M. J. G.; Renault, L. L. R.; Langedijk, J. P. M. Stabilizing the closed SARS-CoV-2 spike trimer. *Nat Commun* **2021**, *12* (1), 244. DOI: 10.1038/s41467-020-20321-x.

(29) Zhang, J.; Cai, Y.; Xiao, T.; Lu, J.; Peng, H.; Sterling, S. M.; Walsh, R. M., Jr.; Rits-Volloch, S.; Zhu, H.; Woosley, A. N.; et al. Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **2021**, *372* (6541), 525-530. DOI: 10.1126/science.abf2303.

(30) Colman, P. M.; Lawrence, M. C. The structural biology of type I viral membrane fusion. *Nat Rev Mol Cell Biol* **2003**, *4* (4), 309-319. DOI: 10.1038/nrm1076.

(31) Benhaim, M. A.; Lee, K. K. New Biophysical Approaches Reveal the Dynamics and Mechanics of Type I Viral Fusion Machinery and Their Interplay with Membranes. *Viruses* **2020**, *12* (4). DOI: 10.3390/v12040413.

(32) Hodge, E. A.; Benhaim, M. A.; Lee, K. K. Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci* **2020**, *29* (4), 843-855. DOI: 10.1002/pro.3790.

(33) Zhang, L.; Jackson, C. B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B. D.; Rangarajan, E. S.; Pan, A.; Vanderheiden, A.; Suthar, M. S.; et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat Commun* **2020**, *11* (1), 6013. DOI: 10.1038/s41467-020-19808-4.

(34) Edwards, R. J.; Mansouri, K.; Stalls, V.; Manne, K.; Watts, B.; Parks, R.; Janowska, K.; Gobeil, S. M. C.; Kopp, M.; Li, D.; et al. Cold sensitivity of the SARS-CoV-2 spike ectodomain. *Nat Struct Mol Biol* **2021**, *28* (2), 128-131. DOI: 10.1038/s41594-020-00547-5.

(35) Costello, S. M.; Shoemaker, S. R.; Hobbs, H. T.; Nguyen, A. W.; Hsieh, C. L.; Maynard, J. A.; McLellan, J. S.; Pak, J. E.; Marqusee, S. The SARS-CoV-2 spike reversibly samples an open-

trimer conformation exposing novel epitopes. *Nat Struct Mol Biol* **2022**, *29* (3), 229-238. DOI: 10.1038/s41594-022-00735-5.

Chapter 3. Human ACE2 Receptor-Induced Allosteric Activation in *Sarbecovirus* Spike Assemblies during Fusion Process

The content in this chapter is reproduced with permission from: Chen, C.; Zhu, R.; Hodge, E. A.; Díaz-Salinas, M. A.; Nguyen, A.; Munro, J. B.; Lee, K. K. hACE2-Induced Allosteric Activation in SARS-CoV versus SARS-CoV-2 Spike Assemblies Revealed by Structural Dynamics. *ACS Infectious Diseases* **2023**, *9* (6), 1180-1189. DOI: 10.1021/acsinfecdis.3c00010.

3.1 Introduction

Viruses in the *Sarbecovirus* subgenus (*Betacoronavirus* genus), have caused two significant outbreaks in the past two decades, greatly impacting the human society. SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus), emerged in 2002 and caused highly pathogenic and deadly infections during the outbreak. Fortunately, it was successfully contained before it could spread widely across large populations.[1] However, SARS-CoV-2 virus, which initiated the COVID-19 pandemic from the first documented case in December 2019,[2, 3] has not yet been contained. Compared to SARS-CoV, SARS-CoV-2 infection exhibits a longer incubation time, greater level of asymptomatic transmission, and overall higher transmissibility.[4] Furthermore, this virus has spread globally, and brought public health challenges with emerging waves of variants that exhibit enhanced transmissibility and altered antigenic profiles.

Both SARS-CoV and SARS-CoV-2 utilize trimeric spike (S) glycoproteins to recognize the same target receptor, human angiotensin-converting enzyme 2 (hACE2) for viral entry into host cells.[5, 6] Although both SARS-CoV S and SARS-CoV-2 S share the same target receptor, their sequence similarity is less than 80%, with their receptor-binding domains (RBD) sharing only

about 50% identity.[7] Another notable difference is at the furin protease cleavage site, where S is recognized and processed by cellular furin protease into receptor-binding subunit (S1) and fusion subunit (S2) prior to hACE2 binding. In SARS-CoV S, furin is ineffective in cleaving after the single arginine residue at position 667. In contrast, SARS-CoV-2 S possesses a polybasic RRAR motif at S1/S2, enhancing furin activity in efficient cleavage of S into S1 and S2 subunits, which has been linked to increased pathogenicity in this virus.[8-10]

A multitude of cryo-EM protein structures on soluble pre-fusion SARS-CoV S and SARS-CoV-2 S ectodomain trimer have been published with S in unbound, hACE2-bound and antibody-bound forms.[11-15] These studies have elucidated the architecture of S assemblies across different strains, revealing open and closed states as well as resolving the detailed contacts with antibodies and the hACE2 receptor. The successful binding of S in the open state to hACE2 is believed to prime the S cell entry machinery, leading the distal regions near the fusion peptide (FP) in S2 subunit (S2') more susceptible to cleavage by host protease TMPRSS2 or cathepsin. This cleavage releases the FP and triggers the refolding of the S2 subunit, leading to membrane fusion.[16-20] Once primed and triggered, similar to other Class I fusion proteins, S undergoes a dramatic structural refolding from pre-fusion to post-fusion conformation to complete the membrane fusion process.[21]

To go beyond static structures and better understand the mechanisms of S priming and activation, techniques that are capable and sensitive in characterizing solution-state protein dynamics are critical. Single-molecule Förster resonance energy transfer (smFRET) studies have utilized fluorescently labeled spikes, which can track the interconversion between RBD up and down conformations and monitor the influence of antibody or hACE2 binding on this conformational landscape.[22-24] However, structural information from smFRET is limited to the relative positioning and orientation of fluorescent probes introduced into the protein.

Complementary methods are greatly needed to comprehensively investigate dynamics throughout complex viral glycoprotein assemblies. Hydrogen/Deuterium-exchange mass spectrometry (HDX-MS) is a powerful technique for monitoring protein structural dynamics under native solution conditions. This bottom-up approach tracks deuterium uptake kinetics at the peptide segment level throughout the entire protein. Therefore, HDX-MS can be particularly informative for identifying conformational changes and allosteric effects induced by ligand binding.[25, 26]

In this chapter, I will describe the use of HDX-MS on tracking dynamic impacts of hACE2-induced S activation. The analysis will feature the comparison of structural dynamics and conformational changes between SARS-CoV S and SARS-CoV-2 S upon hACE2 binding. This comparison aims to provide insights into the distinct mechanisms of viral entry and phenotypic pathogenicity for these two coronaviruses.

3.2 Materials and Methods

Plasmid construction

The gene sequences encoding SARS-CoV S ectodomain (residues 14-1190) and SARS-CoV-2 S ectodomain (residues 14-1211) with di-proline mutations (S-2P: K986P and V987P) and a C-terminal T4 fibritin foldon with His-tag were cloned into pcDNA3.1(-) vector. For the SARS-CoV-2 S, the G614 variant was used with SGAG substitution at the S1/S2 site (RRAR), while the SARS-CoV S retained its original arginine at the S1/S2 site. Plasmids for soluble monomeric and dimeric hACE2 were obtained from Addgene (plasmid #149268 and #154101).

Transient transfection

Suspension Expi293F cell cultures were used for both S-2P and hACE2 protein expression. For each liter of cell culture at confluency (3×10^6 cells/mL), 1 mg plasmid DNA and 3 mL filtered polyethylenimine (PEI) solution at 1 mg/mL were pre-incubated in 25 mL Gibco Opti-MEM (Thermo Fisher Scientific) for 5 minutes before mixing. The mixture was incubated for 15 minutes and then added dropwise into the cell culture. The total transfection volume of each construct depends on the protein yield and experimental need. The transfected Expi293F cultures were incubated at 37 °C and 8% CO_2 with constantly shaking at 125 RPM.

Protein purification

On the fifth day post-transfection, transfected Expi293F cultures were harvested by centrifugation at 2000 RCF for 20 minutes. Supernatant was vacuum-filtered through Nalgene™ 0.45-μm αPES filters (Thermo Fisher Scientific) and supplemented with tris hydrochloride (Tris-HCl, pH 8.0), arginine hydrochloride (ArgCl, pH 6.5), sodium chloride (NaCl) and imidazole to a final concentration of 10 mM Tris-HCl, 50 mM ArgCl, 150 mM NaCl and 5 mM imidazole.[27] For each liter of supernatant, 5 mL of Ni-NTA resin was washed and equilibrated with wash buffer (50 mM Tris-HCl, 100 mM ArgCl, 150 mM NaCl and 5 mM imidazole, pH 8.0), followed by adding into the supernatant for a 2-hour batch binding at room temperature. The mixture was then flowed through a gravity column (Bio-Rad), followed by three times 15 mL (3 column volumes, CV) wash with wash buffer. The bound protein was eluted from the resin with three times 5 mL (1 CV) elution buffer (50 mM Tris-HCl, 100 mM ArgCl, 150 mM NaCl and 500 mM imidazole, pH 8.0). Elution fractions were SDS-PAGE checked, and then combined. Pooled elution was buffer-exchanged with size-exclusion chromatography (SEC) running buffers (50 mM Tris-HCl,

100 mM ArgCl, 150 mM NaCl and 0.02% sodium azide (NaN_3), pH 8.0) to reduce the imidazole level below 1 mM and concentrated to ~1 mg/mL spike or 4 mg/mL hACE2 protein concentration with an Amicon Ultra-15-mL-centrifugal filter (30 K & 10K, Millipore). Concentrated spike and hACE2 samples were flash frozen in liquid nitrogen and stored in -80 °C freezer.

All the protein samples were finally SEC purified before major experiments. 500 μL sample was thawed and spun at 15000 RCF for at least 10 minutes to remove possible protein aggregation pellet before injecting to Superose 6 10/300 GL column (Cytiva) on an AKTA Pure system (GE). The SEC purified spike and hACE2 proteins were concentrated to desired concentrations according to the experiment need.

Dynamic light scattering

Dynamic light scattering (DLS) measurements were performed on a DynaPro NanoStar II (Wyatt) to characterize spike trimer integrity, homogeneity. SEC purified spike samples at 1 mg/mL were first spun down at 15000 RCF for 20 minutes to remove aggregation pellet. 5 μL of the sample was then injected into a 2-μL quartz cuvette (Wyatt). Each DLS run was measured with twenty 10-second acquisitions at 25 °C by LASER auto-attenuation.

Hydrogen/Deuterium-exchange mass spectrometry

For the experiments comparing hACE2 impacts on SARS-CoV and SARS-CoV-2 spike dynamics, Each of 10 μg SARS-CoV S, 10 μg SARS-CoV-2 S and 3 μg trimeric tethered SARS-CoV-2 RBD (tRBD) was incubated with hACE2 with the molar ratio 1:9 (trimer:hACE2) to a final volume 10 μL, overnight at 23 °C, respectively. Both hACE2 bound and unbound samples were deuterium exchanged in 90 μL deuteration buffer (10 mM Tris, pH^* 8.0, 85% D_2O, Cambridge

Isotope Laboratories, Inc.) at 23 °C for 3, 30, 180, and 900 seconds. Exchanged sample was immediately mixed with 100 μL ice-cold quench buffer containing 8 M urea, 200 mM tris(2-chloroethyl) phosphate (TCEP) and 0.2% formic acid (FA) to pH 2.5, followed by flash frozen in liquid nitrogen.

All samples were thawed on ice for 5 minutes, loaded to the LC-MS system on a sample injection platform and analyzed by a Synapt G2 mass spectrometer (Waters) with the settings described in our previous study. Samples were in-line digested into peptides by an immobilized pepsin column (2.1 × 50 mm) and loaded onto CSH C18 trap cartridge (Waters) with loading buffer (2% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA)) at 200 μL/min. Peptides were then separated by a UPLC CSH C18 column (1.0 × 100 mm, 1.7 μm, Waters) with a 20-min linear gradient from 3% to 40% buffer B (buffer A: 2% ACN, 0.1% FA, 0.025% TFA; buffer B: 99.9% ACN, 0.1% FA) at a 40 μL/min flow rate. A series of wash steps with various of solvents were performed to remove carryover and regenerate pepsin column for next run.[28] A fully deuteration control was prepared by collecting digested samples from pepsin column, drying by speed-vaccum, resuspending in 10 μL SEC buffer and incubating in 90 μL deuteration buffer for 1 hour at 85 °C.

Peptides identification and mapping were performed by running digested samples on an Orbitrap Fusion™ mass spectrometer (Thermo Fisher Scientific) with the same settings mentioned in method glycan analysis. Peptide mass spectra were further confirmed on DriftScope (Waters) and identified with specific retention time and drift time. HDExaminer (Sierra Analytics) was used to process the mass spectrometer datasets and calculate deuterium uptake of each peptide identified by the specific retention time and drift time. In the cases of bimodal spectra, the mass envelop was resolved by HX-Express v2 applying bimodal deconvolution and binomial fitting.[29]

SDS-PAGE and western blot

For each SDS-PAGE gel using Coomassie Blue staining, 12 μL spike sample was mixed with 4 μL 4× NuPAGE™ LDS loading dye (Invitrogen) before loading to 4-12% NuPAGE™ Bis-Tris gel (Invitrogen) well. SDS-PAGE gels were run at 120 Volts for 10 minutes and 150 Volts for 60 minutes in MES SDS running buffer (Invitrogen), followed by gel collection and PageBlue™ (Thermo Scientific) staining.

Western blot was used to characterize TMPRSS2 digested S. SARS-CoV S and SARS-CoV-2 S trimers were incubated with hACE2 or negative control Bovine serum albumin (BSA) in the same ratio. Soluble TMPRSS2 (Creative Biomart Inc.) was added into the solution in 1:50, 1:10, and 1:1 ratio (w/w to original S amount), respectively, and incubated for 10 minutes at room temperature. 5 μL digested samples were mixed with 1 μL DTT (0.5M) and 2 μL 4× loading dye, loaded to the SDS-PAGE gel wells, and run at 120 Volts for 80 minutes. The peptides on the gel were then transferred to the Immobilon®-FL PVDF membrane (Millipore) using 10% methanol supplemented NuPAGE™ transfer buffer (Invitrogen) at 30 Volts for 2 hours. The membrane was blocked in 5% non-fat milk overnight, followed by incubation in primary antibody, respectively, in room temperature for 1 hour. Primary antibody against spike RBD (rabbit, pAb, Sino Biological 40592-T62) or against spike S2 subunit (rabbit, pAb, Sino Biological 40590-T62) was used in 1:2000 tris-buffered saline-Tween (TBST) solution. Secondary antibody anti-rabbit IgG Alexa Fluor™ 680 (Invitrogen) was used in 1:10000 5% non-fat milk blocking buffer in room temperature for 1 hour.

Both SDS-PAGE gels and membranes were imaged on the Odyssey® M-XS imaging system (LI-COR Biosciences) using 700 nm channel and optimized exposure intensity.

3.3 Results

Strain-specific differences in trimer structural dynamics

Before investigating the dynamic impacts on the spike proteins in the hACE2-bound state, we first aimed to determine whether similar spike architecture and dynamic profiles could be observed in two distant coronavirus S assemblies with less than 80% sequence homology. Among all the reported SARS-CoV-2 variants of concern (VOCs), we opted to feature the S variant closest to the initial significant spillover event in order to avoid the further evolutionary adaptations that could alter hACE2 interactions and evolve in response to immune selective pressure, highly likely with reshaped structural orders on S. Therefore, the SARS-CoV-2 S-2P with glycine at residue 614 (G614) was used to represent soluble SARS-CoV-2 S in this study, based on its stability and homogeneity under experimental conditions (reported in the previous chapter) as well as its biological significance, aligning with our aims. This S-2P construct also replaced the native polybasic RRAR S1/S2 cleavage site with an SGAG sequence to maintain the S-2P in native pre-fusion conformations.[23] In comparison, the SARS-CoV S-2P form also incorporated the di-proline mutations at the same position as in SARS-CoV-2 S-2P, but retained the native single arginine (R) at S1/S2 site, which remained uncleaved under the expression and experimental conditions.[9]

HDX-MS experiments were performed on SARS-CoV and SARS-CoV-2 S-2P with incubation times ranging from 3 seconds to 20 hours. Peptide identification was determined by the tandem MS/MS analysis using a high-resolution Fusion Orbitrap mass spectrometer with the electron-transfer/higher-energy collision dissociation (EThcD) method along with reference to previously reported glycan profiling.[30-32] Due to differences in pepsin digestion patterns caused by sequence variations, the SARS-CoV-2 S-2P construct showed 85% sequence coverage with 11 N-

glycosylation sites covered, while SARS-CoV S-2P construct showed 70% coverage with 9 N-glycosylation sites covered (Figure 1a, Figure S2).

HDX-MS analysis revealed the dynamic profiles of the pre-fusion SARS-CoV S trimer, depicted in deuterium uptake heatmaps (Figure S1) and butterfly plots showing deuteration trends across the spike sequences (Figure 1a). Due to different enzymatic digestion patterns, the number of homologous peptides with the same number of exchangeable amides that could be quantitatively compared was limited.[3, 33] This limitation restricted the direct comparison of dynamics across the entire spike proteins.

Therefore, in addition to including uptake heatmaps that illustrate the relative spatial dynamic profile of the entire S, we highlighted some homologous peptides in structurally important regions. From these quantitative and direct comparisons, measurable domain-level differences in dynamics could be revealed (Figure 1b-d). Homologous peptides spanning a β-strand in the NTD of SARS-CoV S had similar uptake levels in the early-to-middle exchange times but showed greater dynamics than SARS-CoV-2 S at the 20-hour time point (Figure 1b). Meanwhile, homologous peptides in the RBD featuring the joint to the receptor-binding motif (RBM) and the C-terminal region, were more dynamic in SARS-CoV-2 S (Figure 1c). Based on the findings from the previous chapter, these differences likely resulted from SARS-CoV-2 S adopting more open state conformations than SARS-CoV S, affecting peptide accessibility due to RBD up/down positioning observed in reported structures.[12, 34] It could also be possible that the orientations and exposure of the RBDs differ slightly between the two constructs, even when in the same conformational states. The observation of higher dynamics at the apex of HR1-CH helical bundle in SARS-CoV-2 S (Figure 1d, peptide #5) supported the prior reason which more open states appeared in SARS-CoV-2 S. Additionally, SARS-CoV-2 S showed a slightly more protected

S2 bottom region (Figure 1d, peptide #6 and #7). This stabilized S2 subunit was expected to energetically favor the more open conformations.

Highly dynamic fusion peptide regions (Figure 1d, peptide #4) were observed in both constructs. Our findings support the cryo-EM-based observation that SARS-CoV-2 S adopts more open conformations than SARS-CoV S, while maintaining a relatively conserved architecture at a structural dynamic level.[34-37] The domain-specific dynamics and allosteric effects in response to RBD conformational switches were also conserved, as concluded in the previous chapter of the SARS-CoV-2 S study, despite the significant sequence differences. This highlights the similarities in the structural dynamics of these two viral spike proteins in their unbound states and establishes a common basis for studying the impacts of hACE2 binding.

hACE2-induced activation of SARS-CoV-2 and SARS-CoV spike assemblies

We have observed and reported that the local dynamics in S2 subunit had an allosteric effect in both SARS-CoV and SARS-CoV-2 S. Next, we sought to investigate how the hACE2-RBD interaction activates S, specifically focusing on how this interaction impacts the dynamics of the fusion subunit and fusion peptide, contributing to viral fusion. To mitigate avidity differences arising from variations in binding affinity and the proportion of open conformations, we incubated excess monomeric hACE2 with the SARS-CoV S trimer, SARS-CoV-2 S trimer and trimeric tethered SARS-CoV-2 RBD (served as an all-RBD-up positive control). The molar ratio of 9:1 (3-fold molar excess per protomer) was applied, which could achieve a > 99% hACE2 bound fraction for each protein based on their K_D values (K_D of various S constructs: 1-10 nM; S concentration in HDX buffer: ~200 nM; hACE2 concentration in HDX buffer: ~1800 nM).[14, 38]

Trimeric SARS-CoV-2 RBD construct tethered three RBD domains by a flexible linker to a designed trimerization domain. This served as a simplified mimic of spike with all RBDs in the up conformation, exposing all three receptor-binding motifs (RBMs) to hACE2.[32] The down conformation of the RBD was not favored due to the lack of stabilized packing from other spike domains, such as the NTD. Upon hACE2 binding, HDX-MS revealed local protection across the RBM (Figure 2b) with effects further extending down to the RBD β-strand core (Figure S3). The degree of deuterium uptake difference was significant during early-to-middle exchange times but nearly disappear after the 30-minute exchange. Conversely, sites facing inward towards the trimer center and distal to the RBM (C-terminal) did not show changes in peptide exposure levels upon hACE2 binding (Figure 2b, Figure S3), as anticipated from the all-RBD-up conformations in the trimeric RBD construct.

Moving to the SARS-CoV-2 S trimer construct, the RBD and HR1-CH regions exhibited the most dramatic changes in dynamics upon hACE2 binding (Figure 2a). The RBD responded to hACE2 binding with two distinct modes. Local protection was observed at the RBM and peptides leading to the β-strand core, similar to what was seen in the isolated trimeric RBD (Figure 2c, peptide #1 and #2, Figure S4). The pattern of deuterium uptake differences over time was well-aligned with the trimeric RBD. However, greater exposure of peptides spatially close to the RBD C-terminal pivot was observed in hACE2-bound S trimers (Figure 2c, peptide #3), indicating a consistent conformational bias towards the RBD-up conformation, aligning with our designed > 99% hACE2-bound scenario. Compared to the hACE2-bound isolated trimeric RBD, these peptides in spike RBD showed reduced deuteration due to interactions with the rest of the spike structure (Figure 2c, peptide #3, Figure S3).

Enhanced dynamic behavior was observed in broad regions of the spike, including allosteric changes at sites distant from the RBD, in respond to hACE2-induced conformational changes (Figure 2c, Figure S4). Peptides in the NTD exhibited only modest changes (Figure S4), while the S1 subunit hinge region connecting the RBD C-terminal pivot showed a moderate increase in peptide dynamics (Figure 2c, peptide #5).

The most prominent dynamic changes occurred at the apex of the central helices, from HR1 to the CH domain (Figure 2c, peptide #4 and #6). Specifically, a bimodal mass envelope was observed in peptide #4, which spanned the entire di-proline mutated helical apex, indicating the presence of two co-existing populations exhibiting differences in dynamics at the earliest time point (Figure 2d). This suggested a local conformational switching from the ordered helical structure, in addition to the increased exposure induced by hACE2 binding. The destabilization in the HR1 helix exhibited the most prominent differences, with this region becoming nearly fully exchanged after 30 seconds in the hACE2-bound state, compared to 3 minutes in the unbound state (Figure 2c, peptide #6, Figure S4). These dramatic dynamic behaviors at the central helical apex indicated significant changes in the local structural flexibility of the fusion machinery. These findings are consistent with the distorted helical structure observed in cryo-EM studies upon hACE2 binding.[39]

Increased dynamics resulting from hACE2 binding were observed along the HR1 domain down the helix and identified in the fusion peptide proximal region (FPPR), adjacent to the TMPRSS2 cleavage site (Figure 2c, peptide #6, Figure S4). We speculated that the hACE2-induced dynamic increases in the SARS-CoV-2 S at the FPPR site likely led to a primed conformation with enhanced accessibility for TMPRSS2, which supposed to subsequently cleave the S2' site to release fusion peptide and activate the spike protein's fusion capability.

In comparison, we investigated hACE2's impact on SARS-CoV S. Similar to SARS-CoV-2 S, the RBD and HR1-CH regions exhibited the most significant dynamic changes upon hACE2 binding (Figure 3a). However, the number of residues exchanged, and the degree of the dynamic changes were lower than those observed in the SARS-CoV-2 S.

Local dynamic changes in the RBD of SARS-CoV S upon hACE2 binding showed a similar pattern to those in SARS-CoV-2 S. Direct protection in the RBM (Figure 3b, peptide #1) and increased exposure near the RBD C-terminal region (Figure 3b, peptide #2) were observed. Under the > 99% hACE2-bound conditions, the RBDs were induced to adopt primarily up conformations, as reflected by the peptide (Figure 3b, peptide #2), homologous to residues 388-392 in SARS-CoV-2 S RBD, exhibiting increased deuterium exchange. However, the increase was not comparably significant.

Additionally, a bimodal mass envelope was observed and resolved at the binding interface (Figure 3c, corresponding peptide #1 in Figure 3b). In the unbound state, the deuterium uptake plateaued at 50% from the earliest time point, indicating that half of the exchangeable peptide backbone located at the binding interface was highly exposed while the other half, buried in the RBD β-strand core, was barely deuterated at all time points. In the hACE2-bound state, a significant reduction in local dynamics was observed from the earliest time point. However, this local protection did not last long, as the uptake levels overlapped with the unbound state after 3 minutes (Figure 3b, peptide #1). Compared to the homologous peptide in SARS-CoV-2 S, where the uptake curve in hACE2-bound state met the unbound curve at the 15-minute time point, the short-lived protection at SARS-CoV S-hACE2 interface suggested difference in binding dynamics (Figure S6). This difference is likely associated with the nearly 5-fold faster k_{off} rate for the SARS-CoV S-hACE2 complex (SARS-CoV: 7.1×10^{-4} $M^{-1}s^{-1}$, SARS-CoV-2: 1.6×10^{-4} $M^{-1}s^{-1}$).[14]

Furthermore, the slow-uptake population in the bimodal mass envelope at the 3-second time point resembled the mass envelope in SARS-CoV-2 (Figure S6). In contrast, the fast-uptake population plateaued from the earliest time point, with the same peak m/z values as in the later time points and unbound state (Figure 3c, Figure S6). Both the shorter-lived protection and this exposed population from the bimodal behavior indicated the differences in how SARS-CoV and SARS-CoV-2 S interact with and utilize hACE2 receptor.

Allosteric changes with increased dynamics were also observed at the S1 hinge region connecting the C-terminal of the RBD (Figure 3b, peptide #3), the apex of the HR1-CH domains (Figure 3b, peptide #4), and down the HR1 helix (Figure 3b, peptide #5 and #6) in respond to the hACE2-induced up-conformations of SARS-CoV S RBD (Figure S5). The differences in dynamics were moderate, and no hints of bimodal behavior was observed, in comparison with the homologous peptides in SARS-CoV-2 construct.

Binding of hACE2 resulted in notable allosteric effects extending from the RBD at the S apex down to the central helices in the S2 subunit in both SARS-CoV S and SARS-CoV-2 S (Figure 2c, Figure 3b). These dynamic changes in SARS-CoV-2 S were more widespread, particularly in the ordering of the RBD, through more stable binding to hACE2, and along the S2 subunit's central helical bundles (HR1-CH) at early exchange time points. In SARS-CoV S, the effects were more muted, possibly due to the greater likelihood of hACE2 dissociation during the interaction, reducing the propensity for the S2 subunit to become activated.

hACE2-induced activation impacts on TMPRSS2 processing

A complementary activity assay was conducted to examine the impact of hACE2 binding on subsequent TMPRSS2 accessibility to the spike protein in solution (Figure 4a). Probing with a

primary antibody against conserved RBD epitopes, western blots revealed undigested S bands above 160 kDa and a digested product band at ~90 kDa when TMPRSS2 was added at a 1:1 ratio to S for both SARS-CoV and SARS-CoV-2 S. TMPRSS2 cleavage near the S2' position produced a 90 kDa peptide spanning from the C-terminal of RBD to the S2' site, confirmed through non-reducing conditions and primary antibody blotting of the N-terminal of S2 prior to the S2' site (Figure 4b), consistent with TMPRSS2's known biological activity and previous studies.[40, 41] Even without receptor binding, a faint digest product band appeared for both SARS-CoV S and SARS-CoV-2 S, with a greater extent shown for SARS-CoV-2 S. In both cases, hACE2 binding enhanced TMPRSS2 digestion (Figure 4a), consistent with allosteric structural dynamic changes detected in SARS-CoV-2 S from HDX-MS. These results support our HDX-MS findings in complementary, which hACE2 binding induces allosteric conformational changes extending to the S2' cleavage site, priming the spike protein for TMPRSS2 cleavage activation.

3.4 Discussion

Class I fusion proteins, including the S glycoprotein from *Sarbecovirus* viruses, exhibit a highly dynamic nature, driving significant conformational changes from pre-fusion to post-fusion forms during the viral entry.[42, 43] In the previous chapter, we discussed the dynamic impacts of early SARS-CoV-2 S variants and concluded that allosteric effects extend from the distal part of S2 subunit to the RBD conformational switches. In this chapter, we studied and compared the dynamic behaviors of SARS-CoV S, observing similar allosteric effects despite the different protein sequences from this genetically distant strain. While viral evolution directs the ordering of the S2 subunit to support S open states and facilitate hACE2 recognition, it is logical for us to

hypothesize that hACE2 engagement should also broadly impact the S trimer, in a way extending down to the S2 subunit.

Meanwhile, many HDX-MS studies have investigated the protein dynamics impacted by hACE2, with various focuses. For instance, an HDX-MS and molecular dynamic (MD) simulation study on the ancestral SARS-CoV-2 S D614 reported an increase in local dynamics at the S1/S2 cleavage site upon hACE2 binding.[44] Interestingly, in our experiment using SARS-CoV-2 S G614, the same loop featuring the S1/S2 cleavage site exhibited high flexibility both with and without hACE2 binding. This is probably due to such evolution advantageous D614G mutation, positioned at S1/S2 interface, directly affected the local ordering. This underscored the significance of studying this matured and prevalent variant in our current experiment.

More extreme forms of distorted, splayed open trimers promoted by hACE2 binding have been reported.[45] However, we did not observe hints of exhibiting splayed open trimers in our study. Instead, we detected peptides with fast exchange rates and bimodal behaviors in several regions across the HR1 and CH domains, indicating transient conformational samplings of two structural orderings in these dynamic regions (Figure S4). Despite these behaviors, the core of the central helical bundle remained consistently protected (Figure S4).

Moreover, we found a hACE2-induced allosteric effect that destabilized the fusion peptide proximal region of SARS-CoV-2 S (Figure 2c, Figure S4). This is consistent with recent HDX-MS studies showing hACE2-induced effects in the FPPR,[46] which we further investigated using TMPRSS2 activity assays.

In addition to the complementary biochemical assay to test the dynamic behavior of S during TMPRSS2 digestion, we also tested our hypothesis in another native scenario. We compared the dynamic effects of natively dimeric hACE2 binding on S. This comparison aimed

to investigate whether the potential for bivalent engagement could enhance spike's dynamic activation, a concept suggested in other reports, but not previously investigated using experimental approaches.[47]

When displayed on the cell surface, hACE2 exists as a dimer, and it has been suggested that S which possesses three RBDs may exhibit avid interactions with native, dimeric hACE2 due to the receptor's bivalency.[48] To explore this, we investigated whether dimeric hACE2, which was able to interact with more than one RBD, could potentially lead to greater conformational priming of the spike compared to monomeric hACE2 binding in the saturated ligand conditions. Parallel HDX-MS experiments using dimeric and monomeric hACE2-bound SARS-CoV-2 S at pH 7.4 showed nearly identical changes in structural dynamics for all peptides previously exhibiting significant differences in respond to hACE2 binding (Figure S7). It is also possible that the reported effects of hACE2 bivalency may be more commonly involved in inter-spike crosslinking rather than bivalent intra-spike RBD engagement.[22] Overall, the introduction of bivalent hACE2 did not exhibit significantly novel impacts on S dynamics, suggesting the potential of bivalent binding is not a necessity of the S activation.

Using rigorous experimental conditions and complementary assays, our study extended the dynamic investigations of the *Sarbecovirus* spikes from SARS-CoV-2 to the distant SARS-CoV strain, in examining the S dynamics and receptor-induced activation.

The hACE2 activation pathway we examined is crucial for understanding S priming for subsequent cleavage at the S2' position, by either TMPRSS2 or cathepsin proteases, necessary for releasing the fusion peptide and initiating the conformational changes for membrane fusion. One possible reason for the increased dynamics in the FPPR only observed in SARS-CoV-2 S, but not in SARS-CoV S, might result from the differences in their interactions with hACE2. Biolayer

interferometry (BLI)[14] data suggested that hACE2 dissociates more slowly from SARS-CoV-2 S than from SARS-CoV S, which was also observed by our HDX-MS experiment indicating a faster dissociation rate for SARS-CoV S. The tighter and continuous binding of hACE2 to SARS-CoV-2 S may be necessary to facilitate greater dynamic induction and increased FPPR exposure, enhancing TMPRSS2 access to the S2' cleavage site and subsequent fusion triggering. In contrast, due to the fast dissociation of hACE2, SARS-CoV S likely requires greater saturation of binding sites and possibly higher S avidity on the virion surface with membrane-bound hACE2 to compensate and achieve similar priming effects.

These results reveal critical allosteric pathways and variant-specific differences in dynamic activation among *Sarbecovirus* spike assemblies. This understanding complements high-resolution structural information and provides mechanistic insight into differences in viral infectivity and function during cell entry.

Figures

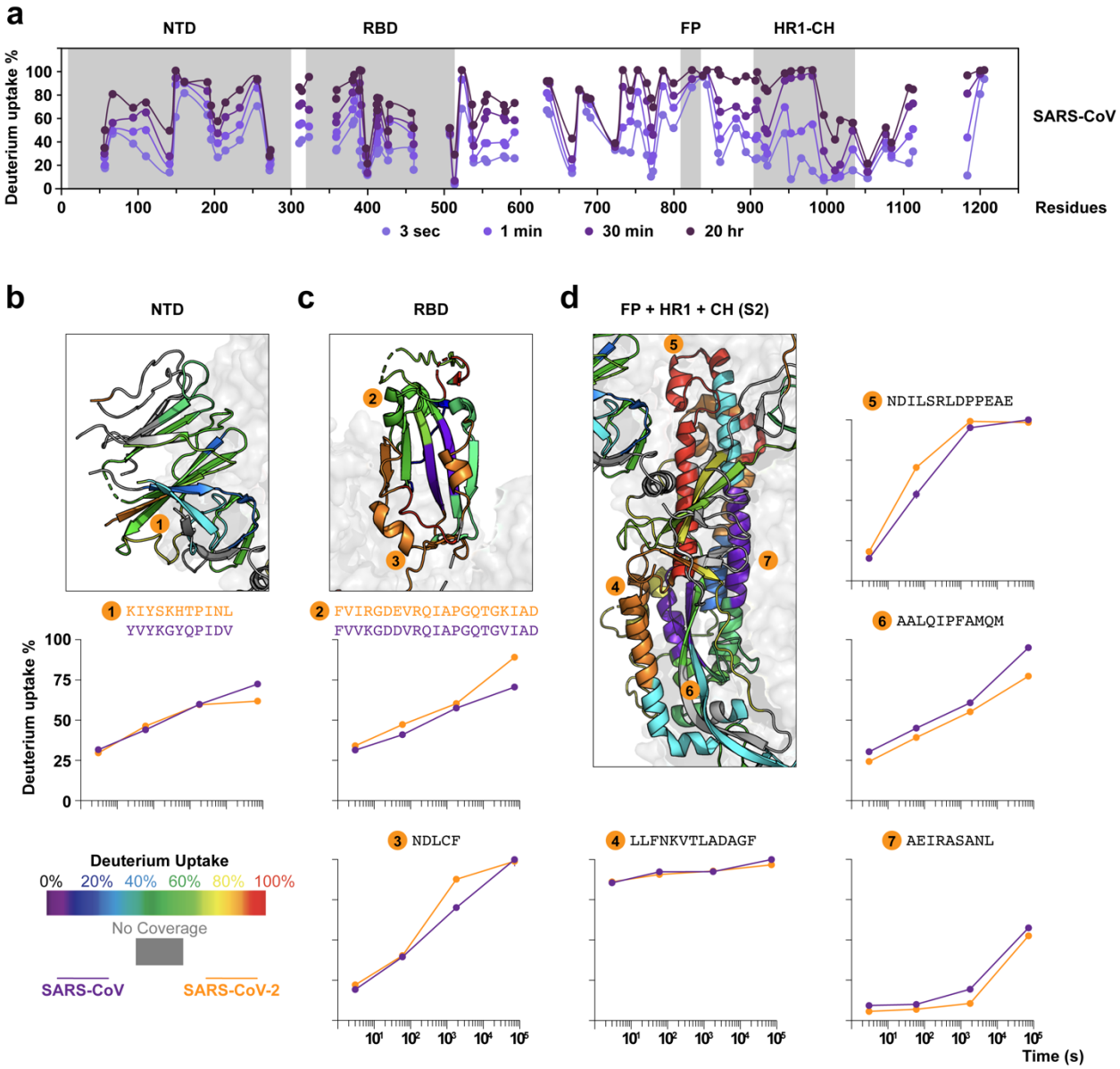

Figure 1. **Dynamic differences in SARS-CoV and SARS-CoV-2 S trimers. (a)** Butterfly plots of SARS-CoV S. Grey shades highlight the RBD, fusion peptide and HR1-CH region. Detailed comparisons for the **(b)** NTD, **(c)** RBD, and **(d)** S2 subunit are displayed between selected homologous peptide sequences from SARS-CoV S and SARS-CoV-2 S.

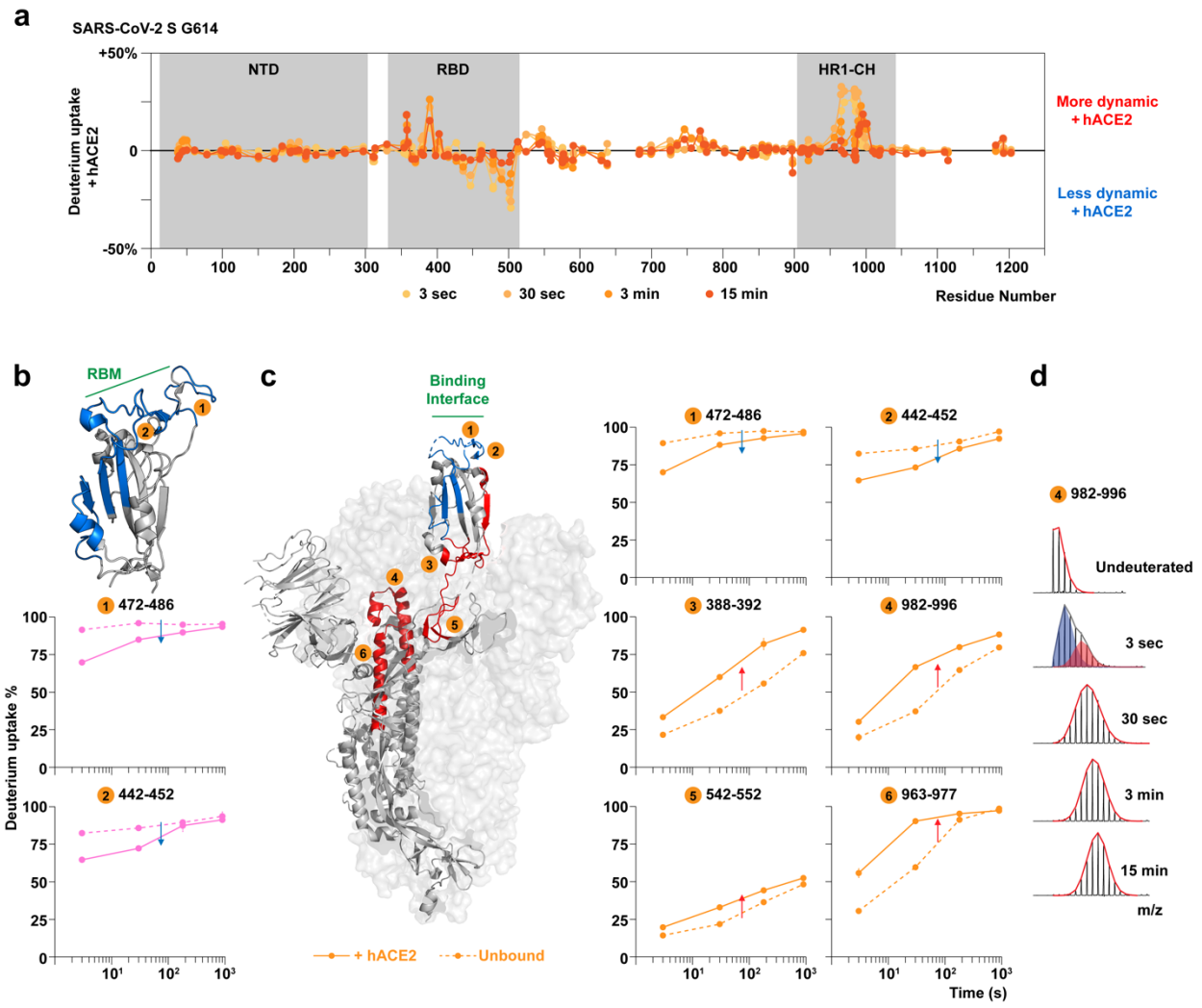

Figure 2. **Local and allosteric effects on SARS-CoV-2 trimeric tethered RBD and trimeric S upon hACE2 binding. (a)** Differential plots of SARS-CoV-2 S indicate footprints and allosteric effects resulting from hACE2 binding. Gray shades highlight the RBD, fusion peptide, and HR1-CH region. **(b)** Only local protection has been observed on the trimeric tethered RBD construct (PDB: 6W41). Blue indicates more protection and red indicates more exposure after hACE2 binding. **(c)** On hACE2-bound S trimer, the corresponding RBD shows protection throughout the RBM (peptide #1 and #2) but more exposure at the C terminal (#3); the hinge region shows increased dynamics (#5), and the central helices show increased accessibility from the fusion peptide proximal region (#6) to the helical apex (#4). **(d)** A region in the central helices' apex (#4

in panel c) shows a bimodal mass envelope at the 3-sec deuteration time point. The bimodal spectrum is binomially fitted to two populations: a protected population is shaded blue, and the heavier in mass population that has sampled an exposed conformation is shaded red.

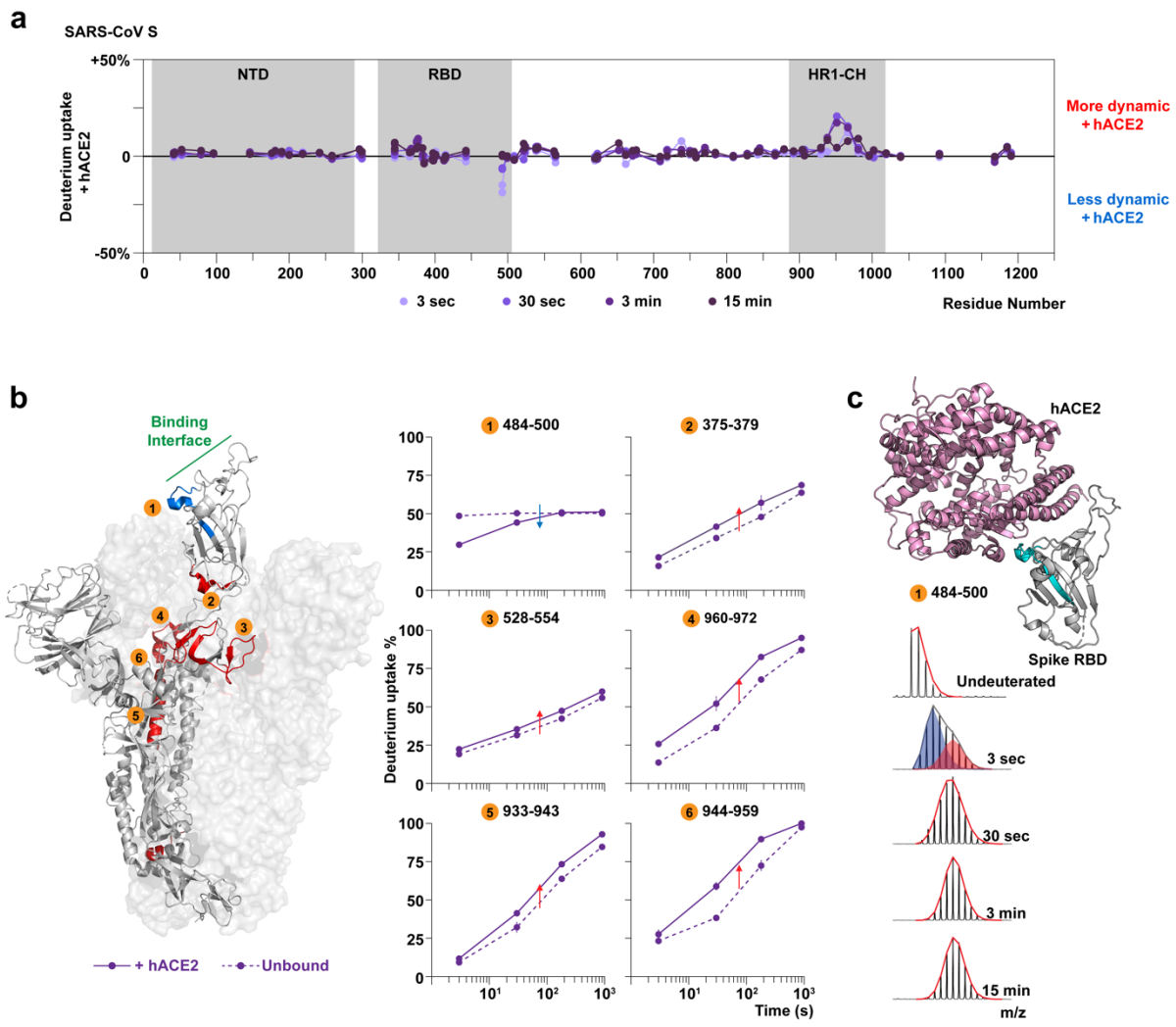

Figure 3. **Local and allosteric effects on SARS-CoV S upon hACE2 binding. (a)** Differential plots of SARS-CoV indicate footprints and allosteric effects resulting from hACE2 binding. Gray shades highlight the RBD, fusion peptide, and HR1-CH region. **(b)** Upon hACE2 binding, the RBD shows protection across the binding interface (peptide #1) and increased exposure at the C terminus (#2); the hinge region shows increased dynamics (#3), and the central helices reveal increased amide accessibility (#4-6). Blue indicates more protection and red indicates more exposure after hACE2 binding (PDB: 6CRZ). **(c)** A region at the binding interface exhibits a bimodal mass envelope at the 3-sec time point (PDB: 3D0G). The bimodal spectrum is binomially

fitted to two populations: a protected population is shaded blue, and the heavier in mass population that has sampled an exposed conformation is shaded red.

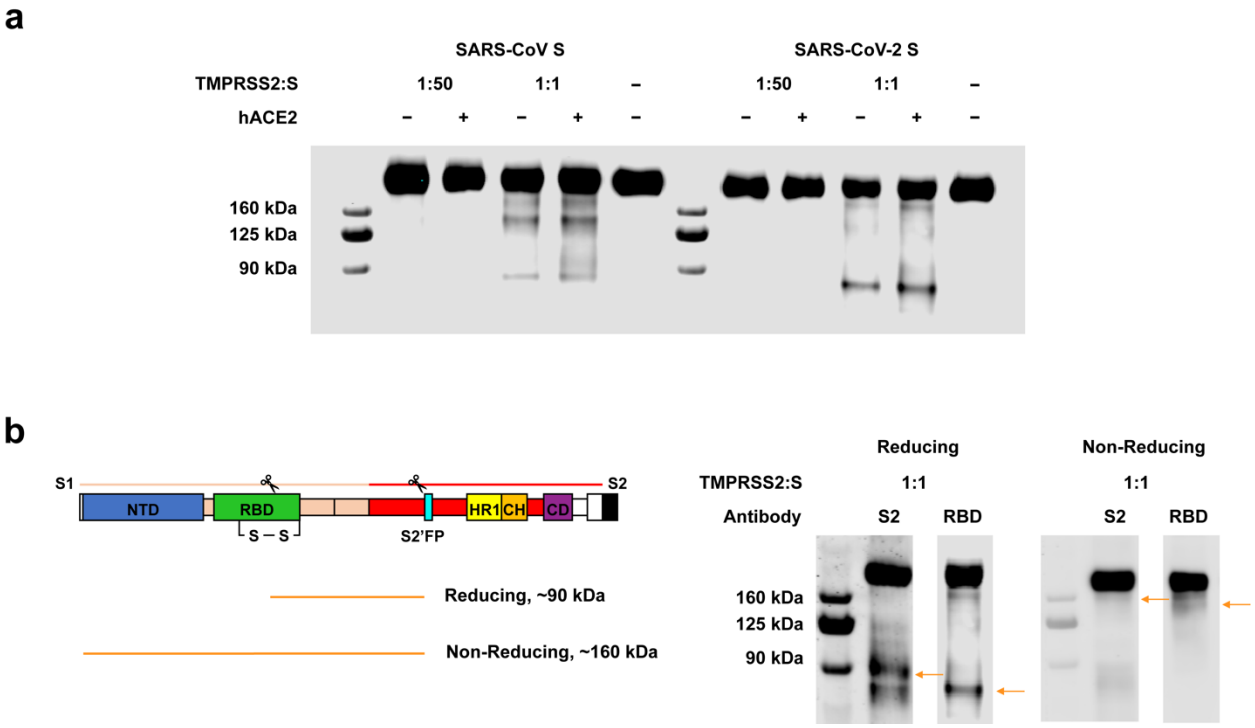

Figure 4. **hACE2-bound SARS-CoV S and SARS-CoV-2 S impact on TMPRSS2 digestion.** **(a)** Western blot indicating TMPRSS2 digests more efficiently on SARS-CoV-2 S. The occurrence of stable digested product at ~90 kDa is enhanced with hACE2 binding. **(b)** Suggested TMPRSS2 cleavage sites and possible digested products in reducing and non-reducing conditions. Western blot on TMPRSS2 digested S using primary anti-S2 pAbs and primary anti-RBD pAbs. Orange arrows indicate the target band shifts due to the reduction of the disulfide bonds.

Supplemental Information

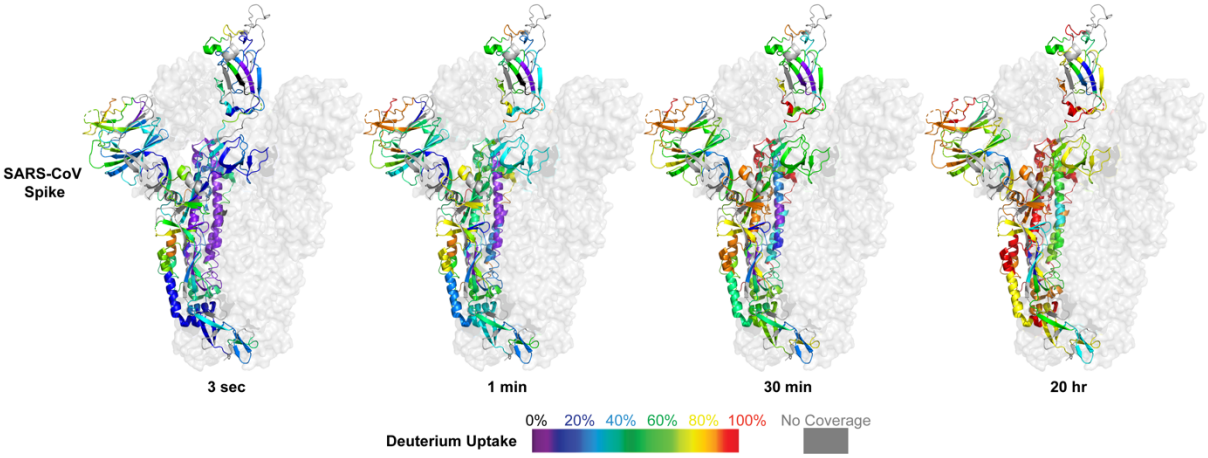

Figure S1. **HDX heatmap of SARS-CoV S-2P.** Color-coded SARS-CoV S deuterium uptake heatmaps are shown in peptide level resolution at 3 sec, 1 min, 30 min, and 20 hr labeling time points.

Figure S2. **HDX coverage map of SARS-CoV S-2P.** Peptide coverage map is from pepsin digestion. Blue shade represents RBD sequence; Orange shade represents engineered mutations at S1/S2 and HR1-CH apex sites; Pink shade represents fusion peptide.

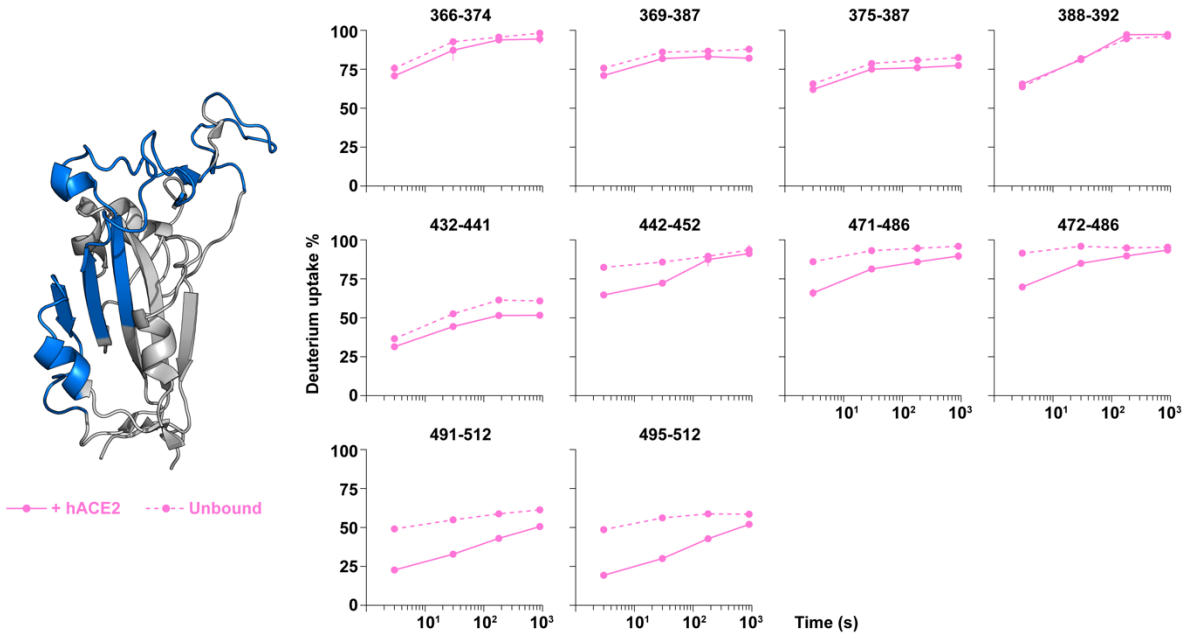

Figure S3. **Deuterium uptake plots show dynamic impacts on trimeric RBD upon hACE2 binding.** This figure is related to Figure 2b.

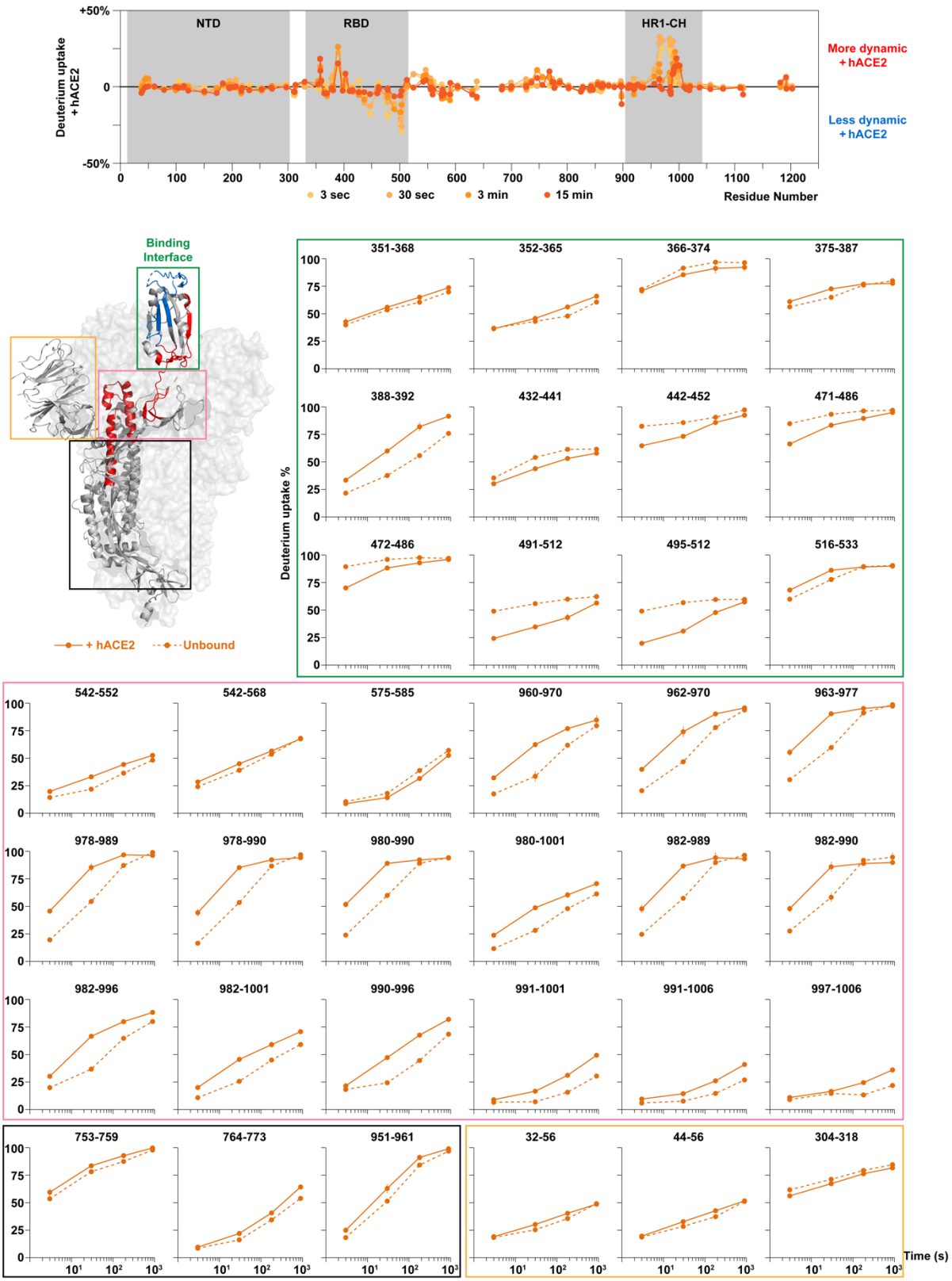
+50%
0
-50%
Deuterium uptake
+ hACE2
NTD
RBD
HR1-CH
More dynamic
+ hACE2
Less dynamic
+ hACE2
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
Residue Number
3 sec
30 sec
3 min
15 min
Binding
Interface
+ hACE2
Unbound
Deuterium uptake %
351-368
352-365
366-374
375-387
388-392
432-441
442-452
471-486
472-486
491-512
495-512
516-533
542-552
542-568
575-585
960-970
962-970
963-977
978-989
978-990
980-990
980-1001
982-989
982-990
982-996
982-1001
990-996
991-1001
991-1006
997-1006
753-759
764-773
951-961
32-56
44-56
304-318
Time (s)

Figure S4. **Deuterium uptake plots show dynamic impacts on SARS-CoV-2 S-2P upon hACE2 binding.** This figure is related to Figure 2c. Green box: RBD; Pink box: S2 central helical apex and hinge region; Black box: bottom S2; Orange box: NTD.

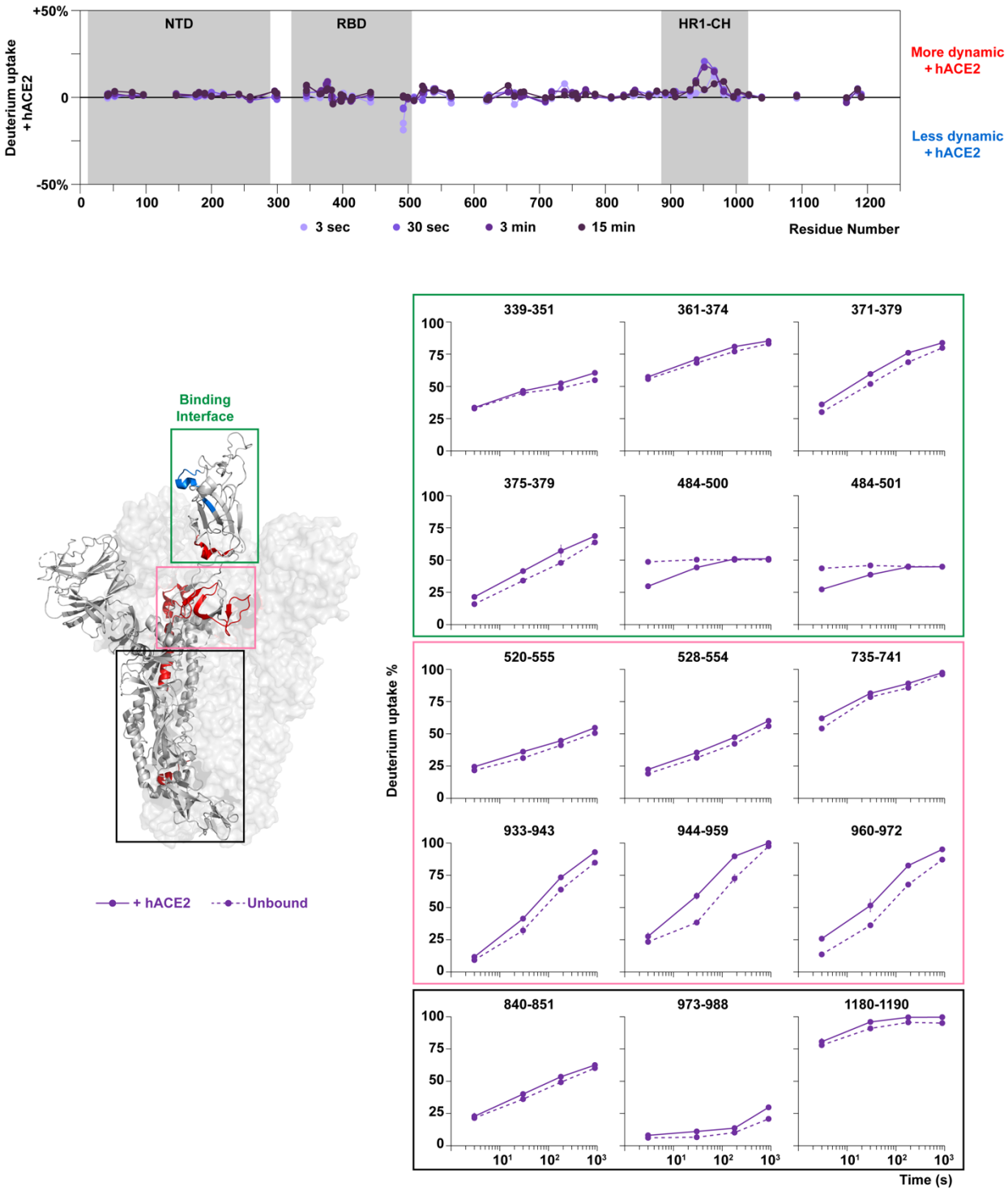

Figure S5. **Deuterium uptake plots show dynamic impacts on SARS-CoV S-2P upon hACE2 binding.** This figure is related to Figure 3b. Green box: RBD; Pink box: S2 central helical apex and hinge region; Black box: bottom S2.

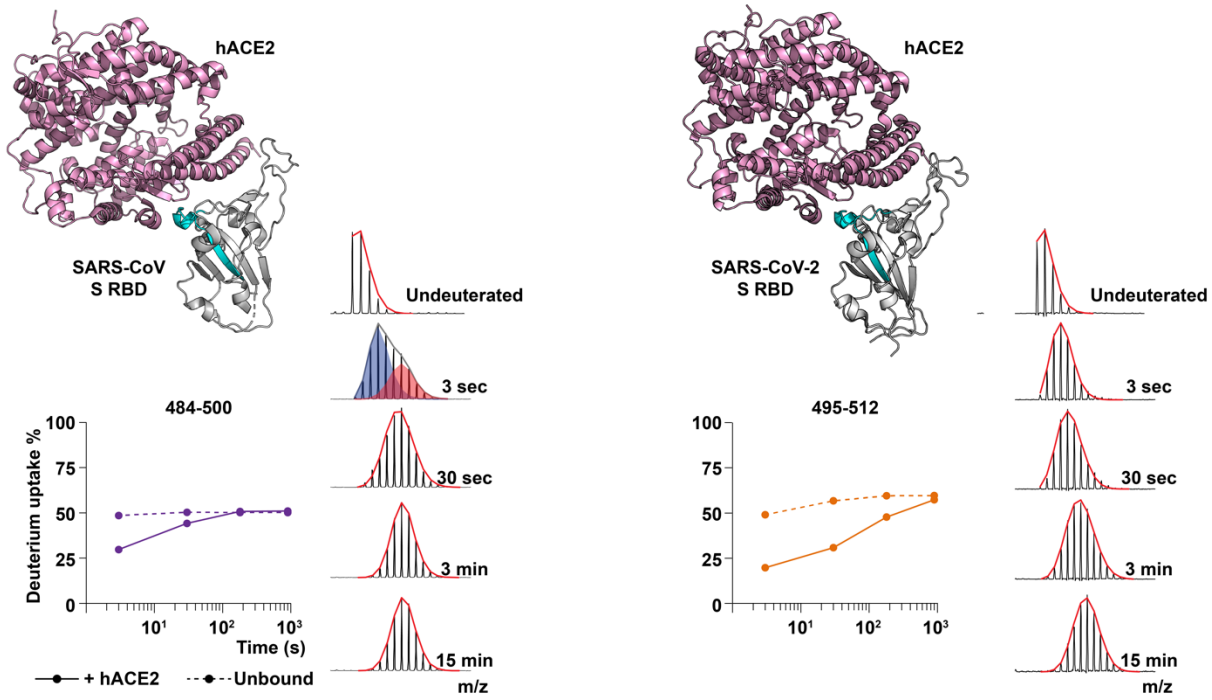

Figure S6. **Different dynamic behaviors over hACE2-interacting peptides between SARS-CoV and SARS-CoV-2 S.** Peptides connecting hACE2-interacting loop and RBD core region of SARS-CoV S (left panel) and SARS-CoV-2 S (right panel) are compared. In SARS-CoV S (PDB: 3D0G), hACE2-bound uptake curve merges unbound curve from 3-min time point with bimodal behavior at 3-sec time point. The bimodal spectrum is binomially fitted to two populations: a protected population is shaded blue, and the heavier in mass population that has sampled an exposed conformation is shaded red. In SARS-CoV-2 S (PDB: 6LZG), hACE2-bound uptake curve overlaps unbound curve at 15-min time point.

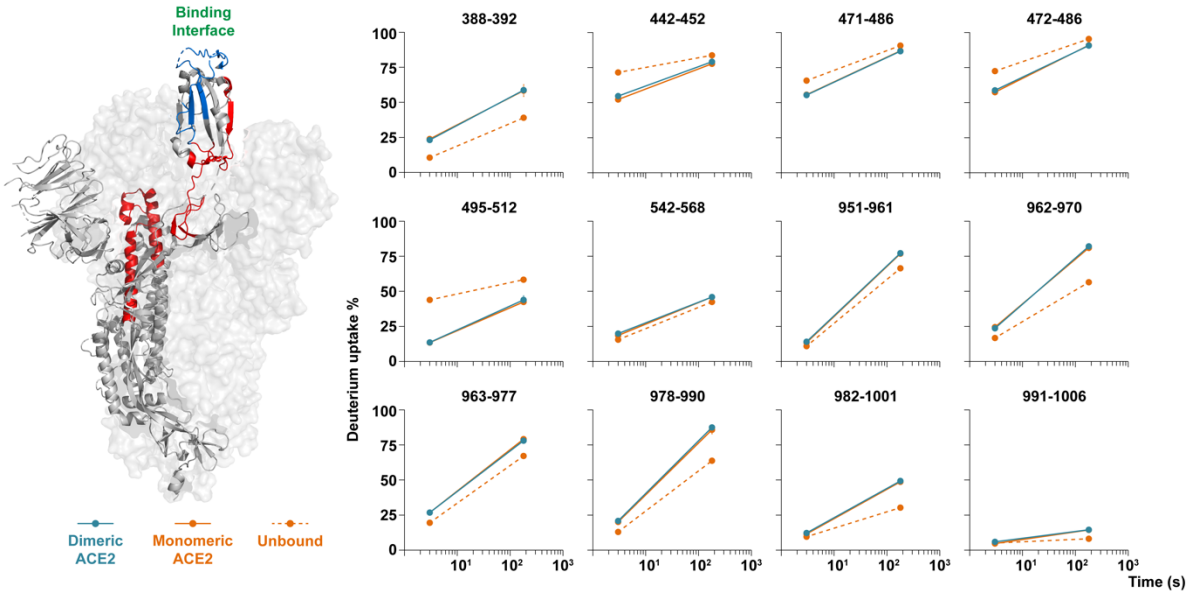

Figure S7. **Similar footprint on S dynamics from monomeric and dimeric hACE2 binding.** Overlapping points at 3-second and 3-minute incubation indicate dynamic impacts from monomeric hACE2 resemble native-like dimeric hACE2.

References

(1) Zhong, N. S.; Zheng, B. J.; Li, Y. M.; Poon; Xie, Z. H.; Chan, K. H.; Li, P. H.; Tan, S. Y.; Chang, Q.; Xie, J. P.; et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* **2003**, *362* (9393), 1353-1358. DOI: 10.1016/s0140-6736(03)14630-2.

(2) Phelan, A. L.; Katz, R.; Gostin, L. O. The Novel Coronavirus Originating in Wuhan, China: Challenges for Global Health Governance. *Jama* **2020**, *323* (8), 709-710. DOI: 10.1001/jama.2020.1097.

(3) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **2020**, *382* (8), 727-733. DOI: 10.1056/NEJMoa2001017.

(4) Hu, B.; Guo, H.; Zhou, P.; Shi, Z. L. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol* **2021**, *19* (3), 141-154. DOI: 10.1038/s41579-020-00459-7.

(5) 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. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **2013**, *503* (7477), 535-538. DOI: 10.1038/nature12711.

(6) Tortorici, M. A.; Veesler, D. Structural insights into coronavirus entry. *Adv Virus Res* **2019**, *105*, 93-116. DOI: 10.1016/bs.aivir.2019.08.002.

(7) 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. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579* (7798), 270-273. DOI: 10.1038/s41586-020-2012-7.

(8) Izaguirre, G. The Proteolytic Regulation of Virus Cell Entry by Furin and Other Proprotein Convertases. *Viruses* **2019**, *11* (9). DOI: 10.3390/v11090837.

(9) Sasaki, M.; Uemura, K.; Sato, A.; Toba, S.; Sanaki, T.; Maenaka, K.; Hall, W. W.; Orba, Y.; Sawa, H. SARS-CoV-2 variants with mutations at the S1/S2 cleavage site are generated in vitro during propagation in TMPRSS2-deficient cells. *PLoS Pathog* **2021**, *17* (1), e1009233. DOI: 10.1371/journal.ppat.1009233.

(10) Johnson, B. A.; Xie, X.; Bailey, A. L.; Kalveram, B.; Lokugamage, K. G.; Muruato, A.; Zou, J.; Zhang, X.; Juelich, T.; Smith, J. K.; et al. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* **2021**, *591* (7849), 293-299. DOI: 10.1038/s41586-021-03237-4.

(11) Kirchdoerfer, R. N.; Cottrell, C. A.; Wang, N.; Pallesen, J.; Yassine, H. M.; Turner, H. L.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B. Pre-fusion structure of a human coronavirus spike protein. *Nature* **2016**, *531* (7592), 118-121. DOI: 10.1038/nature17200.

(12) Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* **2017**, *27* (1), 119-129. DOI: 10.1038/cr.2016.152.

(13) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367* (6483), 1260-1263. DOI: 10.1126/science.abb2507.

(14) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e286. DOI: 10.1016/j.cell.2020.02.058.

(15) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579* (7798), 265-269. DOI: 10.1038/s41586-020-2008-3.

(16) Broer, R.; Boson, B.; Spaan, W.; Cosset, F. L.; Corver, J. Important role for the transmembrane domain of severe acute respiratory syndrome coronavirus spike protein during entry. *J Virol* **2006**, *80* (3), 1302-1310. DOI: 10.1128/jvi.80.3.1302-1310.2006.

(17) Matsuyama, S.; Nagata, N.; Shirato, K.; Kawase, M.; Takeda, M.; Taguchi, F. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J Virol* **2010**, *84* (24), 12658-12664. DOI: 10.1128/jvi.01542-10.

(18) Glowacka, I.; Bertram, S.; Müller, M. A.; Allen, P.; Soilleux, E.; Pfefferle, S.; Steffen, I.; Tsegaye, T. S.; He, Y.; Gnirss, K.; et al. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J Virol* **2011**, *85* (9), 4122-4134. DOI: 10.1128/jvi.02232-10.

(19) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181* (2), 271-280.e278. DOI: 10.1016/j.cell.2020.02.052.

(20) Low, J. S.; Jerak, J.; Tortorici, M. A.; McCallum, M.; Pinto, D.; Cassotta, A.; Foglierini, M.; Mele, F.; Abdelnabi, R.; Weynand, B.; et al. ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies. *Science* **2022**, *377* (6607), 735-742. DOI: 10.1126/science.abq2679.

(21) Bosch, B. J.; van der Zee, R.; de Haan, C. A.; Rottier, P. J. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* **2003**, *77* (16), 8801-8811. DOI: 10.1128/jvi.77.16.8801-8811.2003.

(22) Lu, M.; Uchil, P. D.; Li, W.; Zheng, D.; Terry, D. S.; Gorman, J.; Shi, W.; Zhang, B.; Zhou, T.; Ding, S.; et al. Real-Time Conformational Dynamics of SARS-CoV-2 Spikes on Virus Particles. *Cell Host Microbe* **2020**, *28* (6), 880-891.e888. DOI: 10.1016/j.chom.2020.11.001.

(23) Díaz-Salinas, M. A.; Li, Q.; Ejemel, M.; Yurkovetskiy, L.; Luban, J.; Shen, K.; Wang, Y.; Munro, J. B. Conformational dynamics and allosteric modulation of the SARS-CoV-2 spike. *Elife* **2022**, *11*. DOI: 10.7554/eLife.75433.

(24) Yang, Z.; Han, Y.; Ding, S.; Shi, W.; Zhou, T.; Finzi, A.; Kwong, P. D.; Mothes, W.; Lu, M. SARS-CoV-2 Variants Increase Kinetic Stability of Open Spike Conformations as an Evolutionary Strategy. *mBio* **2022**, *13* (1), e0322721..

(25) Hodge, E. A.; Benhaim, M. A.; Lee, K. K. Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci* **2020**, *29* (4), 843-855. DOI: 10.1002/pro.3790.

(26) Hodge, E. A.; Naika, G. S.; Kephart, S. M.; Nguyen, A.; Zhu, R.; Benhaim, M. A.; Guo, W.; Moore, J. P.; Hu, S. L.; Sanders, R. W.; et al. Structural dynamics reveal isolate-specific differences at neutralization epitopes on HIV Env. *iScience* **2022**, *25* (6), 104449. DOI: 10.1016/j.isci.2022.104449 PubMed-not-MEDLINE.

(27) Herrera, N. G.; Morano, N. C.; Celikgil, A.; Georgiev, G. I.; Malonis, R. J.; Lee, J. H.; Tong, K.; Vergnolle, O.; Massimi, A. B.; Yen, L. Y.; et al. Characterization of the SARS-CoV-2 S Protein: Biophysical, Biochemical, Structural, and Antigenic Analysis. *ACS Omega* **2021**, *6* (1), 85-102. DOI: 10.1021/acsomega.0c03512.

(28) Verkerke, H. P.; Williams, J. A.; Guttman, M.; Simonich, C. A.; Liang, Y.; Filipavicius, M.; Hu, S. L.; Overbaugh, J.; Lee, K. K. Epitope-Independent Purification of Native-Like Envelope

Trimers from Diverse HIV-1 Isolates. *J Virol* **2016**, *90* (20), 9471-9482. DOI: 10.1128/jvi.01351-16.

(29) Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *J Am Soc Mass Spectrom* **2013**, *24* (12), 1906-1912. DOI: 10.1007/s13361-013-0727-5.

(30) 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. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2019**, *176* (5), 1026-1039.e1015. DOI: 10.1016/j.cell.2018.12.028.

(31) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, *369* (6501), 330-333. DOI: 10.1126/science.abb9983.

(32) 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.; et al. Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell* **2020**, *183* (5), 1367-1382.e1317. DOI: 10.1016/j.cell.2020.10.043.

(33) 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* **2020**, *94* (7). DOI: 10.1128/jvi.00127-20.

(34) Benton, D. J.; Wrobel, A. G.; Roustan, C.; Borg, A.; Xu, P.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. *Proc Natl Acad Sci U S A* **2021**, *118* (9). DOI: 10.1073/pnas.2022586118.

(35) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **2020**, *183* (3), 739-751.e738. DOI: 10.1016/j.cell.2020.09.032.

(36) Zhang, J.; Cai, Y.; Xiao, T.; Lu, J.; Peng, H.; Sterling, S. M.; Walsh, R. M., Jr.; Rits-Volloch, S.; Zhu, H.; Woosley, A. N.; et al. Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **2021**, *372* (6541), 525-530. DOI: 10.1126/science.abf2303.

(37) Gobeil, S. M.; Janowska, K.; McDowell, S.; Mansouri, K.; Parks, R.; Manne, K.; Stalls, V.; Kopp, M. F.; Henderson, R.; Edwards, R. J.; et al. D614G Mutation Alters SARS-CoV-2 Spike Conformation and Enhances Protease Cleavage at the S1/S2 Junction. *Cell Rep* **2021**, *34* (2), 108630. DOI: 10.1016/j.celrep.2020.108630.

(38) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* **2020**, *117* (21), 11727-11734. DOI: 10.1073/pnas.2003138117.

(39) Benton, D. J.; Wrobel, A. G.; Xu, P.; Roustan, C.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **2020**, *588* (7837), 327-330. DOI: 10.1038/s41586-020-2772-0.

(40) Heurich, A.; Hofmann-Winkler, H.; Gierer, S.; Liepold, T.; Jahn, O.; Pöhlmann, S. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J Virol* **2014**, *88* (2), 1293-1307. DOI: 10.1128/jvi.02202-13.

(41) Fraser, B. J.; Beldar, S.; Seitova, A.; Hutchinson, A.; Mannar, D.; Li, Y.; Kwon, D.; Tan, R.; Wilson, R. P.; Leopold, K.; et al. Structure and activity of human TMPRSS2 protease implicated

in SARS-CoV-2 activation. *Nat Chem Biol* **2022**, *18* (9), 963-971. DOI: 10.1038/s41589-022-01059-7.

(42) Colman, P. M.; Lawrence, M. C. The structural biology of type I viral membrane fusion. *Nat Rev Mol Cell Biol* **2003**, *4* (4), 309-319. DOI: 10.1038/nrm1076.

(43) Benhaim, M. A.; Lee, K. K. New Biophysical Approaches Reveal the Dynamics and Mechanics of Type I Viral Fusion Machinery and Their Interplay with Membranes. *Viruses* **2020**, *12* (4). DOI: 10.3390/v12040413.

(44) Raghuvamsi, P. V.; Tulsian, N. K.; Samsudin, F.; Qian, X.; Purushotorman, K.; Yue, G.; Kozma, M. M.; Hwa, W. Y.; Lescar, J.; Bond, P. J.; et al. SARS-CoV-2 S protein:ACE2 interaction reveals novel allosteric targets. *Elife* **2021**, *10*. DOI: 10.7554/eLife.63646.

(45) Costello, S. M.; Shoemaker, S. R.; Hobbs, H. T.; Nguyen, A. W.; Hsieh, C. L.; Maynard, J. A.; McLellan, J. S.; Pak, J. E.; Marqusee, S. The SARS-CoV-2 spike reversibly samples an open-trimer conformation exposing novel epitopes. *Nat Struct Mol Biol* **2022**, *29* (3), 229-238. DOI: 10.1038/s41594-022-00735-5.

(46) Calvaresi, V.; Wrobel, A. G.; Toporowska, J.; Hammerschmid, D.; Doores, K. J.; Bradshaw, R. T.; Parsons, R. B.; Benton, D. J.; Roustan, C.; Reading, E.; et al. Structural dynamics in the evolution of SARS-CoV-2 spike glycoprotein. *Nat Commun* **2023**, *14* (1), 1421. DOI: 10.1038/s41467-023-36745-0.

(47) Pak, A. J.; Yu, A.; Ke, Z.; Briggs, J. A. G.; Voth, G. A. Cooperative multivalent receptor binding promotes exposure of the SARS-CoV-2 fusion machinery core. *Nat Commun* **2022**, *13* (1), 1002. DOI: 10.1038/s41467-022-28654-5.

(48) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367* (6485), 1444-1448. DOI: 10.1126/science.abb2762.

Chapter 4. Dynamic Study on Mosaic Spike Heterotrimers' Stability and Antigenic Impacts Provides Insights into Vaccine Development

4.1 Introduction

SARS-CoV-2, which is responsible for the COVID-19 pandemic, has been in circulation for nearly five years. The evolutionary pressure from vaccines and the human immune system has driven this virus to adapt and evolve.[1] The viral spike (S) glycoproteins of SARS-CoV-2, displayed on the viral surface mediate binding to angiotensin-converting enzyme 2 (ACE2) receptor during host cell entry, are also the primary antigens commonly targeted by the immune systems.[2, 3]

The monopartite genome (genome consists of a single molecule of nucleic acid) and proofreading function of exoribonuclease (ExoN) during SARS-CoV-2 genome replication, have limited its mutation space and rate, compared to many other infectious pathogens.[4, 5] However, significant mutations have emerged and new variants spread due to the massive number of infections that occurred globally. The D614G mutation was one of the first significant mutations on S, increasing the stability and infectivity of the virus.[6] Most neutralizing antibodies elicited after natural infection target the receptor-binding domain (RBD) of the spike protein.[7, 8] Subsequent mutations in the RBD, where neutralizing antibodies compete with host ACE2 receptor binding, have further enhanced viral fitness. Notable hotspot mutations in the RBD, such as N501Y, E484K/A, and K417N/T, combined with D614G mutation, have led to peaks in infection waves associated with variants of concern (VOCs).[1, 9, 10]

Omicron variant, identified in late 2021, is notable for its extensive mutations in the spike RBD, enhancing its ability to evade immune responses and increase transmissibility.[11, 12] Omicron

strain rapidly spread worldwide and diverged into many subvariants such as BA.1, BA.2, and BA.5, and their offspring subvariants.

These VOCs of SARS-CoV-2 have primarily been driven by antigenic drift, where point mutations occurred and gradually altered the antigenic profile of the spike protein, improving the viral fitness. However, studies on the emerging strain, XBB, have suggested the presence of antigenic shift through genomic recombination in SARS-CoV-2 strains.[13-15] XBB is a recombinant variant resulting from the combination of two different Omicron subvariants, BM.1 and BJ.1, at genomic position 22920.[13] This recombinant nature allows XBB to have the N-terminal homologous amino acid sequence to BJ.1 up to position 459, and the C-terminal sequence from BM.1, inheriting structural features from both subvariants with enhancements in human ACE2 (hACE2) binding and immune evasion.[16, 17] Although the original host of this recombinant strain is unclear, the emergence of the XBB variant indicates the capability of SARS-CoV-2 viruses to co-infect a single cell with different strains, leading to the production and proliferation of recombinant strains.

At the same time, vaccines featuring entire S or RBD antigens have also made progress in preventing severe infections.[18-23] In particular, the efficacy of mRNA vaccines was demonstrated in the COVID-19 pandemic. Advancements in mRNA vaccine technology over the last decade, such as 5’ capping, N1-methylpseudouridine base modification, and ionizable lipid-PEG encapsulation, have significantly improved the translation efficiency and stability of mRNA vaccine.[24, 25] The Pfizer-BioNTech (BNT162b2)[18] and Moderna (mRNA-1273)[21] vaccines, two most prominent mRNA vaccines developed for COVID-19, demonstrated high efficacy on a global scale by leveraging the human body's cellular machinery to produce antigens and elicit an immune response.

In response to the emergence of Omicron strains circulation, a special booster, bivalent mRNA vaccines were developed and deployed in fall 2022. 1:1 mix of mRNAs encoding Wuhan-Hu-1 WT and Omicron BA.4/5 were encapsulated in ionizable lipid nanoparticles and delivered to human cells in order to display both antigens for human immune responses. Clinical trials indicated prominently improved protection compared to non-vaccinated and monovalent-vaccinated groups.[26]

In both scenarios, where two viral strains infect the same host cell during SARS-CoV-2 infection or when two encapsulated mRNA sequences are delivered to the same cell during mRNA vaccination, it is possible that different spike mRNAs could be translated concurrently and resulted in translated peptides co-assembling into a mosaic heterotrimer form. However, to our knowledge it has not been determined whether protomers from different constructs can co-assemble as a mosaic heterotrimer and how this might impact the trimer's stability, dynamics and antigenicity. Few structural studies have attempted to address this question. Understanding whether such mosaic heterotrimers can form and their implications on the structural properties of spike protein is crucial, as it could influence both viral responses to host immune system and vaccine efficacy.

To investigate this, in this chapter, I focus on experiments demonstrating successful purification of mosaic trimer assembled by genetically close strains Omicron BA.2 and XBB (OX), as well as distant strains Omicron BA.2 and original Wuhan-Hu-1 G614 (OG). The trimer stability and integrity are well-maintained in the mosaic spike trimers. Hydrogen/deuterium-exchange mass spectrometry (HDX-MS) and biolayer-interferometry (BLI) were used to investigate the mosaic spike dynamics and impacts on RBD structural orders and receptor binding. These are important issues to investigate as numerous vaccine studies employing gene delivery such as with the mRNA-lipid nanoparticle modality that are pursuing pan-coronavirus or universal influenza

vaccines rely upon co-administration of panels of antigen variants. The study we report in this chapter provides structural dynamic information to help understand the impact of co-transfection and co-expression on the combinatorial antigens that are produced in these scenarios.

4.2 Materials and Methods

Plasmid construction

The gene sequences encoding SARS-CoV-2 S ectodomain with hexa-proline mutations (S-6P: F817P, A892P, A899P, A942P, K986P and V987P)[27], GSAS substitution at S1/S2 site (RRAR) and C-terminal T4 fibritin foldon were cloned into pCMV plasmid. The ORF sequence of G614 construct is the same as mentioned in previous chapter with a C-terminal His-tag. Omicron BA.2 construct was assembled from the ORF in the pαH Omicron BA.2 S-2P plasmid (Addgene plasmid #184829) into pCMV plasmid and mutated with four additional prolines (NEB). XBB construct was assembled from the ORF in the pVRC8400 XBB S-2P plasmid (Addgene plasmid #160474) into pCMV plasmid and mutated with four additional prolines and a C-terminal His-tag (NEB).

Transient transfection

293T adherent cell line was used for optimizing co-transfection condition. Cells were split into 12-well plate (1 mL per well) at seeding density (0.1×10^6 cells per well). Confluent cells (0.5×10^6 cells per well) were transfected with total 2 μg DNA and 2 μL Lipofectamine™ 3000 (Invitrogen) based on Lipofectamine™ 3000 reagent protocol. After 2 days in the 37 °C incubator (5% CO_2), 200 μL supernatant from each well was collected for SDS-PAGE to check expression level.

Suspension Expi293F cell culture was used for spike protein expression. For a 700 mL cell culture at confluency (3×10^6 cells/mL), 0.7 mg total DNA (1:1 for co-transfection) and 2.1 mL ExpiFectamine™ 293 reagent were mixed in Gibco Opti-MEM and added into the cell culture following ExpiFectamine™ 293 user guide. The transfected culture was incubated at 37 °C and 5% CO_2 with constantly shaking at 125 RPM for 5 days before harvest.

Protein purification

On the fifth day post-transfection, Expi293F cells were harvested by centrifugation at 2000 RCF for 30 minutes. Supernatant was vacuum-filtered through 0.45 μm αPES filters and supplemented with Tris-HCl (pH 8.0) and NaCl.

For G614 and XBB spike homotrimers purification, imidazole was added to the supernatant to reach a final concentration of 50 mM Tris-HCl (pH 8.0), 200 mM NaCl and 20 mM imidazole, same as the Ni-IMAC binding buffer composition. A 2 mL Ni-IMAC resin (Thermo Fisher Scientific) was washed and equilibrated with binding buffer and packed for a gravity flow column. Supernatant was loaded and flowed through this gravity flow column, followed by four times 10 mL (5 CV) wash with binding buffer. The bound protein was then eluted from the column with three times 6 mL (3 CV) Ni-IMAC elution buffer (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 300 mM imidazole). Elution fractions were SDS-PAGE checked, and then combined. Combined spike protein solution was buffer-exchanged to reduce the imidazole level below 1 mM and concentrated to ~1 mg/mL with an Amicon Ultra-15-mL-centrifugal filter (30 K, Millipore), followed by flash frozen in liquid nitrogen and stored in -80 °C freezer.

For Omicron, mosaic Omicron-G614, and mosaic Omicron-XBB trimers purification (Figure S1), ethylenediaminetetraacetic acid (EDTA) was added to a reach a final concentration

of 100 mM Tris-HCl (pH 8.0), 150 mM NaCl and 1 mM EDTA, same as the Strep-Tactin® XT wash buffer composition. A 3.5 mL Strep-Tactin® XT 4Flow® high capacity resin was washed and equilibrated with wash buffer and packed for a gravity flow column. Supernatant was loaded and flowed through this gravity flow column, followed by four times 8.5 mL (2.5 CV) wash. The bound protein was then eluted from the column with three times 5 mL (1.5 CV) Strep-Tactin® XT elution buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM biotin). Elution fractions were combined, buffer-exchanged to reduce the biotin level below 1 mM and concentrated to a final volume below 1 mL. HRV 3C protease (Pierce™) was added accordingly based on the protein solution volume for the digestion to remove TwinStrep tag overnight at 37 °C. On the next day, digested samples were spun at 15000 RCF for 10 minutes to remove possible protein aggregation pellet. The Omicron homotrimer protein could now be flash frozen in liquid nitrogen and stored in -80 °C freezer. For both mosaic heterotrimer purifications, samples were flowed through a 1 mL pre-equilibrated prepacked StrepTrap™ XT column (Cytiva) at 1 mL/min to separate undigested protein. Flow-through was reloaded to the column for better purification results. The final flow-through containing heterotrimer with TwinStrep tag removed was collected and buffer-exchanged to the nickel column binding buffer. Samples were flowed through a 1 mL pre-equilibrated prepacked HisTrap™ HP column (Cytiva) at 1 mL/min, followed by 7 mL (7 CV) wash. The bound protein was then eluted from the column with 3 mL (3 CV) elution buffer. SDS-PAGE checked elution fractions were combined, buffer-exchanged to reduce the imidazole level below 1 mM and concentrated to ~1 mg/mL with an Amicon Ultra-15-mL-centrifugal filter (30 K, Millipore). The duo-affinity-column purified mosaic spike heterotrimers were then flash frozen in liquid nitrogen and stored in -80 °C freezer.

All the spike trimers were finally size-exclusion chromatography (SEC) purified before major experiments. 500 μL sample was thawed and spun at 15000 RCF for 10 minutes to remove possible protein aggregation pellet before injecting to Superdex 200 column (GE) on an AKTA Pure system (GE). With the SEC running buffer (10 mM HEPES, 200 mM NaCl, 0.02% NaN_3, pH 7.5) and flow rate at 0.5 mL/min, spike trimer was eluted at around 8.5-9 mL. The SEC purified spike trimers were concentrated to desired concentrations according to the experiment need.

Dynamic light scattering

Dynamic light scattering (DLS) measurements were performed on a DynaPro NanoStar II (Wyatt) to characterize spike trimer integrity, homogeneity and thermal stability. SEC purified spike samples at 0.5 mg/mL were first spun down at 15000 RCF for 20 minutes to remove aggregation pellet. 5 μL of the sample was then injected into a 2-μL quartz cuvette (Wyatt). Each DLS run was measured with twenty 10-second acquisitions at 25 °C by LASER auto-attenuation.

Hydrogen/Deuterium-exchange mass spectrometry

For the hydrogen/deuterium-exchange (HDX) experiments in this chapter, 10 μg each of Hu-1 G614, Omicron BA.2, XBB, mosaic Omicron-G614 (OG) and mosaic Omicron-XBB (OX) spike trimer was incubated in the deuteration buffer (10 mM HEPES, pH^* 7.5, 85% D_2O, Cambridge Isotope Laboratories, Inc.) at 22 °C for three different timepoints: 3, 60, 900 seconds (five replicates for 3 seconds and four replicates for 60 and 900 seconds). All exchanged samples were immediately mixed with an equal volume of ice-chilled quench buffer containing 8 M urea, 200 mM tris(2-chloroethyl) phosphate (TCEP) and 0.2% formic acid (FA) to pH 2.5, transferred into an autosampler vial with metal cap and flash frozen in liquid ethanol. Samples were loaded to

the LC-MS system by an automating platform and analyzed by a Synapt G2 mass spectrometer (Waters) with the settings described in previous chapter.[28] Samples were in-line digested into peptides by an immobilized pepsin column (2.1 × 50 mm) and loaded onto CSH C18 trap cartridge (Waters) with loading buffer (2% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA)) at 800 μL/min. Peptides were then separated by a UPLC CSH C18 column (1.0 × 50 mm, 1.7 μL, Waters) with a 15-min linear gradient from 3% to 40% buffer B (buffer A: 2% ACN, 0.1% FA, 0.025% TFA; buffer B: 99.9% ACN, 0.1% FA) at a 40 μL/min flow rate.

Omicron and XBB undeuterated samples were also run on a Orbitrap™ Ascend Tribrid™ mass spectrometer (Thermo Fisher) to obtain MS/MS information and identify pepsin digested peptides of these two constructs by Byonic. Peptide mass spectra were further confirmed on DriftScope (Waters) and identified with specific retention time and drift time. HDExaminer (Sierra Analytics) was used to parse the mass spectrometer datasets and calculate deuterium uptake of each peptide identified by the specific retention time and drift time.

SDS-PAGE and western blot

For all the SDS-PAGE gels using Coomassie Blue staining, 11 μL spike samples were mixed with 1 μL DTT (0.5M) and 4 μL 4× NuPAGE™ LDS loading dye (Invitrogen) before loading to 4-12% NuPAGE™ Bis-Tris gel (Invitrogen) wells. SDS-PAGE gels were run at 120 Volts for 10 minutes and 150 Volts for 60 minutes in MES SDS running buffer (Invitrogen), followed by gel collection and PageBlue™ (Thermo Scientific) staining.

For SDS-PAGE gels proceeding to western blots, 5 μL spike samples were mixed with 1 μL DTT (0.5M) and 2 μL 4× loading dye, loaded to the SDS-PAGE gel wells, and run at 120 Volts for 80 minutes. The peptides on the gel were then transferred to the Immobilon®-FL PVDF

membrane (Millipore) using 10% methanol supplemented NuPAGE™ transfer buffer (Invitrogen) at 30 Volts for 2 hours. Membrane with total secreted proteins from 293T culture supernatant was stained with Ponceaus S for imaging supernatant loading control from Bio-Rad gel imager prior to blocking. All membranes were blocked in 5% non-fat milk overnight, followed by incubation in primary antibody, respectively, in room temperature for 1 hour. Primary antibody against spike S2 subunit (1A9, mouse, mAb, Invitrogen, MA5-35946) was used in 1:2000; against Strep-tag (mouse, mAb, GenScript, A01732S) was used in 1:2000; against His-tag (mouse, mAb, Invitrogen) was used in 1:2000 tris-buffered saline-Tween (TBST) solution. Secondary antibody anti-mouse IgG was used in 1:5000 5% non-fat milk blocking buffer in room temperature for 1 hour. The membranes were imaged on the Odyssey® M-XS imaging system (LI-COR Biosciences) using 700 nm channel and optimized exposure intensity.

Bio-layer interferometry

Bio-layer interferometry (BLI) was used to measure binding kinetics between spikes and hACE2 receptor or neutralizing antibodies. Octet® anti-human Fc capture (AHC) biosensors (Sartorius) were pre-incubated in assay binding buffer (10 mM phosphate-buffered saline (PBS), pH 7.4, 0.1% bovine serum albumin (BSA), 0.05% Tween 20) for at least 15 minutes. Spike samples were 2-fold serial diluted with binding buffer from 200 nM to 3.1 nM. Ligand samples were also prepared with binding buffer at 2 μg/mL concentration for dimeric hACE2-Fc and 1 μg/mL for antibodies. The binding assay was performed on an Octet® BLI system (Sartorius) with 60s ligand loading, 180s association and 180s dissociation processes. BLI kinetic data were processed by ForteBio Data Analysis software (v11.0) with 1:1 binding model fitting.

Negative-stain electron microscopy

SEC purified spike trimers were diluted to 20 ng/μL for negative-stain electron microscopy (nsEM). Carbon film 300 mesh copper grids (Electron Microscopy Sciences) were glow discharged for 30 seconds in a vacuum chamber. 3 μL of each diluted spike sample was added onto the grid (carbon side), incubated for 1 minute before blotting off with a piece of folded filter paper. The grid was then incubated in one drop of nano-W (Methylamine Tungstate, Nanoprobes) for 1 minute, followed by blotting off excess stain solution to leave a thin layer. The air-dried grid was saved in the grid box and imaged on a FEI Tecnai G2 Spirit TEM (120 kV) with a Gatan Ultrascan 4000 CCD detector. Data collection was performed using Leginon with an image pixel size of 1.6 Å.

4.3 Results

Mosaic spike trimers can be stably formed from genetically close strains (Omicron and XBB) and distant strains (Hu-1 and Omicron)

To understand if co-transfection mimicking bivalent mRNA vaccine function could result in mosaic SARS-CoV-2 S heterotrimer formation, we first compared the total spike expression levels between co-transfections and single plasmid transfections. Equal amounts of total plasmid DNA were transfected or co-transfected into equal amounts of cell cultures under nine conditions: Omicron only, G614 only, XBB only, Omicron-G614 co-transfection (2:1, 1:1, and 1:2), and Omicron-XBB co-transfection (2:1, 1:1, and 1:2). Western blot analysis using an antibody targeting a conserved linear epitope in the S2 subunit across all three strains showed that all conditions expressed similar amounts of spike protein into the cell culture supernatant (Figure 1b). This indicated that all three hexaproline-mutated ectodomain S (S-6P) variants used in this chapter

showed similar transfection efficiency and expression level, further with no significant interference from co-transfection between Omicron and G614, and Omicron and XBB.

While SDS-PAGE denatured trimers, to evaluate if mosaic heterotrimers could be formed and how they affected trimer stability, we performed protein preparations of mosaic Omicron-G614 (OG) and Omicron-XBB (OX) side by side with control purifications of Omicron, G614 and XBB homotrimers. Using the optimized purification protocols and low-passage viable Expi293F cells, the best yields from recent parallel purifications of all five constructs were obtained (Table S1). Due to different steps of affinity purifications applied among the five constructs (G614 and XBB: His-tag purification only; Omicron: strep-tag purification only; OG and OX: 2 × strep-tag purification and His-tag purification), the crude protein yields were calculated from the yields after the first affinity purification with the adjustments of possible protein loss estimated from previous affinity purification trials (~25%). Therefore, the data suggested that all S-6P constructs generally produced 8-11 mg of crude spike trimers per liter of culture at the beginning of the affinity purification steps.

Nearly 90% of the protein input appeared as a single peak at 8.5-9 mL in the final SEC purification, indicating well-behaved trimers with comparable yields around 4-5 mg per liter of culture (Figure 1c). Furthermore, dynamic light scattering (DLS) measurements (Table S2) and negative-stain electron microscopy (nsEM) images (Figure 1d) also confirmed that the size and trimeric morphology of mosaic heterotrimers were similar to their homotrimer counterparts. Therefore, both Omicron and XBB protomers, which possess exactly the same S2 subunit sequence, and Omicron and G614 protomers, which carry more mutations around the S1/S2 site and S2 subunit (Figure 1a), could form mosaic heterotrimers with limited alterations in trimeric structure and morphology.

Furthermore, based on the yield ratios of the isolated mosaic heterotrimer to the Omicron homotrimer from the His-tag purification (Table S1, OG: 4.7 mg/1.9 mg; OX: 3.6 mg/1.2 mg), the mosaic trimers were estimated to be ~60% of the total trimer yield from co-transfection, closely aligning with the theoretical percentage (75%) for stochastic trimer assemblies from 1:1 co-transfection. This suggests that, in both genetically close strains (Omicron and XBB) and distant strains (Omicron and G614) co-transfection, homotrimer formation did not overwhelm heterotrimer formation. It could be extrapolated that in biological environments, such as two strains co-infecting a single cell or bivalent mRNA-LNP uptake by a single cell, co-assembled mosaic spike trimers are likely to form and be displayed as antigens.

Dynamic features of Omicron and XBB protomers in the mosaic heterotrimer are similar to their homotrimer counterparts

If natural co-infection and bivalent mRNA vaccination could result in mosaic heterotrimer formation as our experiments suggested, we next investigated how mosaic trimer formation impacted the structural dynamics of S conformational states, and trimer integrity. Herein, we have set up comprehensive HDX-MS experiments among three S-6P homotrimers: G614, Omicron and XBB, and two heterotrimers: Omicron-XBB and Omicron-G614.

Comparisons among G614, Omicron and XBB homotrimers indicated differences in conformational preferences and HR1 dynamics due to sequence variations of these three constructs (Figure 2a). Increased dynamics in peptide 388-392 near the C-terminal of RBD facing the apex of central helices and in peptide 982-990 at the central helical apex are representative of the S open conformation (Figure 2b), as discussed in previous chapters. When S is in the open state, the RBD up conformation also affects the S1 hinge region dynamics connecting RBD to the rest of the S.

The N-terminal peptide 516-533 connecting to the RBD, showed increased dynamics in response to RBD up shift, while the other end, peptide 542-568, is more protected by the neighboring protomer towards the trimer core. These dynamic changes suggested the preference for the S open state is more prominent in G614 than in Omicron and XBB (Figure 2b).

Sequence variations in the NTD and RBD, especially in the RBM regions among the three constructs resulted in varied pepsin digestion patterns and limited homologous peptide profiles that could provide informative direct comparisons in these two regions (Figure 2a). Omicron and XBB share the same S2 sequences and digestion patterns, with only dynamic differences along S2 subunit shown in the HR1 region in respond to less open states in XBB (Figure 2b, peptide 962-981 and 982-990). However, mutations acquired around the S2' and HR1 region in Omicron from original Hu-1 G614 strain affected the S2 subunit dynamics near the S2' site (peptide 732-740 and 962-981), beyond the impacts on S conformation preferences (Figure 2b). Overall, the HDX-MS data indicated that Hu-1 G614 S could adopt more open states while Omicron and XBB S possessed higher dynamics in HR1 domain extending to the fusion peptide proximal region (FPPR).

To understand the dynamic impacts of forming mosaic heterotrimers, we compared the homologous peptides in the mosaic trimer with those in the corresponding homotrimers. In Omicron-XBB (OX) mosaic trimer, the two participating constructs mostly retained their dynamics as in their homotrimers. The dynamics of NTD and RBD from each protomer were not significantly interfered by mosaic formation. The peptide uptake profiles in the NTD of OX were intermediate between corresponding peptides in Omicron and XBB homotrimers (Figure 3, peptide #1 and #2), averaging the uptake levels from indistinguishable mass spectra. Due to the peptide length and uptake effect size, the bimodal behaviors in these OX peptides could not be deconvoluted. A comprehensive example is shown in peptide 177-191 (Figure 3, peptide #3),

where the Q183E mutation in XBB makes the homologous peptides two separate amino acid sequences to compare. The peptide in Omicron protomer of the OX mosaic trimer showed the identical dynamic profile as in the Omicron homotrimer and the peptide in XBB protomer matched the XBB homotrimer peptide as well. Ratios for the RBD adopting the up conformation were also averaged between Omicron and XBB. Peptides refection RBD conformation switches such as C-terminal RBD (Figure 3, peptide #4) and apex of the central helix (Figure 3, peptide #5 and #6) lay in the middle of the two homotrimers. Overall, the dynamics changes in mosaic OX trimer were a compromise between Omicron and XBB (Figure 3). Similar dynamics were retained in NTD and RBD compared to their homotrimer counterparts when such two strains sharing the identical S2 subunit forming mosaic heterotrimer.

Distinct structural dynamics reveal in homologous peptides across different protomers in the Omicron-G614 mosaic heterotrimers

To understand how the mosaic S trimer is packed from non-identical S2 subunits and if allosteric effects occur, we next investigate the dynamic behavior in Omicron-G614 (OG) mosaic heterotrimer. Similar effects were observed in the RBD as in the OX case, with unresolved deuterium uptake mass envelopes appearing intermediate between those in Omicron and G614 homotrimers (Figure 4, peptide #1 and #2). Peptide 471-486, spanning the loop connecting the RBD core and the receptor-binding motif (RBM) located at the surface of RBD, is highly mutated during viral evolution for enhanced antibody escaping. The S477N, T478K, and E484A mutations in the Omicron make the homologous peptide 471-486 ideal for direct comparison between G614 protomer and Omicron protomer. As expected, the uptake level in each protomer overlapped on the peptide in the homotrimer, respectively (Figure 4, peptide #3). Unlike the OX case, the

conformational preferences in OG favored the open conformation. Since G614 spike could exhibit higher avidity and possibly adopt two or three RBD up conformations, the peptides reflecting RBD conformations in OG behaved more similarly to the G614 homotrimer (Figure 4, peptide #1, #2 and #6). One exception was the hinge region (Figure 4, peptide #4), where dynamic behavior of this peptide in OG trimer was closer to the Omicron pattern. As previously discussed in chapter 2, a higher portion of open states should show decreased dynamic in peptide 553-568. While increased dynamics at C-terminal RBD and helical apex are directly responded to the exposure in open conformation, the dynamics of the peptide 553-568 loop is quite related to the positioning and packing of the NTD on the neighboring protomer. Therefore, even though the open conformation is more favored in OG mosaic heterotrimer than Omicron homotrimer, the hinge loop could be as exposed as in Omicron trimer since the NTD on the neighboring protomer retained similar conformation and dynamics as in Omicron homotrimer.

One prominent dynamic change in OG heterotrimer occurred in the HR1 domain, near the helical apex and fusion peptide proximal region. Homologous peptides in the region 962-981 were differentiated from two protomers due to an N969K mutation in Omicron. In homotrimer case, this peptide was intrinsically more dynamic in Omicron homotrimer than G614 homotrimer (Figure 2). It also showed the bimodal mass envelope at 1-minute time point with slow-uptake peak resembling uptake level in G614 homotrimer (Figure S5). The existence of the fast-uptake population indicated that FPPR in Omicron homotrimer could adopt more exposed conformations, which agreed with previous study.[29] However, the same region on the Omicron protomer in OG mosaic trimer showed decreased uptake levels with less detectable peak broadening and bimodal at 1-minute time point (Figure 4, peptide #7 dark purple, Figure S5). On the other hand, the homologous peptide in G614 protomer showed dramatically higher dynamics in OG mosaic trimer

than in G614 homotrimer (Figure 4, peptide #7 light purple). The occurrence of bimodal mass envelope at 1-minute time point consisted of a slow-uptake population resembling uptake profile in G614 homotrimer and a large fast-uptake population (Figure S5). This suggested that the G614 protomer could adopt the more exposed conformations at FPPR when OG mosaic trimer formed. This bimodal behavior shifting between the two distinct protomers, resulted in OG mosaic heterotrimer maintaining both the favored open states as in G614 homotrimer and highly exposed FPPR as in Omicron homotrimer. Despite these prominent dynamic changes in mosaic trimer formation, the trimer integrity and stability were maintained, as indicated by well-protected central helix domains (Figure 4, peptide #8, Figure S4).

Retained structural ordering of mosaic S trimer receptor-binding interface suggests similar antigenic impacts

The HDX experiments of both mosaic OX and OG trimers suggest that highly similar S2 subunits result in well-behaved trimer, and the dynamics in the NTD and RBD are less affected by the mosaic heterotrimer formation. Previous studies have reported that over 90% of neutralizing antibodies targeted RBD and NTD. Retaining the structural orders in NTD and RBD is important for assessing the antigenicity as well as vaccine efficacy.

To further investigate this, we used bio-layer interferometry (BLI) to measure the binding kinetics changes of mosaic S trimer formation. In the dimeric hACE2 binding experiments, all five constructs showed strong binding affinity below 10 nM (Figure 5). 2 μg/mL concentration of dimeric hACE2 was immobilized on the Fc-probes with 1:1 binding mode applied for calculating the binding affinity, which was aimed to alleviate the avidity effect. XBB homotrimer showed the weakest binding affinity at 10 nM to dimeric hACE2. Mosaic OX trimers showed K_{on} value

between Omicron and XBB homotrimer and comparable K_{off} value, reflecting a 9.5 nM hACE2 binding affinity (Figure 5). Mosaic OG trimer exhibited enhanced binding affinity from comparable K_{on} rate as G614 and K_{off} rate as Omicron.

We then tested the binding with monoclonal antibody (mAb) C68.13, which was identified from the patient serum. In the published competing binding assays against the hACE2 receptor, C68.13 showed strong binding on S RBD, competing hACE2 at RBM binding. However, C68.13 exhibited 2-3 orders of magnitude weaker binding EC50 on XBB and BQ strains.[30] In our binding assay, 1 μg/mL concentration of C68.13 was immobilized on the Fc-probes in all the experiments. As expected, the G614 homotrimer, Omicron homotrimer and OG mosaic trimer showed ~1 nM K_D. Flat dissociation curves resulted in K_{off} rates in lower 10E-4 levels, close to the detection limit, indicating very strong binding. The XBB homotrimer showed 95 nM K_D, 50-to-100-fold weaker binding to this mAb than G614 and Omicron trimers. The mosaic OX showed comparably strong binding from almost 5-fold K_{off} rate changes from XBB homotrimers.

In both hACE2- and selected mAb-binding cases, the interface structural order of mosaic spike heterotrimer were shown to be well-conserved. Our BLI results align with the inferences from the HDX experiments, suggesting that mosaic heterotrimers (OX and OG) impose limited restriction on RBD dynamics and receptor / antibodies binding, and can be efficient as displayed antigens.

4.4 Discussion

In this chapter, we have studied and evaluated the formation of mosaic S heterotrimer through co-transfection. We selected three SARS-CoV-2 variants: Hu-1 G614, Omicron BA.2 and XBB, based on their genetic distances and prevalence. The Hu-1 G614 variant, first detected and

sequenced in early 2020, led to subsequent infection waves worldwide before vaccination was deployed across large populations. In contrast, Omicron subvariants and XBB emerged in late 2021 and 2022, respectively, following the administration of two or three vaccine doses in most developed countries. The genetic distance between Hu-1 G614 and Omicron BA.2 is relatively large compared to the distance between Omicron and XBB. Particularly, XBB is the recombinant variant from two Omicron subvariants, resulting in notably different S1 but identical S2 subunit sequences. Therefore, the choices of co-assembled spike from Hu-1 G614 and Omicron BA.2 (OG) and Omicron and XBB (OX) could be representative of studying the impacts of S2 subunit sequences on mosaic heterotrimer formation as well as the outcomes from bivalent mRNA vaccine.

The S homotrimers of these three variants showed distinct intrinsic dynamic behaviors. The G614 variant favored more open states compared to Omicron and XBB, which exhibited higher FPPR dynamics, in agreement with previous HDX findings featuring S dynamics in Delta and Omicron strains.[29] While adopting more open states in G614 variant increased the likelihood of successful hACE2 binding, the more exposed FPPR dynamics in Omicron and XBB could facilitate spike activation and fusion peptide release. Both structural features resulted in enhancement of host cell recognition and viral fusion.

Upon forming mosaic S trimers from co-transfection, the overall protein stability and morphology did not alter from corresponding homotrimers. The yield of mosaic trimer closely matched the theoretical yield from arbitrary matching, suggesting a negligible extra energy cost for folding protomers from different variants into mosaic trimer compared to their homotrimers. We speculate that this well-folded and packed S morphology is due to the highly similar S2 subunit sequence and shape.

In previous chapters, we discussed how allosteric effects influence S conformational states. Sequence variants in the NTD and RBD could affect RBD-NTD crosstalk, resulting in Omicron S having more open states than XBB, despite sharing identical S2 sequence. When forming the mosaic OX trimer, the dynamics in the S2 subunit were retained, as expected. The RBD-NTD crosstalk was found to be intermediate, so the movements of RBD and NTD in the mosaic OX were not impaired, reflecting similar structural orders in RBD and NTD regions compared to their respective homotrimers. The impacts on mosaic spike RBD and NTD were further evaluated by BLI from ligand binding, confirming the structural order of the RBD, especially the RBM, was well-conserved.

For the OG trimer where the protomers do not share the identical S2 subunit, the local dynamics in the FPPR were affected by both the point mutation in the HR1 domain and the transfer of conformational preferences over protomers (bimodal behaviors). The bimodal behavior was not observed in the G614-S2-bundled homotrimer but Omicron-S2-bundled homotrimer, indicating the Omicron-S2-bundled structure induced strong dynamic potential in FPPR (Figure S5). When G614 protomer replaced Omicron protomer in OG trimer, this highly dynamic potential in FPPR was transferred from the Omicron protomer to the G614 protomer, making G614 protomer preferable in adopting more exposed conformation. Therefore, the deuterium uptake levels in this region showed the Omicron-down-G614-up changes (Figure 4, peptide #7, Figure S5).

This prominent dynamic change did not have a detrimental impact on trimer integrity and stability, as revealed by the well-protected trimer central helix domains (Figure 4, peptide #8, Figure S4). Therefore, the formation of the OG mosaic trimer retained both the preferences in open states from G614 and highly dynamic FPPR from Omicron, with limited impacts on spike stability and a low risk of misfolding.

Throughout the COVID-19 period, various structural-based vaccine designs focused on S- or RBD-based antigen display.[23] Our experiments successfully purified the stabilized mosaic trimers and investigated their dynamic behavior, providing insights into the possibly natural setup of displaying multivalent antigens. The sequence conservation of S2 subunits in most strains from early 2020 to 2023, on the one hand, enable the possibility of displaying mosaic heterotrimers for multivalency with minimal engineered sequence optimizations.

The first bivalent mRNA vaccine, consisting of spike-coding mRNA from both Hu-1 and Omicron BA.4/5 strains, revealed promising clinical trial results on providing protection from various VOCs.[26] Our co-transfection experiment mimicking the process of vaccine antigen display, showed that mosaic spike trimer could be stably formed and antigenically functional. However, there remains a gap in understanding whether the co-assembly occurred *in vivo*. Possible pseudovirus display experiments and immune response studies on mosaic antigens, are needed to investigate if mosaic S antigens is capable to induce novel and/or broadly neutralizing antibodies that enhances the bivalent vaccine efficacy.

Seeking pan-coronavirus or universal influenza vaccines is believed to be a solution for generating broader immune responses and protecting people from highly mutable and infectious viruses. Typical broader antibody generation from multivalent vaccination relies on the co-administration of antigen variants as a cocktail. Advanced technology utilizing self-assembled nanoparticles can display multiple antigens on a single platform.[31] However, mRNA-LNP vaccines do not administer the protein antigens directly but rather the mRNAs encoding them, leading to a new approach for expanding the antigen pool and diversity. While current multivalent mRNA vaccines aim to increase antigen diversity by mixing different mRNAs, the concurrent translation could result in the formation of mosaic protein multimers. Our experiments have shown that, at

least in the SARS-CoV-2 spike fusion protein system, the formation of mosaic heterotrimers retained antigenicity and induced novel dynamic behaviors related to selected strains. Therefore, if this phenomenon occurs with multivalent mRNA vaccines, the displayed antigen space could theoretically be more diverse, providing valuable insights into the design strategies for pan-coronavirus or universal influenza vaccines.

Figures

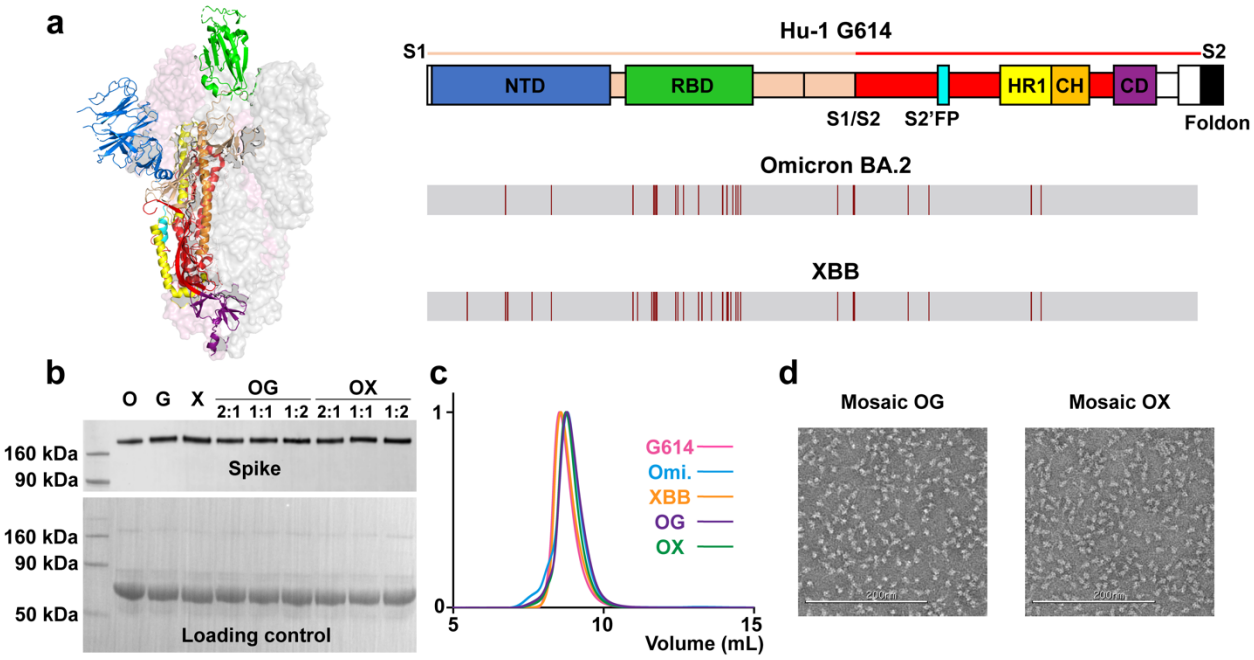

Figure 1. **Expression and structural characterization of mosaic S-6P heterotrimers. (a)** Color-coded S domains on cartoon diagram (PDB: 6VSB) and sequence alignments. Dark red bands in Omicron BA.2 and XBB indicate mutations with reference to Hu-1 G614 sequence. **(b)** Western blot images of spike expression in various conditions. Spike bands are blotted by anti-S2 primary antibody. Loading control bands are stained by Ponceau S. **(c)** SEC chromatograms of homotrimers and mosaic heterotrimers purifications. Major peak eluting at ~9 mL reveals the trimeric S integrity. **(d)** Negative-stain images illustrate the morphology of Omicron-G614 heterotrimer and Omicron-XBB heterotrimer.

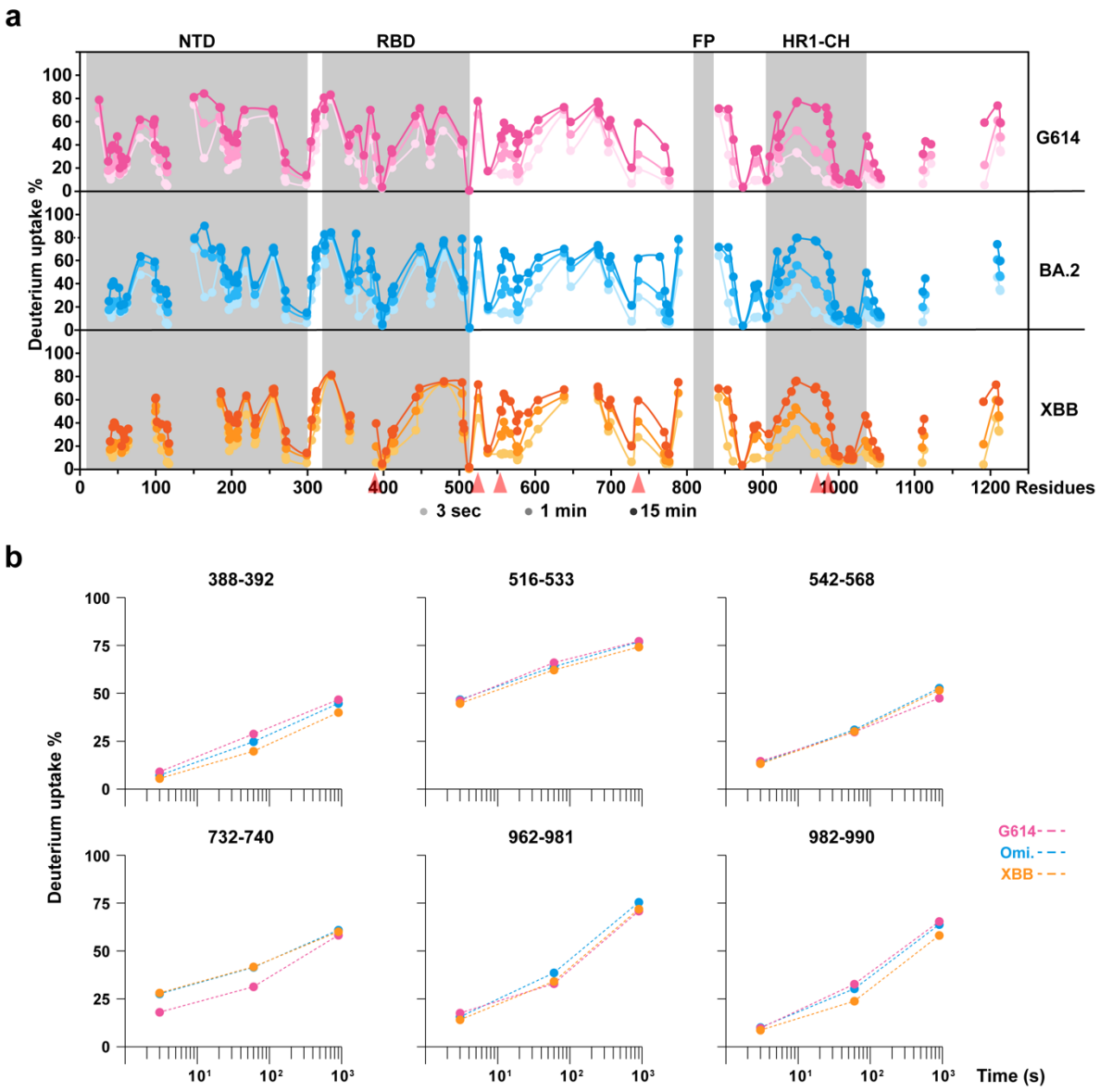

Figure 2. **Dynamic comparison among G614, Omicron and XBB S-6P homotrimers. (a)** Butterfly plots and **(b)** zoom-in homologous peptide comparisons in the region exhibiting major differences, reveal the strain-specific dynamic profile. Peptide locations are indicated by red arrows in panel a.

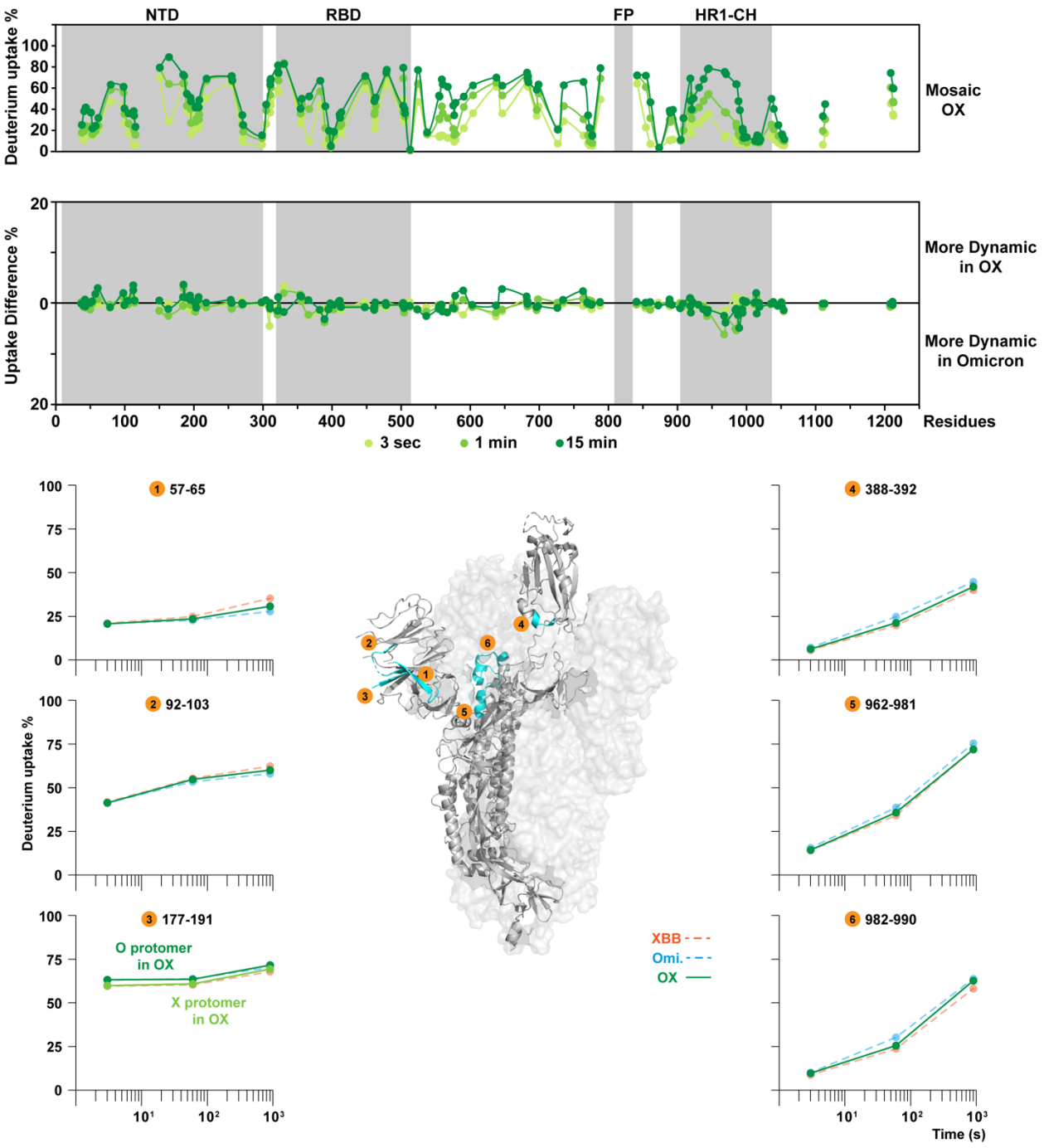

Figure 3. **Dynamic differences upon forming mosaic Omicron-XBB heterotrimers.** Peptide 1-3 featuring NTD, reveal the NTD dynamics are retained as in homotrimers. In peptide 3, the darker green plot indicates the behavior of Omicron protomer in the OX mosaic trimer while the lighter

green plot indicates the behavior of XBB protomer in the OX mosaic trimer. Peptide 4-6 showed the dynamic impacts on spike conformational states, where OX has the intermediate uptake level.

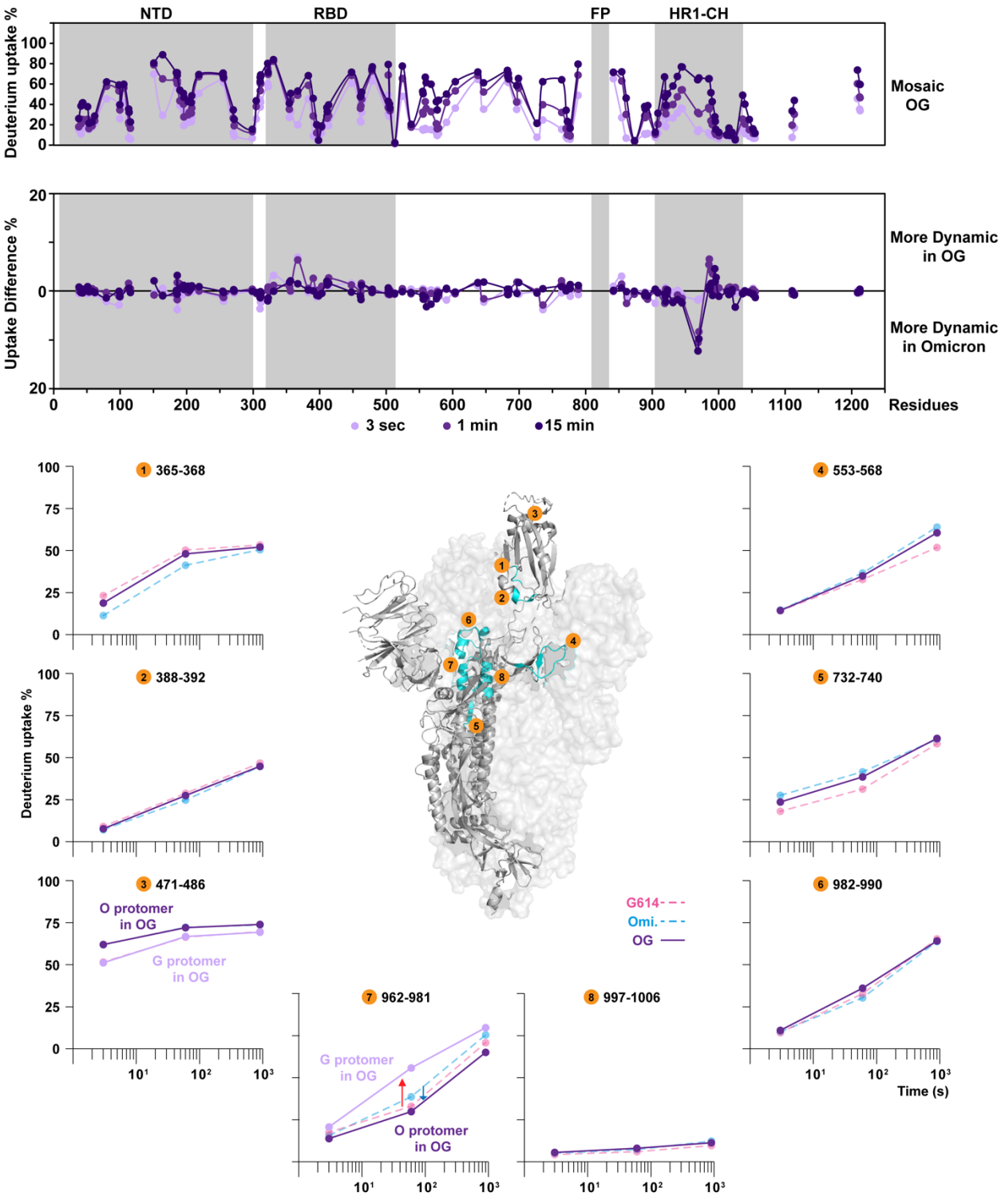

Figure 4. **Dynamic differences upon forming mosaic Omicron-G614 heterotrimers.** Peptide 1-3 featuring RBD, reveal the RBD dynamics are retained as in homotrimer. In peptide 3, the

darker purple plot indicates the behavior of Omicron protomer in the OG mosaic trimer while the lighter purple plot indicates the behavior of G614 protomer in the OG mosaic trimer. Peptide 4 is on the S1 hinge region, interacting with neighboring NTD, shows similar dynamic profile to Omicron. Peptide 5-7 indicate the increasing dynamics over FPPR in the OG heterotrimer. In peptide 7, the darker purple plot indicates the behavior of Omicron protomer in the OG mosaic trimer while the lighter purple plot indicates the behavior of G614 protomer in the OG mosaic trimer. Peptide 8 confirms the spike integrity reflected by the well-protected central helix core.

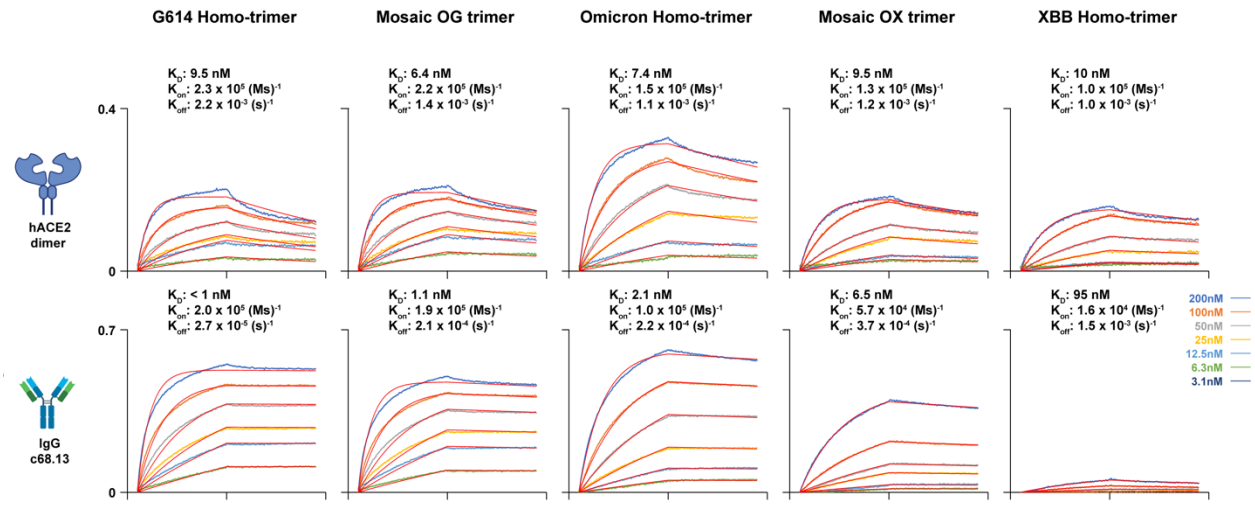

Figure 5. **RBD structural orders of mosaic S-6P heterotrimers.** BLI measures K_D on five spike trimers interacting with dimeric hACE and monoclonal antibody c68.13. Binding curves are aligned at the start of association step and fitted globally with a 1:1 ligand binding model.

Supplemental Information

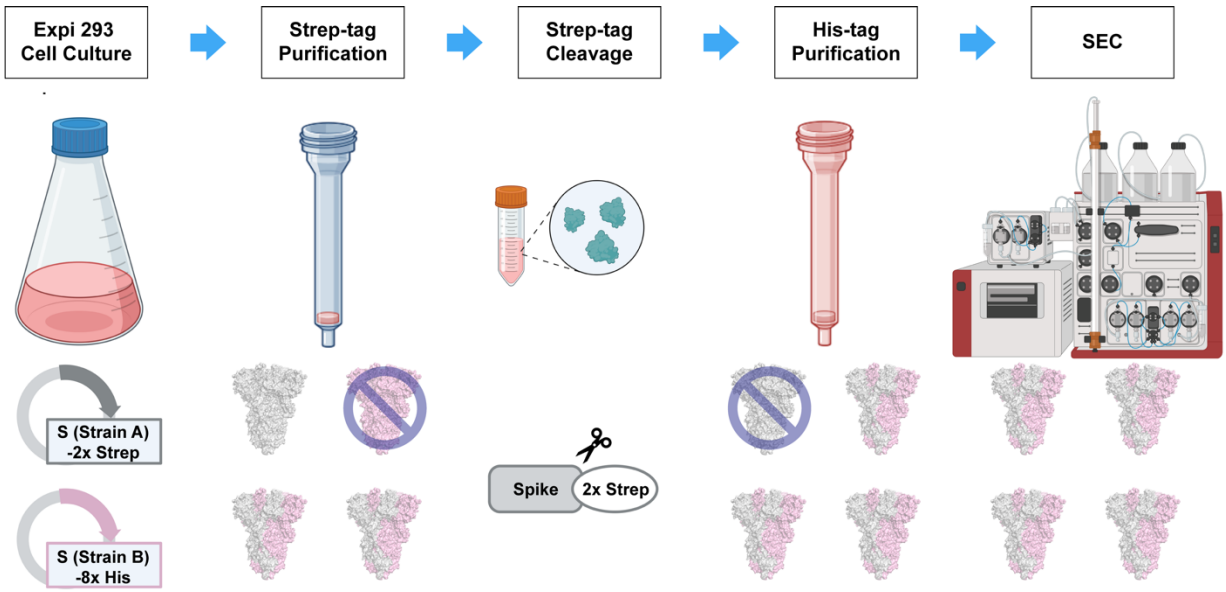

Figure S1. **Scheme of mosaic heterotrimer purification process.** Strep-tag affinity purification removes homotrimer of construct with His-tag. His-tag affinity purification continues removing the homotrimer of construct originally with Strep-tag.

Table S1. Experimental yields in mg of five protein preparations from 1 liter of Expi293F cell culture.

Spike	1st Strep	3C & 2nd Strep	His*
G614	-	-	6.8[II]
Omicron	4.4[I], 10.0[II]	2.3[I], 7.7[II]	-
XBB	-	-	5.8[II]
OG	4.4[I], 10.8[II]	2.5[I], 8.4[II]	(1.5 + 0.5)[I] (4.7 + 1.9)[II]
OX	3.4[I], 9.6[II]	1.7[I], 6.0[II]	(1.2 + 0.4)[I] (3.6 + 1.2)[II]

Protein constructs used in this chapter are only from single or two batches of preparations after purification protocol was optimized.

[I] Parallel preparations of Omicron, OG and OX from Expi293F cell (Sep-Nov 2023).

[II] Parallel preparations of G614, Omicron, XBB, OG and OX from Expi293F cell (Jan-Feb 2024).

* Yields of OG and OX are shown by mosaic heterotrimer mass from His-tag affinity purification elution + Omicron homotrimer mass from flow-through.

Table S2. DLS measurements of five spike trimers

Spike	Radius (nm)	PD (%)
G614	10.7±0.1	14.1±2.1
Omicron	11.7±0.1	17.1±2.4
XBB	11.6±0.1	15.1±1.8
OG	10.8±0.1	17.7±3.3
OX	10.0±0.2	11.8±3.8

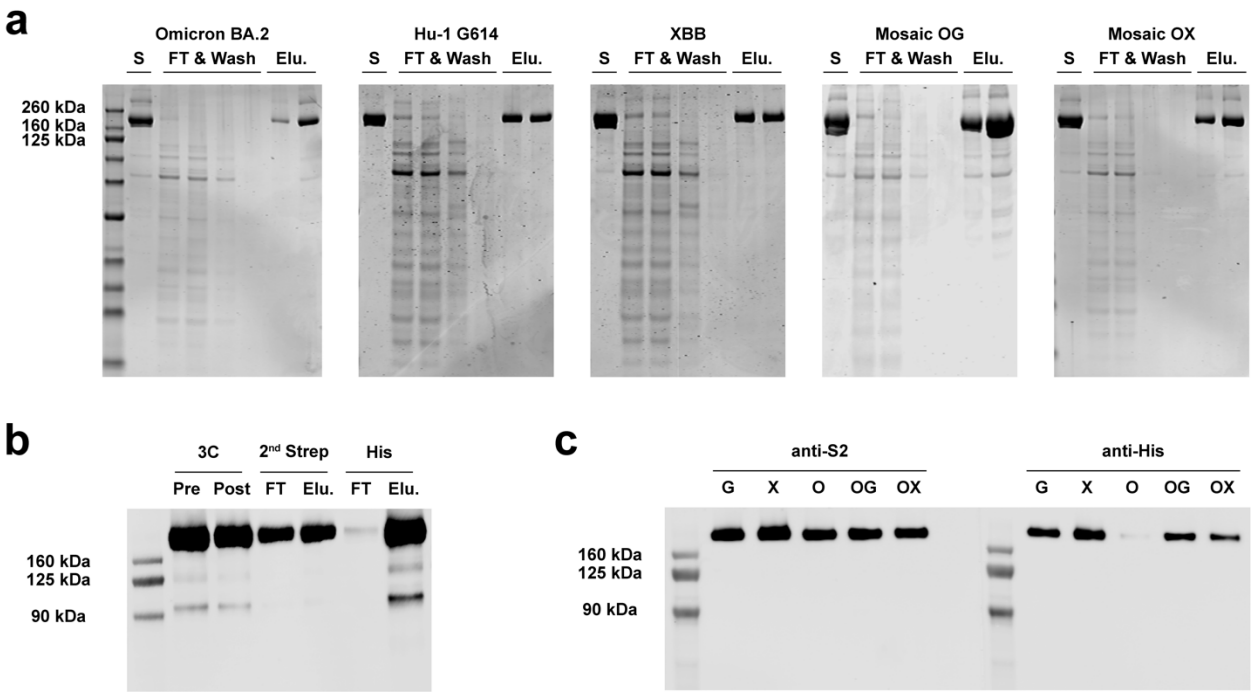

Figure S2. **Mosaic S-6P heterotrimer purification. (a)** SDS-PAGE gel with Coomassie blue staining on affinity purification of each spike trimer. **(b)** Western blot with anti-S2 primary antibody indicated the efficacy and purity of 3C protease digestion and tandem-affinity purification in mosaic heterotrimer purification. **(c)** Western blot characterization of homotrimers and heterotrimers using anti-S2 and anti-His primary antibodies. This figure is related to Figure 1.

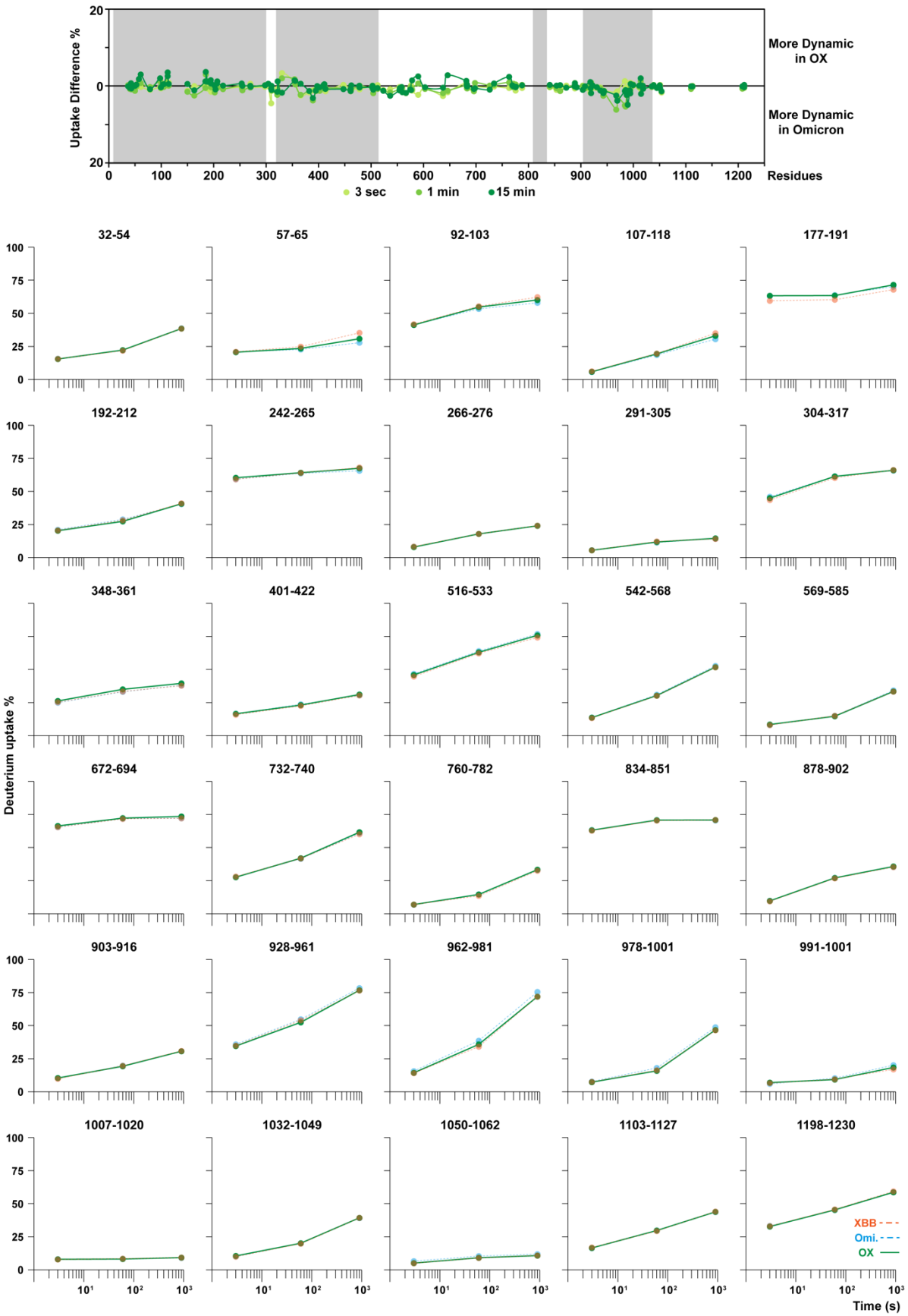
Uptake Difference %
20
0
20
More Dynamic in OX
More Dynamic in Omicron
0 100 200 300 400 500 600 700 800 900 1000 1100 1200
Residues
3 sec
1 min
15 min
Deuterium uptake %
32-54
57-65
92-103
107-118
177-191
192-212
242-265
266-276
291-305
304-317
348-361
401-422
516-533
542-568
569-585
672-694
732-740
760-782
834-851
878-902
903-916
928-961
962-981
978-1001
991-1001
1007-1020
1032-1049
1050-1062
1103-1127
1198-1230
100
75
50
25
0
10^1 10^2 10^3
XBB
Omi.
OX
Time (s)

Figure S3. **Deuterium uptake plots of mosaic Omicron-XBB S-6P.** This figure is related to Figure 3.

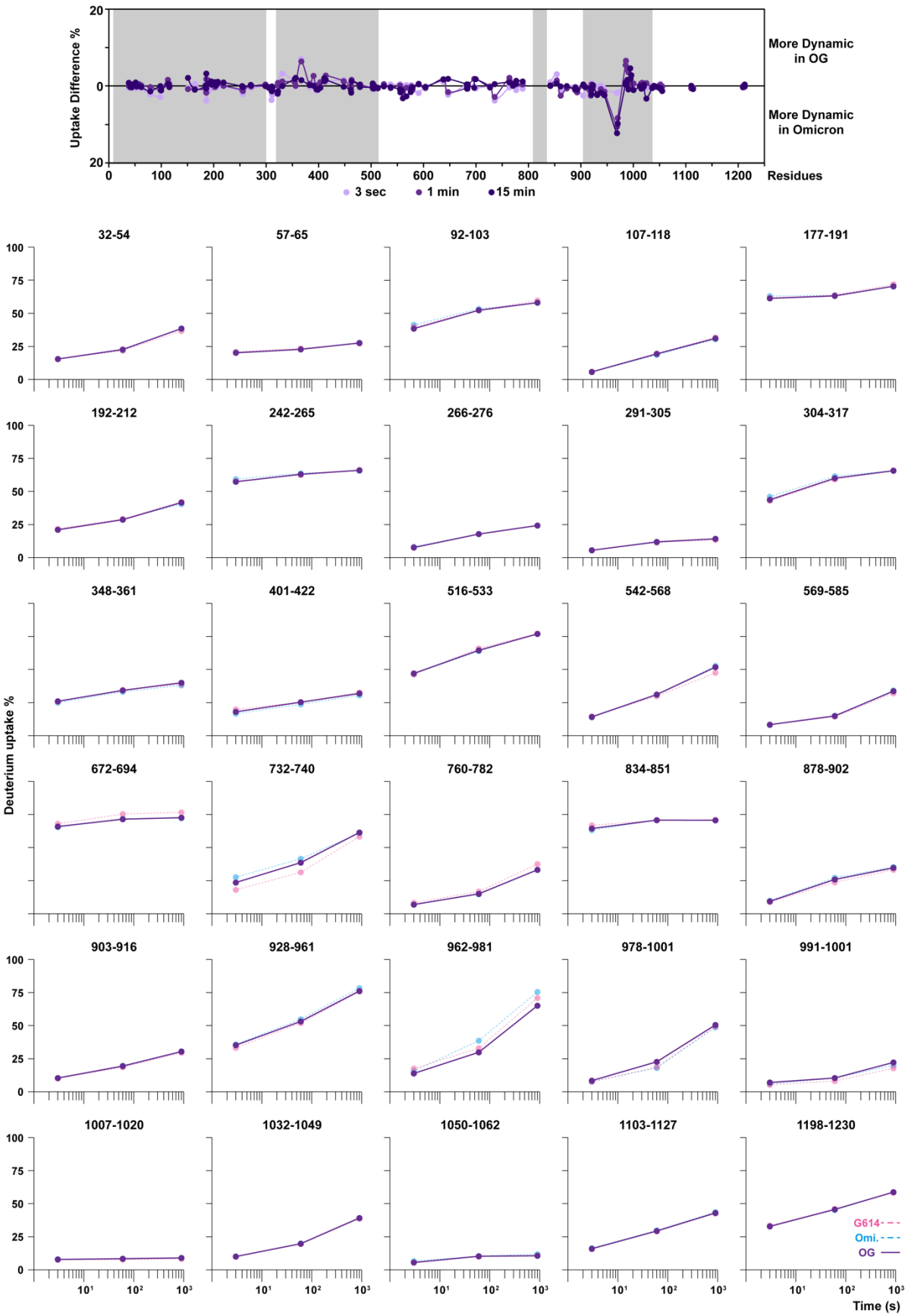
Uptake Difference %
20
0
20
More Dynamic in OG
More Dynamic in Omicron
0 100 200 300 400 500 600 700 800 900 1000 1100 1200
Residues
3 sec
1 min
15 min
32-54
57-65
92-103
107-118
177-191
192-212
242-265
266-276
291-305
304-317
348-361
401-422
516-533
542-568
569-585
672-694
732-740
760-782
834-851
878-902
903-916
928-961
962-981
978-1001
991-1001
1007-1020
1032-1049
1050-1062
1103-1127
1198-1230
Deuterium uptake %
100
75
50
25
0
10^1 10^2 10^3
Time (s)
G614
Omi.
OG

Figure S4. **Deuterium uptake plots of mosaic Omicron-G614 S-6P.** This figure is related to Figure 4.

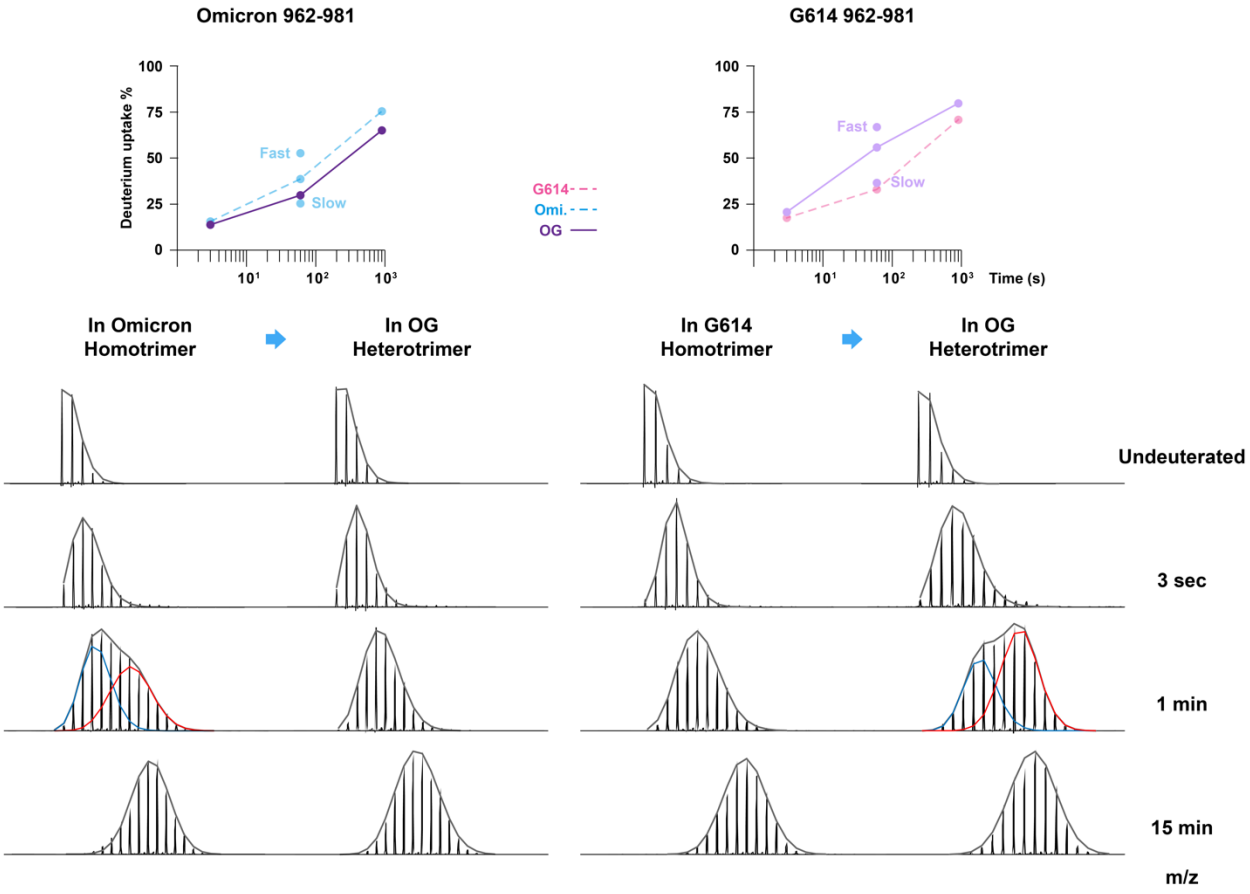

Figure S5. **Protomer-specific dynamics near S2 subunit apex in the Omicron-G614 heterotrimers.** Peptides 962-981 containing a N969K mutation in Omicron construct is used to differentiate deuterium uptake profiles between Omicron and G614 protomers. Dynamic changes in Omicron protomers (left panel) and G614 protomers (right panel) from their homotrimers to OG heterotrimer are compared. Fast and slow uptake levels labeled on the plots, correspond to red (fast) and blue (slow) populations of the bimodal mass envelopes at 1-minute time point.

References

(1) Harvey, W. T.; Carabelli, A. M.; Jackson, B.; Gupta, R. K.; Thomson, E. C.; Harrison, E. M.; Ludden, C.; Reeve, R.; Rambaut, A.; Peacock, S. J.; et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology* **2021**, *19* (7), 409-424. DOI: 10.1038/s41579-021-00573-0.

(2) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367* (6483), 1260-1263. DOI: 10.1126/science.abb2507.

(3) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e286. DOI: 10.1016/j.cell.2020.02.058.

(4) Gussow, A. B.; Auslander, N.; Faure, G.; Wolf, Y. I.; Zhang, F.; Koonin, E. V. Genomic determinants of pathogenicity in SARS-CoV-2 and other human coronaviruses. *Proceedings of the National Academy of Sciences* **2020**, *117* (26), 15193-15199. DOI: doi:10.1073/pnas.2008176117.

(5) Moeller, N. H.; Shi, K.; Demir, Ö.; Belica, C.; Banerjee, S.; Yin, L.; Durfee, C.; Amaro, R. E.; Aihara, H. Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN. *Proceedings of the National Academy of Sciences* **2022**, *119* (9), e2106379119. DOI: doi:10.1073/pnas.2106379119.

(6) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **2020**, *182* (4), 812-827.e819. DOI: 10.1016/j.cell.2020.06.043.

(7) 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.; et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **2020**, *588* (7839), 682-687. DOI: 10.1038/s41586-020-2852-1.

(8) 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. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **2020**, *584* (7821), 437-442. DOI: 10.1038/s41586-020-2456-9.

(9) Khan, A.; Zia, T.; Suleman, M.; Khan, T.; Ali, S. S.; Abbasi, A. A.; Mohammad, A.; Wei, D. Q. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *J Cell Physiol* **2021**, *236* (10), 7045-7057. DOI: 10.1002/jcp.30367.

(10) Volz, E.; Mishra, S.; Chand, M.; Barrett, J. C.; Johnson, R.; Geidelberg, L.; Hinsley, W. R.; Laydon, D. J.; Dabrera, G.; O'Toole, Á.; et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* **2021**, *593* (7858), 266-269. DOI: 10.1038/s41586-021-03470-x.

(11) Chen, J.; Wang, R.; Gilby, N. B.; Wei, G. W. Omicron Variant (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance. *J Chem Inf Model* **2022**, *62* (2), 412-422. DOI: 10.1021/acs.jcim.1c01451.

(12) Chen, J.; Wei, G.-W. Omicron BA.2 (B.1.1.529.2): High Potential for Becoming the Next Dominant Variant. *The Journal of Physical Chemistry Letters* **2022**, *13* (17), 3840-3849. DOI: 10.1021/acs.jpclett.2c00469.

(13) Tamura, T.; Ito, J.; Uriu, K.; Zahradnik, J.; Kida, I.; Anraku, Y.; Nasser, H.; Shofa, M.; Oda, Y.; Lytras, S.; et al. Virological characteristics of the SARS-CoV-2 XBB variant derived from

recombination of two Omicron subvariants. *Nat Commun* **2023**, *14* (1), 2800. DOI: 10.1038/s41467-023-38435-3.

(14) He, Q.; Wu, L.; Xu, Z.; Wang, X.; Xie, Y.; Chai, Y.; Zheng, A.; Zhou, J.; Qiao, S.; Huang, M.; et al. An updated atlas of antibody evasion by SARS-CoV-2 Omicron sub-variants including BQ.1.1 and XBB. *Cell Rep Med* **2023**, *4* (4), 100991. DOI: 10.1016/j.xcrm.2023.100991.

(15) Wang, Q.; Iketani, S.; Li, Z.; Liu, L.; Guo, Y.; Huang, Y.; Bowen, A. D.; Liu, M.; Wang, M.; Yu, J.; et al. Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* **2023**, *186* (2), 279-286.e278. DOI: 10.1016/j.cell.2022.12.018.

(16) Uraki, R.; Ito, M.; Furusawa, Y.; Yamayoshi, S.; Iwatsuki-Horimoto, K.; Adachi, E.; Saito, M.; Koga, M.; Tsutsumi, T.; Yamamoto, S.; et al. Humoral immune evasion of the omicron subvariants BQ.1.1 and XBB. *Lancet Infect Dis* **2023**, *23* (1), 30-32. DOI: 10.1016/s1473-3099(22)00816-7.

(17) Yue, C.; Song, W.; Wang, L.; Jian, F.; Chen, X.; Gao, F.; Shen, Z.; Wang, Y.; Wang, X.; Cao, Y. ACE2 binding and antibody evasion in enhanced transmissibility of XBB.1.5. *Lancet Infect Dis* **2023**, *23* (3), 278-280. DOI: 10.1016/s1473-3099(23)00010-5.

(18) Polack, F. P.; Thomas, S. J.; Kitchin, N.; Absalon, J.; Gurtman, A.; Lockhart, S.; Perez, J. L.; Marc, G. P.; Moreira, E. D.; Zerbini, C.; et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **2020**, *383* (27), 2603-2615. DOI: doi:10.1056/NEJMoa2034577.

(19) 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.; et al. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* **2020**, *586* (7830), 567-571. DOI: 10.1038/s41586-020-2622-0.

(20) Martínez-Flores, D.; Zepeda-Cervantes, J.; Cruz-Reséndiz, A.; Aguirre-Sampieri, S.; Sampieri, A.; Vaca, L. SARS-CoV-2 Vaccines Based on the Spike Glycoprotein and Implications of New Viral Variants. *Front Immunol* **2021**, *12*, 701501. DOI: 10.3389/fimmu.2021.701501.

(21) Baden, L. R.; Sahly, H. M. E.; Essink, B.; Kotloff, K.; Frey, S.; Novak, R.; Diemert, D.; Spector, S. A.; Rouphael, N.; Creech, C. B.; et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine* **2021**, *384* (5), 403-416. DOI: doi:10.1056/NEJMoa2035389.

(22) Yang, J.; Wang, W.; Chen, Z.; Lu, S.; Yang, F.; Bi, Z.; Bao, L.; Mo, F.; Li, X.; Huang, Y.; et al. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature* **2020**, *586* (7830), 572-577. DOI: 10.1038/s41586-020-2599-8.

(23) 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.; et al. Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell* **2020**, *183* (5), 1367-1382.e1317. DOI: 10.1016/j.cell.2020.10.043.

(24) Pardi, N.; Hogan, M. J.; Porter, F. W.; Weissman, D. mRNA vaccines — a new era in vaccinology. *Nature Reviews Drug Discovery* **2018**, *17* (4), 261-279. DOI: 10.1038/nrd.2017.243.

(25) Nance, K. D.; Meier, J. L. Modifications in an Emergency: The Role of N1-Methylpseudouridine in COVID-19 Vaccines. *ACS Central Science* **2021**, *7* (5), 748-756. DOI: 10.1021/acscentsci.1c00197.

(26) Tseng, H. F.; Ackerson, B. K.; Sy, L. S.; Tubert, J. E.; Luo, Y.; Qiu, S.; Lee, G. S.; Bruxvoort, K. J.; Ku, J. H.; Florea, A.; et al. mRNA-1273 bivalent (original and Omicron) COVID-19 vaccine effectiveness against COVID-19 outcomes in the United States. *Nat Commun* **2023**, *14* (1), 5851. DOI: 10.1038/s41467-023-41537-7.

(27) 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. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **2020**, *369* (6510), 1501-1505. DOI: doi:10.1126/science.abd0826.

(28) Watson, M. J.; Harkewicz, R.; Hodge, E. A.; Vorauer, C.; Palmer, J.; Lee, K. K.; Guttman, M. Simple Platform for Automating Decoupled LC-MS Analysis of Hydrogen/Deuterium Exchange Samples. *J Am Soc Mass Spectrom* **2021**, *32* (2), 597-600. DOI: 10.1021/jasms.0c00341.

(29) Calvaresi, V.; Wrobel, A. G.; Toporowska, J.; Hammerschmid, D.; Doores, K. J.; Bradshaw, R. T.; Parsons, R. B.; Benton, D. J.; Roustan, C.; Reading, E.; et al. Structural dynamics in the evolution of SARS-CoV-2 spike glycoprotein. *Nat Commun* **2023**, *14* (1), 1421. DOI: 10.1038/s41467-023-36745-0.

(30) Guenthoer, J.; Lilly, M.; Starr, T. N.; Dadonaite, B.; Lovendahl, K. N.; Croft, J. T.; Stoddard, C. I.; Chohan, V.; Ding, S.; Ruiz, F.; et al. Identification of broad, potent antibodies to functionally constrained regions of SARS-CoV-2 spike following a breakthrough infection. *Proceedings of the National Academy of Sciences* **2023**, *120* (23), e2220948120. DOI: doi:10.1073/pnas.2220948120.

(31) Dosey, A.; Ellis, D.; Boyoglu-Barnum, S.; Syeda, H.; Saunders, M.; Watson, M. J.; Kraft, J. C.; Pham, M. N.; Guttman, M.; Lee, K. K.; et al. Combinatorial immune refocusing within the influenza hemagglutinin RBD improves cross-neutralizing antibody responses. *Cell Rep* **2023**, *42* (12), 113553. DOI: 10.1016/j.celrep.2023.113553.

Chapter 5. Summary and Future Directions

5.1 Summary

This thesis work illustrates the dynamic impacts on SARS-CoV-2 spike's structural conformations, activations and antigenicity, as revealed by HDX-MS analysis from various controlled experiments. The sudden prevalence of the SARS-CoV-2 virus and the research interests in spike glycoprotein have contributed to numerous cryo-EM studies on spike structures. Due to its associations with viral membrane fusion functions and the dynamic nature of Class I fusion proteins, the more protein dynamics analysis on spike protein are needed to gain insights on the biological relevance beyond static structures.

In Chapter 1, the basic studies on spike structure organization and functions are reviewed, revealing the biological process on spike activation is significantly affected by differences in spike dynamics, a topic that has not been thoroughly studied. The biophysics and recent applications of HDX-MS are discussed, demonstrating the feasibility and advantages of HDX-MS for our project's needs.

In chapter 2-4, three major impacts on spike dynamics: structural conformations (Chapter 2), activation mechanisms (Chapter 3) and antigenicity (Chapter 4), are rigorously investigated and discussed. While only spike in open conformations is capable of receptor binding, spike variants show different dynamic sampling preferences on conformations. The D614G mutation dynamically stabilizes spike S2 subunits, significantly increases the preferences for open conformations. The characterization of open states spike is also discussed in SARS-CoV (Chapter 3) and later SARS-CoV-2 variants Omicron and XBB (Chapter 4), highlighting strain-specific behavior. The conservation of N-glycosylation at spike RBD N343, is shown to facilitate RBD

conformational shifts, further demonstrates dynamic motions of various domains in spike supporting RBD conformational shifts. Differences in RBD conformational preferences are proposed to be induced by the sequence-specific S2 stabilization and structural organization that influences the packing and relative dynamic motion of NTD, which crosstalk with RBD through interaction with the N343 glycan.

While the ratio of spike in open states affect the chances of hACE2 receptor binding, the differences in hACE2-induced spike activation also impact the subsequent fusion functions. In SARS-CoV-2 spike, hACE2 binding induces major allosteric effects spanning the spike domains. Increased dynamics are transduced from the RBD to the spike trimer central helices, extending through the helices and loops in HR1 domain towards the FPPR near the fusion peptide. This likely facilitates TMPRSS2 processing on the fusion peptide during this activation. In contrast, SARS-CoV spike, despite binding with hACE2, does not show prominent dynamics increases in the FPPR, suggesting less enhancement in hACE2-binding activation.

In Chapter 4, the dynamics and antigenicity of various VOC spike trimers are examined. The mosaic spike heterotrimer is expressed from co-transfection to mimic the possible outcome of bivalent mRNA vaccination. Both genetically close strains Omicron and XBB and distant strains Omicron and Wuhan-Hu-1 G614, which resembles the formulation in mRNA vaccines, can form mosaic heterotrimer without significant misfolding and aggregation. The mosaic spike trimer shows protomer-specific behaviors similar to their corresponding variants, while additional mosaic-specific dynamics are also observed. Mosaic trimer formed from Omicron and Wuhan-Hu-1 G614, containing differences in amino acid sequences in S2 subunit, show prominent dynamic changes near FPPR. The dynamics in G614 protomer is significantly enhanced at the cost of reduced dynamics in Omicron protomer, resulting in the OG mosaic heterotrimer possessing

more open state conformations like G614 variant as well as highly dynamic fusion peptide region similar to Omicron variant's phenotype. In both OG and OX mosaic formations, the hACE2 binding and antigenicity on the RBD are not impaired, as suggested by HDX-MS and direct BLI binding affinity measurements with dimeric hACE2 and specific RBD-targeting antibodies.

Overall, this thesis project demonstrates the advantages of HDX-MS in investigating the highly dynamic SARS-CoV-2 spike fusion protein. The versatility of HDX-MS specifies various types of S dynamic motions including local structural dynamics, conformational changes and allosteric effects. This work provides a structural dynamic basis for understanding receptor-induced conformational changes during activation as well as the impact on antigenicity. The findings support our hypothesis on the impact of structural dynamic differences on variant phenotypes in membrane fusion functions, and further provide insights into structure-based vaccine design, highlighting the effects of multivalent display.

5.2 Future Directions

Structural characterization of TMPRSS2-spike complex

In Chapter 3, we discussed the activation mechanism of spike protein upon binding to hACE2 receptor. For the spike to fully express its biological function in membrane fusion, it requires TMPRSS2 cleavage at S2' site, which exposes fusion peptide necessary for membrane insertion.[1] Many structural studies, published around the same time as ours, have independently observed the effects on the spike fusion peptide (FP) region upon hACE2 induced activation, supporting our results.[2-4] However, few studies have provided structural interpretations of how TMPRSS2 processes S upon recognizing the FP. Some studies, like ours, performed TMPRSS2

ectodomain cleavage assays on S under various conditions, revealing that TMPRSS2 has proteolytic activities at several sites on S, including the RBD, S1/S2 site and S2' site.[5]

As a type 2 transmembrane serine protease (TTSP), TMPRSS2 glycoprotein is expressed as a single peptide chain zymogen and undergoes activation by auto-cleavage at I256/V257. A crystal structure of TMPRSS2 has been reported in complex with the Nafamostat inhibitor.[5] Due to the auto-cleavage mechanism of this zymogen protease, structural studies featuring TMPRSS2-target interaction require the activated TMPRSS2 form for target binding.[2] While the soluble activated TMPRSS2 ectodomain shows functional digestion of spike protein, its activity can be overwhelming and the specificity at cleavage site can be largely affected, limiting the structural study of TMPRSS2 in complex with protein targets such as spike proteins.

Understanding the spike activation mechanism with TMPRSS2 digestion as a downstream effect of hACE2 binding is essential. Investigating how differences in dynamics upon hACE2 activation result in TMPRSS2 recognition orientation and digestion efficiency is a particularly interesting, related to our dynamic study in Chapter 3. As one future direction in studying the basic science of SARS-CoV-2 spike activation and its biological function in membrane fusion, we plan to study the structural and dynamic impacts of the interaction between TMPRSS2 and SARS-CoV-2 S, aiming to determine TMPRSS2-spike complex structures.

To facilitate this study, an activated stable but proteolytic function impaired TMPRSS2 will be required. This can be achieved through varying reaction conditions and protein engineering strategies. One direct method is to silence the possible cleavage site on spike RBD and S1/S2, leaving the S2' as the only recognition site, and to set up the TMPRSS2 binding at lower temperature and less enzyme units. Some possible TMPRSS2 engineering directions have been suggested by other studies, including typical mutations at enzyme catalytic triad H296-D345-S441.

A recently posted pre-print showed interacting between HKU1B spike RBD and S441A mutated TMPRSS2.[6] Experiments involving different reaction conditions and/or engineered proteins could be characterized by SDS-PAGE and native PAGE gel electrophoresis to check the interaction and peptide pattern from protein digestion, providing insights into TMPRSS2 activity and specificity. Subsequent structural studies using cryo-EM and HDX-MS would provide the structural information on the TMPRSS2-spike interaction, such as accessibility, TMPRSS2 orientation, and structural features at the binding interface.

We anticipate that this future direction will comprehensively enhance our understanding of the basic sciences related to how spike activating mechanism impacts its fusion machinery. It will reinforce our HDX-MS studies on the strain specific dynamics affecting spike functions over its evolution and providing insights into possible vaccine and antivirals development.

Serum polyclonal antibody epitope mapping

Besides the direction focusing on understanding basic biological questions on TMPRSS2-spike interactions, another future direction from my thesis work is more applied: developing HDX-MS-based serum polyclonal antibody epitope mapping.

Serum antibody mapping is an important approach to understand the immune response generated against pathogens like SARS-CoV-2 and influenza. This method involves identifying and characterizing the epitopes on an antigen that are recognized by antibodies in the serum of individuals exposed to the pathogens or vaccinated.

One significant advantage of serum antibody mapping is its ability to provide a comprehensive overview of both antigen's structural features and antibody sequence information. By mapping these interactions, we can gain insights into the immune system response to specific

antigens and obtain structural and immunological information essential for developing broad-spectrum vaccines or antiviral treatments.

Polyclonal antibody epitope mapping has been developed from various techniques. HDX-MS, powerful in detecting local structural orderings from antibody binding, has been employed in numerous studies.[7] Another advanced technique for serum antibody mapping is electron microscopy polyclonal epitope mapping (EMPEM).[8, 9] EMPEM allows for high-resolution visualization of antibody-antigen interactions, revealing detailed structural information about how antibodies bind to their epitopes. EMPEM study on the diversity of human antibody responses to H5N1 vaccination, demonstrates its capability to capture comprehensive maps of immune responses.[10]

The progress made from these studies has raised the potential for simultaneous sequencing and mapping of polyclonal antibodies directly from serum. While EMPEM technique can reconstruct 3D structures of antibody-antigen complexes, the resolution is a limiting factor of determining accurate antibody sequence. Therefore, various studies have coupled cryo-EM images with mass spectrometry proteomic data to enhance the resolution and accuracy of epitope mapping.[11]

Based on previous studies showing that HDX-MS aids in antibody mapping, one direction for serum polyclonal antibody mapping is to utilize HDX-MS for both epitope identification and antibody sequence determination. Since HDX-MS can be conducted under physiological conditions, it can reflect both native interactions and structural dynamics. Optimization of HDX-MS in serum samples is urgently needed to avoid non-specific binding and enhance detectable signals. In the SARS-CoV-2 case, since over 90% of antibody responses target the RBD or NTD domain,[12, 13] which are susceptive to viral adaptation for antibody escape, filtering out RBD and

NTD targeting antibodies could potentially enrich antibodies with novel epitopes. Identifying such novel epitopes, coupled with subsequent HDX-MS analysis on the dynamic impacts of antigen-antibody interactions, could provide valuable information for broader-spectrum vaccine design or antibody therapy.

Another challenge of HDX-MS-based antibody epitope mapping is the lack of cryo-EM images, making antibody sequence determination largely dependent on digestion efficiency and MS sensitivity. To better facilitate this sequence-to-structure prediction, coupling HDX-MS with computational *de novo* tools like RoseTTAFold and AlphaFold could be a viable solution.

Utilizing HDX-MS for serum polyclonal antibody mapping could help us understand the connection between immune responses and antigenicity, as well as bridging the advantages of structural mass spectrometry and computational structural prediction. This integration could significantly advance our knowledge of immune responses and improve strategies for vaccine and therapeutic antibody development.

References

(1) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181* (2), 271-280.e278. DOI: 10.1016/j.cell.2020.02.052.

(2) Chen, C.; Zhu, R.; Hodge, E. A.; Díaz-Salinas, M. A.; Nguyen, A.; Munro, J. B.; Lee, K. K. hACE2-Induced Allosteric Activation in SARS-CoV versus SARS-CoV-2 Spike Assemblies Revealed by Structural Dynamics. *ACS Infectious Diseases* **2023**, *9* (6), 1180-1189. DOI: 10.1021/acsinfecdis.3c00010.

(3) Calvaresi, V.; Wrobel, A. G.; Toporowska, J.; Hammerschmid, D.; Doores, K. J.; Bradshaw, R. T.; Parsons, R. B.; Benton, D. J.; Roustan, C.; Reading, E.; et al. Structural dynamics in the evolution of SARS-CoV-2 spike glycoprotein. *Nat Commun* **2023**, *14* (1), 1421. DOI: 10.1038/s41467-023-36745-0.

(4) Low, J. S.; Jerak, J.; Tortorici, M. A.; McCallum, M.; Pinto, D.; Cassotta, A.; Foglierini, M.; Mele, F.; Abdelnabi, R.; Weynand, B.; et al. ACE2-binding exposes the SARS-CoV-2 fusion peptide to broadly neutralizing coronavirus antibodies. *Science* **2022**, *377* (6607), 735-742. DOI: 10.1126/science.abq2679.

(5) Fraser, B. J.; Beldar, S.; Seitova, A.; Hutchinson, A.; Mannar, D.; Li, Y.; Kwon, D.; Tan, R.; Wilson, R. P.; Leopold, K.; et al. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat Chem Biol* **2022**, *18* (9), 963-971. DOI: 10.1038/s41589-022-01059-7.

(6) Fernández, I.; Saunders, N.; Duquerroy, S.; Bolland, W. H.; Arbabian, A.; Salazar, E. B.; Blanc, C.; Lafaye, P.; Haouz, A.; Buchrieser, J.; et al. Structural basis of TMPRSS2 zymogen activation

and recognition by the HKU1 seasonal coronavirus. *bioRxiv* **2024**, 2024.2002.2021.581378. DOI: 10.1101/2024.02.21.581378.

(7) Ständer, S.; R. Grauslund, L.; Scarselli, M.; Norais, N.; Rand, K. Epitope Mapping of Polyclonal Antibodies by Hydrogen–Deuterium Exchange Mass Spectrometry (HDX-MS). *Analytical Chemistry* **2021**, *93* (34), 11669-11678. DOI: 10.1021/acs.analchem.1c00696.

(8) Antanasijevic, A.; Bowman, C. A.; Kirchdoerfer, R. N.; Cottrell, C. A.; Ozorowski, G.; Upadhyay, A. A.; Cirelli, K. M.; Carnathan, D. G.; Enemuo, C. A.; Sewall, L. M.; et al. From structure to sequence: Antibody discovery using cryoEM. *Science Advances* **2022**, *8* (3), eabk2039. DOI: doi:10.1126/sciadv.abk2039.

(9) Guthmiller, J. J.; Han, J.; Utset, H. A.; Li, L.; Lan, L. Y.-L.; Henry, C.; Stamper, C. T.; McMahon, M.; O'Dell, G.; Fernández-Quintero, M. L.; et al. Broadly neutralizing antibodies target a haemagglutinin anchor epitope. *Nature* **2022**, *602* (7896), 314-320. DOI: 10.1038/s41586-021-04356-8.

(10) Han, J.; Schmitz, A. J.; Richey, S. T.; Dai, Y.-N.; Turner, H. L.; Mohammed, B. M.; Fremont, D. H.; Ellebedy, A. H.; Ward, A. B. Polyclonal epitope mapping reveals temporal dynamics and diversity of human antibody responses to H5N1 vaccination. *Cell Reports* **2021**, *34* (4), 108682. DOI: https://doi.org/10.1016/j.celrep.2020.108682.

(11) Schulte, D.; Šiborová, M.; Käll, L.; Snijder, J. Simultaneous polyclonal antibody sequencing and epitope mapping by cryo electron microscopy and mass spectrometry – a perspective. *bioRxiv* **2024**, 2024.2006.2021.600107. DOI: 10.1101/2024.06.21.600107.

(12) 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. Convergent antibody responses to SARS-CoV-

2 in convalescent individuals. *Nature* **2020**, *584* (7821), 437-442. DOI: 10.1038/s41586-020-2456-9.

(13) 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.; et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **2020**, *588* (7839), 682-687. DOI: 10.1038/s41586-020-2852-1.