Functional *De Novo* Proteins as Custom Molecular Tools for Bioengineering: from Targeted Cancer Immunotherapy to Modular Biosensors

Alfredo Quijano Rubio

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

David Baker, Chair

Andre Berndt

Hao Kueh

Gaurav Bhardwaj

Program Authorized to Offer Degree:

Bioengineering

University of Washington

Abstract

Functional *De Novo* Proteins as Custom Molecular Tools for Bioengineering: from Targeted Cancer Immunotherapy to Modular Biosensors

Alfredo Quijano Rubio

Chair of the Supervisory Committee:

David Baker

Department of Biochemistry

Traditional protein engineering methods use naturally existing proteins as a starting point and are intrinsically limited to small perturbations of a protein's original structure and function. Computational de novo protein design is free of such limitations and enables the creation of designer proteins with custom functions and ideal biochemical properties, which make excellent tools to address diverse bioengineering challenges. In this work, I focus on two applications. First, I describe the development of conditionally active cytokine mimics for targeted cancer immunotherapy. We take advantage of the robust folding of a de novo designed Interleukin-2 mimetic protein to design a split version consisting of two components that do not elicit signaling individually but regain activity when co-localized. With this approach, we aim to overcome the severe side-effects of systemic cytokine therapies. Second, I describe the design of modular and tunable biosensors from broadly applicable design principles by using a robust de novo designed protein switch. We demonstrate the modularity of this biosensor platform by designing functional biosensors for multiple analytes of interest.

Acknowledgements

I want to express my gratitude to the many people that have supported me in this adventure, my graduate studies in Bioengineering at the University of Washington. To David Baker, my thesis advisor, for his guidance and support to pursue my scientific curiosity. Furthermore, for encouraging (by example) scientific collaboration and creating a thriving lab environment, which he refers to as the "communal brain", where everyone's ideas are highly valued and interconnected. To all my colleagues at the Institute for Protein Design (staff, students, postdocs, faculty, and the Core) for making a place of utmost scientific excellence while keeping it fun and collaborative. In particular, to the colleagues with whom I have done research side-by-side: to the *de novo* protein mimetics team (Daniel-Adriano Silva, Umut Ulge, Carl Walkey, and Jung Ho Chun), the *de novo* biosensors team (Andy Yeh, Jooyoung Park, Byung-Ha Oh, Bobby Langan, Scott Boyken, Marc Lajoie, Cameron Chow, and Marcos Miranda) and the *de novo* protein logic group (a long list of people). Especial thanks to Daniel-Adriano Silva for his mentorship in protein design, experimental skills and for showing me how to change the world one protein at a time.

I also want to thank all the collaborators outside the Institute for Protein Design that have made this research possible. Prof. Michael Dougan, Prof. Stephanie Dougan, and their team (Aladdin Bhuiyan, Elisa Bello, Lestat Ali, and Kevin Zhangxu); Prof. Jamie Spangler and her team (Huilin Yang, Wentao Wang, and Yun-Huai Kuo); and Prof. Stanley Riddell and his team (Isabel Leung). It has been a privilege.

Thanks to "La Caixa Foundation" for their support in embarking me on this journey.

Finalmente, a mi padre, mi madre, mi hermana y al resto de mi familia por su apoyo incondicional desde la otra parte del mundo. Y a Camilla, por formar parte de esta aventura independientemente de la distancia y franja horaria.

Table of Contents

Chapter 1

Introduction: The advent of de novo proteins as tools for cancer immunotherapy and synthetic biology

Adapted from: Quijano-Rubio, A., Ulge, U. Y., Walkey, C. D., & Silva, D.-A. (2020). The advent of de novo proteins for cancer immunotherapy. Current Opinion in Chemical Biology, 56, 119–128.

Abstract

Traditional protein engineering methods use naturally existing proteins as a starting point and, therefore, are intrinsically limited to small perturbations of a protein's natural structure and function. Conversely, computational de novo protein design is free of such limitations since it can produce a virtually infinite number of novel protein sequences, folds, and functions. Recent advances in de novo protein design have enabled the creation of designer proteins with custom functions and ideal biochemical properties, which make excellent tools to address diverse bioengineering challenges. This chapter describes some of the characteristics that make de novo proteins promising candidates to create novel therapeutic and synthetic biology tools, such as their extreme stability, high binding affinity, mutational robustness, unique sequence and re-engineerability. These properties can help overcoming many of the challenges usually associated with protein drug development, such as instability, manufacturability or immunogenicity. While this chapter focuses primarily on the promise of de novo protein design to create novel therapeutic proteins, the unique properties of de novo proteins can also be exploited to solve challenges in other fields, such as the development of highly modular and tunable biosensors.

Introduction

During the last decade, immunotherapy has revolutionized cancer treatment and become the forefront of oncology [1,2]. This revolution stemmed from engineering efforts aimed at developing more effective (mostly protein-based) cancer therapeutics [3]. Remarkably, the advent of antibody-based checkpoint inhibitors (engineered antibodies that selectively inhibit immunosuppressive pathways in immune cells [4,5]) and Chimeric Antigen Receptors (engineered receptors that arm T cells against malignant cells [6,7]) have shown the potential of protein engineering to enable novel immunotherapeutic approaches. Protein engineering is spearheading immuno-oncology innovation, and new advances in this field may bring a revolution in medicine.

The development of protein immunotherapeutics has traditionally required the modification of naturally occurring proteins. Earlier immuno-oncology therapies relied on the inherent immunomodulatory properties of natural proteins, such as interferon-α [8] and interleukin-2 [9,10]. While effective to a certain degree, many of these therapies have resulted in severe adverse effects limiting their therapeutic use [11,12]. Since then, more advanced protein engineering techniques have been applied to improve the efficacy and reduce the toxicity of novel therapeutics. Notably, techniques such as directed evolution [13,14], chemical modification [15] and antibody/antibody-fragment engineering [16–18], have been pivotal for engineering highly specialized immunotherapeutic proteins. Some examples include bispecific antibodies [19], immunocytokines [20,21], bispecific T-cell engagers [22] and antibody-drug conjugates [23]. These advances have led to many clinical trials, and several approved drugs for use in humans [24].

Despite the success of traditional protein engineering approaches, there is still much room for improvement. The reliance on natural proteins as building blocks imposes several constraints in the design of novel therapeutics. For instance, the biochemical properties of the final products, such as stability, (fundamental) mode of action, and protein-sequence are inherently tied to the protein used as a starting point. Furthermore, natural proteins often only tolerate a limited number of modifications before they are rendered unstable or non-functional [25]. As a consequence, therapeutic proteins developed from natural proteins frequently exhibit poor stability and manufacturability, and in many cases, significant immunogenicity [26–28].

The recent advent of de novo protein design has enabled the creation of a new class of proteins (i.e., designer-proteins) that share no homology with natural ones [29]. De novo protein design creates proteins using computational algorithms, and it can access an exponentially larger protein-sequence space than what nature has sampled through evolution to find new sequences that could solve challenges unmet by natural proteins (Fig. 1) [29–31]. Algorithms used in de novo protein design apply our understanding of biophysics to create ideal protein structures [32], allowing us to confer them with unique and potentially superior biochemical properties and enhanced and/or novel functions [33,34]. De novo designed proteins offer an innovative way to overcome the inherent constraints of natural protein building blocks, opening the door to new tools with a wide range of applications from therapeutics and synthetic biology.

Exploring new functions through computational *de novo* protein design

The advent of computational protein design has been made possible by a combination of advances in computational power, gene synthesis technologies, and protein-structure prediction and design algorithms [29,33,35]. Rosetta is currently the leading software suite for protein modeling and design

[35]. Since initially developed more than two decades ago, it has continuously been improved by the scientific community, and new methods have been instrumental in enabling the rapid evolution of the field. The first applications of computational protein design focused on redesigning natural proteins, and its functional efforts aimed to solve a variety of challenges in biomedicine and synthetic biology [36,37]. Examples are the stabilization of proteins [38,39] for the development of more effective vaccines against malaria [40], enhanced enzymes to treat celiac disease [41], self-assembling nanoparticles for the presentation of immunogenic epitopes [42], and the optimization of numerous antibodies [43–47]. Nevertheless, these approaches still relied on the use of natural proteins as starting points, which limited the overall scope to the improvement (or adaptation) of pre-existing natural protein structures and functions.

Eventually, computational protein design methods advanced to a point where it became possible to create fully de novo proteins, free of the constraints imposed by evolution into natural proteins [29]. Initial efforts in de novo design focused on establishing the basic design principles that could allow the systematic generation of de novo proteins [35,48–50]. The great effort invested in early research paid off, not only by producing proteins with new folds but also by delivering novel proteins with remarkable stability [48]. Ultimately, the first examples of de novo designed proteins demonstrated that computational protein design can be applied to accurately engineer entirely new proteins with physicochemical properties that can (often) exceed those of naturally occurring proteins. Nevertheless, none of these de novo proteins was designed for (or capable of) accomplishing a biologically relevant function. Since then, a large number of de novo "protein scaffolds" (i.e. well-folded protein structures without function) spanning a wide range of shapes and sizes have been designed and confirmed experimentally [51–59]. Such a diverse repertoire of novel hyper-stable proteins can indeed also serve as starting points for the development of new functional proteins. For instance, introducing functional motifs into such inert de novo protein scaffolds (known as motif grafting) allowed protein designers to create the first examples of functional de novo proteins [60,61]. More recently, studies have demonstrated that this approach is widely applicable and that a broad array of functions can be grafted into suitable (previously inert) de novo proteins. Remarkable examples include the creation of thousands of high-affinity mini-proteins with (inhibitory) therapeutic potential against influenza and botulinum neurotoxin (Fig. 2A)[62] and bioactive protein switches with customizable cellular functions (Fig. 2D)[63]. These designed protein switches were subsequently used to develop designed protein logic for cell targeting [64] , biological feedback control [65], and modular biosensors (Chapter 3) showcasing their broad use in synthetic biology.

A second, more challenging approach, is creating functional de novo proteins that are tailored from their conception to achieve a given desired function. Simultaneously designing a protein-fold and its function can be more challenging than adapting/functionalizing an existing protein scaffold. Such a "purposefully-functional" de novo protein design requires a high level of control over the shape, size, and function of the final designs (as well as prior knowledge/or/assumptions about the structure of the targeted functional site). Thus, the de novo design of functional proteins remained an elusive landmark in the field. The earliest successes of this design paradigm include the design of an Epstein-Barr viral protein inhibitor [66], a designed scaffold for stabilizing a neutralizing epitope of the respiratory syncytial virus [67], and a transmembrane Zn^{2+}-transporting helical protein [68]. More recently, designers have achieved the development of pH-sensitive protein oligomers (Fig. 2C) [69]; a de novo designed fluorescent protein (Fig. 2B) [70]; and Neo-2/15, a de novo protein

with biologically-relevant immunomodulatory function (Fig. 2F) [71]. Altogether, these examples illustrate how advances in de novo computational protein design are enabling the creation of novel proteins with functions and (potential) applications in a wide range of scientific and technological fields.

De novo designed proteins exhibit exceptional robustness

Natural protein evolution, in most cases, prioritizes protein function over stability [25,72]. As a result, most natural proteins have structural irregularities, which translate in non-ideal properties that pose challenges for the development of therapeutic applications, such as low solubility and poor mutational robustness [25]. While these irregularities may not compromise the function of natural proteins in their natural context, they can severely hinder the process of repurposing them as therapeutic agents [28]. De novo proteins can overcome many of the limitations imposed by natural proteins because they can be designed to follow the principles of an idealized protein structure: regular secondary structure elements with short and structured connecting loops, well-packed hydrophobic cores, mostly polar exposed-surfaces, and highly satisfied polar interactions. Furthermore, current protein design methods excel at creating molecules with strong intramolecular interactions (in particular between consecutive residues), which frequently results in de novo proteins with unusually high intrinsic stability and folding propensity (when compared to natural proteins) [32,58,73]. As a result, de novo proteins with idealized structures regularly exhibit unusual/enhanced biochemical properties, such as extreme thermostability and mutational robustness [30,38,74].

Figure 3 illustrates some examples of the unusual thermostability of de novo functional proteins. Furthermore, it has been demonstrated that further enhancement of a de novo protein's stability is frequently an easily attainable task (due to the aforementioned mutational robustness). In general, the improvement (for therapeutic/application purposes) of de novo proteins can be approached computationally by a myriad of methods such as introducing further-stabilizing mutations or introducing designed chemical protein-staples (e.g. disulfide bridges) [54,62,71,75] (Fig. 3D). Additionally, it is within the realm of de novo proteins to easily introduce designed chemical-handles using natural-amino acids (e.g. cysteine residues), that for instance can be used to enhance the physical properties of the final conjugated molecule (Fig. 3E). Such modifications can, in general, be achieved without significantly affecting the protein's function.

In summary, de novo proteins can, by design, overcome many of the biochemical limitations inherent to natural proteins, thereby enabling the creation of potentially therapeutic proteins with highly desirable biochemical properties such as high-thermostability, manufacturability, solubility, and tunability.

De novo design can be used to build potent proteins for immunotherapy

We recently described the first example of a de novo protein immunotherapeutic: Neo-2/15, a de novo mimic of the function of both interleukin-2 (IL-2) and interleukin-15 (IL-15)[71]. Recombinant interleukin-2 (rIL-2), also known as aldesleukin, was the first immunotherapy approved for the treatment of melanoma and renal cell carcinoma [9], but its clinical use has been limited due to severe dose-limiting toxicities [11,76,77]. To accomplish its biological function, IL-2 induces the hetero-dimerization of two IL-2 cell-membrane receptors, namely the IL-2 beta (IL-2Rβ) and the common-gamma (γ_c or IL-2Rγ) receptors, also called the IL-2Rβγ complex. The dimerization of

IL-2Rβγ initiates a signaling cascade that ultimately stimulates immune cells (expressing the receptors), including those with anti-tumor function, such as naive effector T-cells and Natural killer cells (NK). However, many off-target cell types express a higher-affinity 3-component IL-2-receptor containing, in addition, the IL-2 alpha subunit (IL-2Rα, or CD25). Although CD25 is not directly involved in IL-2 signal transduction, its presence greatly enhances the affinity and activity of IL-2 (i.e., IL-2Rαβγ >> IL-2Rβγ, meaning that CD25 is a potent (allosteric) activator of IL-2). As a result, off-target CD25+ cells, including endothelial cells and immunosuppressive regulatory T-cells, respond to IL-2 therapy with much higher potency than the intended target cells (e.g. CD8+ naive T-cells). The strong bias towards the activation of cells expressing CD25 leads to severe side effects and a mixed, pleomorphic immune response [77,78].

Many efforts have been devoted to trying to improve the safety and efficacy of IL-2 by targeting to eliminate its natural binding preference for CD25. Nevertheless, all efforts have failed to completely abrogate CD25 binding (while preserving IL-2Rβγ binding), and most of these reengineered IL-2 proteins have simultaneously introduced significant challenges to the potency, manufacturability, and/or stability [79–81]. To solve this decades-long challenge, we built on current Rosetta design methodology (extensively described in these two reviews [32,33]) to develop a de novo computational and experimental protein design method to designing proteins capable of replicating the potent immunostimulatory function of IL-2 (and IL-15), while fully eliminating the binding site (and hence the dependency) for CD25 (or IL-15Rα, also known as CD215). Our scientific efforts resulted in the creation of a novel and robust de novo computational design method and a number of fully de novo IL-2/IL-15 protein-mimetics that are truly independent of CD25 and CD215. The best-optimized design, Neo-2/15 (Fig. 2F), is a highly thermostable (Fig. 3B), tunable (Fig 3C-E), easily manufactured protein that demonstrates potent signaling on both human and mouse cells, and exhibits anti-tumor activity in syngeneic mouse tumor models with less toxicity than recombinant IL-2 (rIL-2)[71].

Discussion, opportunities, and challenges

The exploration of novel protein functions through de novo design is enabling the development of new proteins with promising synthetic biology and therapeutic applications [62,71]. In just two decades, computational protein engineering has advanced from being able to optimize natural proteins, to creating entirely de novo proteins with extreme stability, a wide range of shapes, and custom functions. In contrast to current antibody-based therapies, de novo proteins can be small, thermally stable, easy to manufacture (usually in recombinant bacterial systems), and highly tunable. For the first time, we are now able to design custom proteins with a predefined functional purpose. The emerging understanding that we have acquired in creating de novo proteins is already allowing us to create a new class of custom-built proteins with unique (and potentially superior) biochemical properties. Furthermore, de novo design can enable more complex protein systems, such as conditionally-active signaling proteins (Chapter 2) or modular and tunable protein biosensors (Chapter 3). We anticipate that future iterations of these methods will be a powerful driver for the development of next-generation diagnostics and therapeutics. It is certainly an exciting time in de novo protein design! With the emergence of the first examples of biologically-relevant designed de novo proteins, it is clear that the field is coming of age.

Figures and figure legends

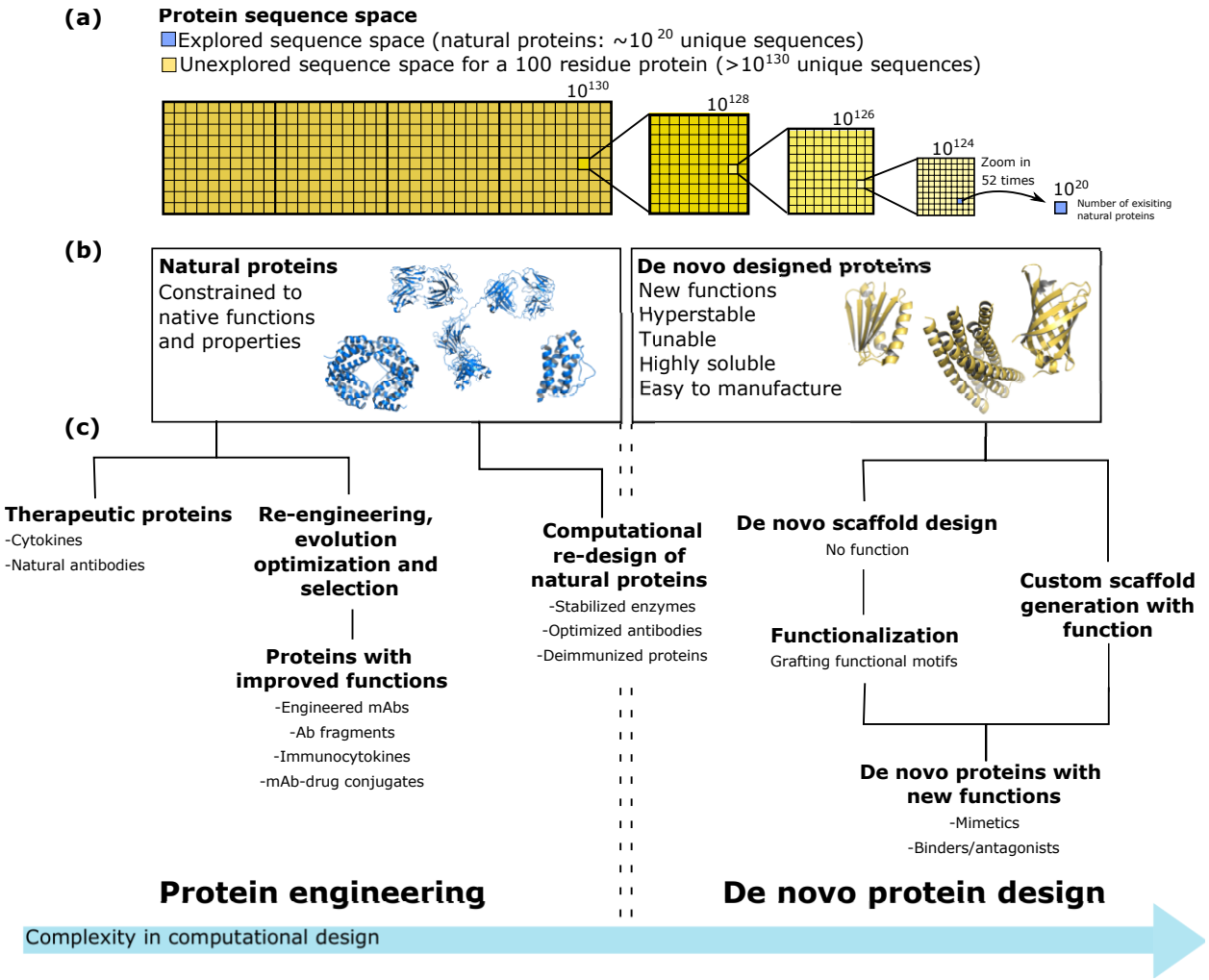


Figure 1. ***De novo*** **protein design enables the exploration of new protein sequences, structures, and functions. (a)** The vast number of sequences possible for proteins remains largely unexplored by nature (yellow squares). If we consider a protein of 100 residues, the possible number of sequences combining all 20 natural amino acids is approximately 10^{123}. However, the total number of proteins sampled through evolution and protein engineering is estimated to be on the order of 10^{20} different sequences (blue square)[30], which is a very small fraction of the available sequence space. **(b)** Representation of natural protein scaffolds (blue) and *de novo* designed protein scaffolds (yellow). Left panel: Interferon-γ (PDB ID: 1HIG), an antibody (PDB ID: 1IGT) and Interleukin-2 (PDB ID: 1QYS). Right panel: Top7 (PDB ID: 1HIG) [48], a *de novo* designed helical bundle oligomer (PDB ID: 6MSR) [53], and a *de novo* designed beta-barrel (PDB ID: 6D0T) [70]. **(c)** Advances in computational protein design methods are enabling the creation of *de novo* proteins with new functions. *De novo* design shows promise for solving current limitations of traditional protein engineering, enabling the creation of custom proteins for the next generation of immunotherapeutics and other medical areas.

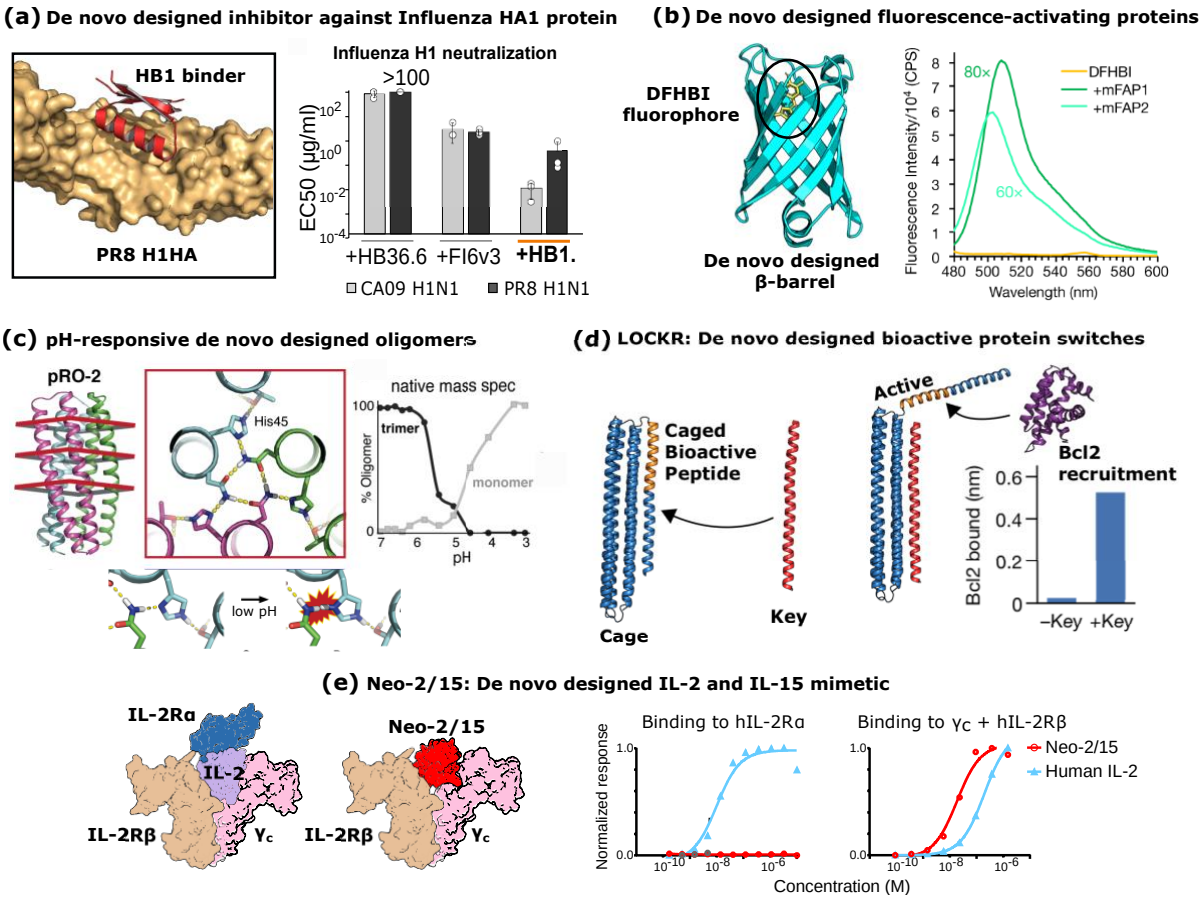


Figure 2. Recent *de novo* designed proteins with novel functions. (a) The left panel shows the structure of the *de novo* designed mini-protein inhibitor, HB1.6928.2.3 (red, cartoon representation), binding the stem region of influenza HA1 protein (beige, surface representation) (PDB ID: 5VLI) [58]. This protein showed improved efficacy neutralizing H1N1 CA09 (gray columns) and H1N1 PR8 (black columns) viral infection of canine kidney cells (right panel), compared to a previously designed binder (HB36.6) and the widely validated FI6v3 neutralizing antibody. **(b)** Structure of the *de novo* designed beta-barrel, mFAP1, bound to the DFHBI fluorophore group (PDB ID: 6CZI). The right panel shows the weak fluorescence intensity of free DFHBI (yellow line), which is significantly increased when DFHBI is bound to two mFAP protein variants (green and cyan lines) [70]. **(c)** Structure of a pH-responsive helical protein homotrimer with designed hydrogen bond networks in the core. These networks contain pre-organized histidine residues, which can be protonated at low pH, causing a disruption of core interactions and disassembly of the trimer (right panel, black line) into monomers (right panel, gray line) [69]. **(d)** Structure of a bioactive protein switch with customizable cellular functions. A small bioactive peptide (e.g. BIM, orange protein helix in cartoon representation) can be strategically grafted into the de novo scaffold (blue protein in cartoon representation) to "cage" (or "sequester") it, effectively making it inactive. In the presence of a high-affinity competitive peptide, called "key" (red protein in cartoon representation), the protein switch changes conformation, releasing the previously caged peptide, which can then perform its function (e.g., recruitment of Bcl2, purple protein in cartoon representation)[63]. **(e)** *De novo* designed IL-2 and IL-15 mimetic (Neo-2/15) tailored for high biological activity with no dependency on the IL-2Rα receptor. The upper panel compares the structure of IL-2 bound to the receptors IL-2Rα, IL-2Rβ and γ_c; and the structure of

Neo-2/15 bound to IL-2Rβ and γ_c. The lower panel shows Surface Plasmon Resonance binding experiments of IL-2 (blue lines) and Neo-2/15 (red lines) binding to the IL-2 receptors. Neo-2/15 shows no binding to IL-2Rα, but very potent binding to IL-2Rβ and γ_c receptors[71].

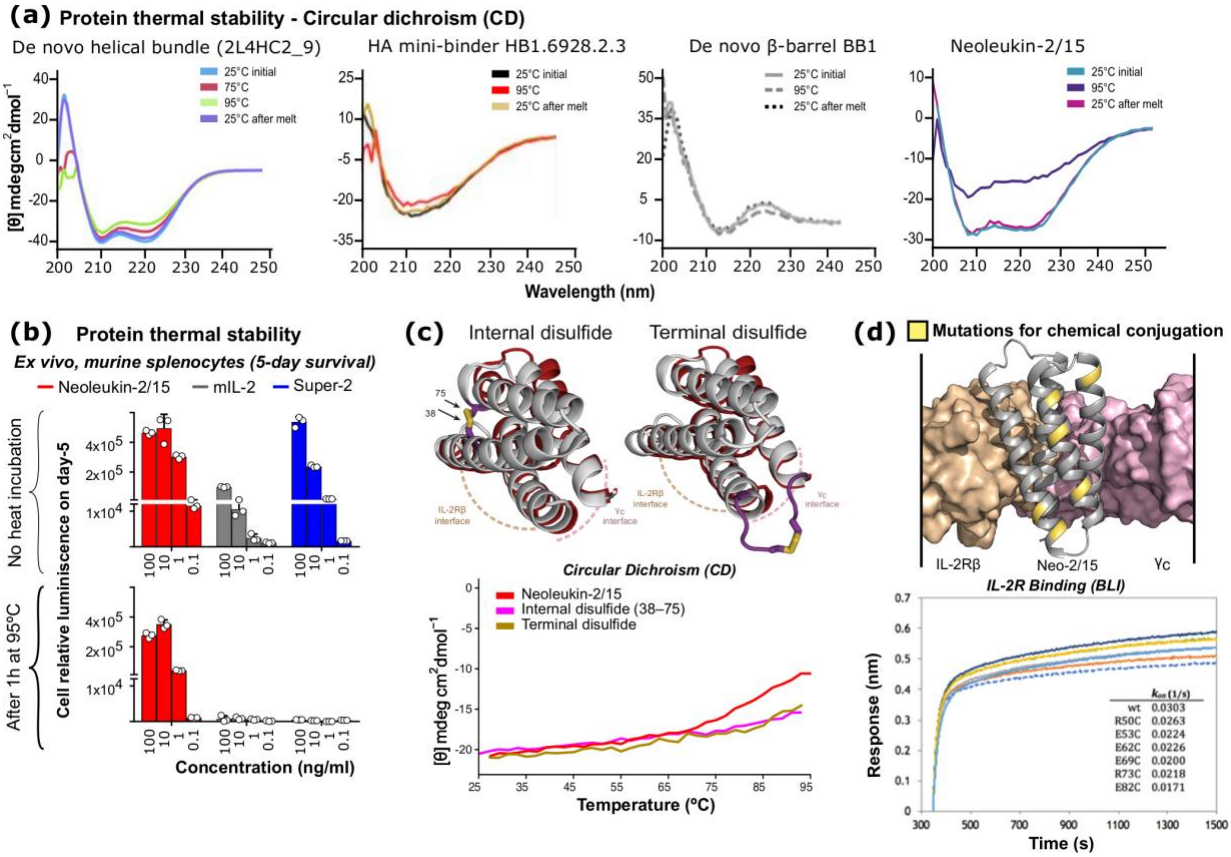


Figure 3. ***De novo*** **proteins present superior biophysical properties. (a)** Circular dichroism (CD) experiments illustrating the extreme thermostability of some *de novo* proteins. The plots show the spectra for four proteins, a de novo helical bundle homooligomer 2L4HC2_23 [53], HB1.6928.2.3 [58], BB1 *de novo* beta-barrel [70] and Neo-2/15 [71]. The proteins were incubated sequentially at 25°C, 95°C and then again at 25°C to evaluate re-folding. As with most *de novo* proteins, the examples here demonstrate extreme stability and nearly complete reversibility of thermal denaturation after heating. **(b)** The ability of Neo-2/15, mIL-2 and Super-2 [79] to expand murine splenocytes *ex vivo* was evaluated before and after heat incubation at 95°C. Neo-2/15 resisted the treatment and maintained its function [71]. **(c)** *De novo* proteins have been proven to be amenable to extensive modifications to improve their properties. For example, disulfide bridges can be seamlessly engineered into these proteins to improve their thermostability. The upper panel shows the structural models of two disulfide-stabilized variants of Neo-2/15 (grey), superposed on the ternary crystal structure of Neo-2/15 (red). The thermal melting of each variant demonstrated that the variants with disulfide bonds are indeed more thermostable than the parent protein [71]. **(d)** Mutational robustness of *de novo* proteins was demonstrated by introducing individual cysteine-mutations on the surface of Neo-2/15 scaffold (at non-interface/hotspot positions). As expected from the in-silico design, these mutations had minimal effect on the function or stability of the protein, at the same time they enable an easy and powerful way to perform chemical conjugations to the protein [71].

References

1. Dougan, M., Dranoff, G. & Dougan, S. K. Cancer Immunotherapy: Beyond Checkpoint Blockade. *Annu. Rev. Cancer Biol.* **3**, 55–75 (2019).
2. Zhang, H. & Chen, J. Current status and future directions of cancer immunotherapy. *J. Cancer* **9**, 1773–1781 (2018).
3. Dougan, M. & Dranoff, G. Immune therapy for cancer. *Annu. Rev. Immunol.* **27**, 83–117 (2009).
4. Wei, S. C., Duffy, C. R. & Allison, J. P. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. *Cancer Discov.* **8**, 1069–1086 (2018).
5. Darvin, P., Toor, S. M., Sasidharan Nair, V. & Elkord, E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp. Mol. Med.* **50**, 165 (2018).
6. Lim, W. A. & June, C. H. The Principles of Engineering Immune Cells to Treat Cancer. *Cell* **168**, 724–740 (2017).
7. Hyrenius-Wittsten, A. & Roybal, K. T. Paving New Roads for CARs. *Trends Cancer Res.* **5**, 583–592 (2019).
8. Belardelli, F., Ferrantini, M., Proietti, E. & Kirkwood, J. M. Interferon-alpha in tumor immunity and immunotherapy. *Cytokine & Growth Factor Reviews* vol. 13 119–134 (2002).
9. Rosenberg, S. A. IL-2: the first effective immunotherapy for human cancer. *J. Immunol.* **192**, 5451–5458 (2014).
10. Jiang, T., Zhou, C. & Ren, S. Role of IL-2 in cancer immunotherapy. *Oncoimmunology* **5**, e1163462 (2016).
11. Conlon, K. C., Miljkovic, M. D. & Waldmann, T. A. Cytokines in the Treatment of Cancer. *Journal of Interferon & Cytokine Research* vol. 39 6–21 (2019).
12. Postow, M. A., Sidlow, R. & Hellmann, M. D. Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. *N. Engl. J. Med.* **378**, 158–168 (2018).
13. Vasserot, A. P. *et al.* Optimization of protein therapeutics by directed evolution. *Drug Discovery Today* vol. 8 118–126 (2003).
14. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).
15. Spicer, C. D. & Davis, B. G. Selective chemical protein modification. *Nat. Commun.* **5**, 4740 (2014).
16. Chiu, M. L. & Gilliland, G. L. Engineering antibody therapeutics. *Current Opinion in Structural Biology* vol. 38 163–173 (2016).
17. Strohl, W. R. & Strohl, L. M. *Therapeutic Antibody Engineering: Current and Future Advances Driving the Strongest Growth Area in the Pharmaceutical Industry*. (Elsevier, 2012).
18. Muyldermans, S. Nanobodies: natural single-domain antibodies. *Annu. Rev. Biochem.* **82**, 775–797 (2013).
19. Dahlén, E., Veitonmäki, N. & Norlén, P. Bispecific antibodies in cancer immunotherapy. *Ther Adv Vaccines Immunother* **6**, 3–17 (2018).
20. Fercher, C., Keshvari, S., McGuckin, M. A. & Barnard, R. T. Evolution of the magic bullet: Single chain antibody fragments for the targeted delivery of immunomodulatory proteins. *Exp. Biol. Med.* **243**, 166–183 (2018).
21. Neri, D. Antibody–Cytokine Fusions: Versatile Products for the Modulation of Anticancer Immunity. *Cancer Immunol Res* **7**, 348–354 (2019).

22. Huehls, A. M., Coupet, T. A. & Sentman, C. L. Bispecific T-cell engagers for cancer immunotherapy. *Immunology and Cell Biology* vol. 93 290–296 (2015).
23. Gébleux, R. & Casi, G. Antibody-drug conjugates: Current status and future perspectives. *Pharmacol. Ther.* **167**, 48–59 (2016).
24. Carter, P. J. & Lazar, G. A. Next generation antibody drugs: pursuit of the 'high-hanging fruit'. *Nature Reviews Drug Discovery* vol. 17 197–223 (2018).
25. Taverna, D. M. & Goldstein, R. A. Why are proteins marginally stable? *Proteins* **46**, 105–109 (2002).
26. Sauerborn, M., Brinks, V., Jiskoot, W. & Schellekens, H. Immunological mechanism underlying the immune response to recombinant human protein therapeutics. *Trends Pharmacol. Sci.* **31**, 53–59 (2010).
27. Shire, S. J., Shahrokh, Z. & Liu, J. Challenges in the development of high protein concentration formulations. *J. Pharm. Sci.* **93**, 1390–1402 (2004).
28. Hermeling, S., Crommelin, D. J. A., Schellekens, H. & Jiskoot, W. Structure-immunogenicity relationships of therapeutic proteins. *Pharm. Res.* **21**, 897–903 (2004).
29. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
30. Baker, D. What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683 (2019).
31. Woolfson, D. N. *et al.* De novo protein design: how do we expand into the universe of possible protein structures? *Curr. Opin. Struct. Biol.* **33**, 16–26 (2015).
32. Marcos, E. & Silva, D.-A. Essentials of de novo protein design: Methods and applications. *WIREs Comput Mol Sci* **8**, e1374 (2018).
33. Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* **20**, 681–697 (2019).
34. Dawson, W. M., Rhys, G. G. & Woolfson, D. N. Towards functional de novo designed proteins. *Curr. Opin. Chem. Biol.* **52**, 102–111 (2019).
35. Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* **77**, 363–382 (2008).
36. Gainza-Cirauqui, P. & Correia, B. E. Computational protein design—the next generation tool to expand synthetic biology applications. *Curr. Opin. Biotechnol.* **52**, 145–152 (2018).
37. Grayson, K. J. & Anderson, J. L. R. Designed for life: biocompatible de novo designed proteins and components. *J. R. Soc. Interface* **15**, (2018).
38. Goldenzweig, A. & Fleishman, S. J. Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **87**, 105–129 (2018).
39. Dantas, G., Kuhlman, B., Callender, D., Wong, M. & Baker, D. A large scale test of computational protein design: folding and stability of nine completely redesigned globular proteins. *J. Mol. Biol.* **332**, 449–460 (2003).
40. Campeotto, I. *et al.* One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 998–1002 (2017).
41. Gordon, S. R. *et al.* Computational design of an α-gliadin peptidase. *J. Am. Chem. Soc.* **134**, 20513–20520

(2012).

42. Marcandalli, J. *et al.* Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **176**, 1420–1431.e17 (2019).
43. Adolf-Bryfogle, J. *et al.* RosettaAntibodyDesign (RAbD): A general framework for computational antibody design. *PLoS Comput. Biol.* **14**, e1006112 (2018).
44. Baran, D. *et al.* Principles for computational design of binding antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 10900–10905 (2017).
45. Nimrod, G. *et al.* Computational Design of Epitope-Specific Functional Antibodies. *Cell Rep.* **25**, 2121–2131.e5 (2018).
46. Leaver-Fay, A. *et al.* Computationally Designed Bispecific Antibodies using Negative State Repertoires. *Structure* **24**, 641–651 (2016).
47. Froning, K. J. *et al.* Computational design of a specific heavy chain/κ light chain interface for expressing fully IgG bispecific antibodies. *Protein Sci.* **26**, 2021–2038 (2017).
48. Kuhlman, B. *et al.* Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science* vol. 302 1364–1368 (2003).
49. Koga, N. *et al.* Principles for designing ideal protein structures. *Nature* **491**, 222–227 (2012).
50. Lin, Y.-R. *et al.* Control over overall shape and size in de novo designed proteins. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E5478–85 (2015).
51. King, I. C. *et al.* Precise assembly of complex beta sheet topologies from de novo designed building blocks. *Elife* **4**, (2015).
52. Huang, P.-S. *et al.* De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nat. Chem. Biol.* **12**, 29–34 (2016).
53. Boyken, S. E. *et al.* De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* vol. 352 680–687 (2016).
54. Bhardwaj, G. *et al.* Accurate de novo design of hyperstable constrained peptides. *Nature* **538**, 329–335 (2016).
55. Hosseinzadeh, P. *et al.* Comprehensive computational design of ordered peptide macrocycles. *Science* **358**, 1461–1466 (2017).
56. Marcos, E. *et al.* Principles for designing proteins with cavities formed by curved β sheets. *Science* **355**, 201–206 (2017).
57. Lu, P. *et al.* Accurate computational design of multipass transmembrane proteins. *Science* **359**, 1042–1046 (2018).
58. Rocklin, G. J. *et al.* Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* **357**, 168–175 (2017).
59. Hsia, Y. *et al.* Design of a hyperstable 60-subunit protein icosahedron. *Nature* vol. 535 136–139 (2016).
60. Silva, D. A., Correia, B. E. & Procko, E. Motif-driven design of protein–protein interfaces. *Computational Design of Ligand Binding* (2016).
61. Berger, S. *et al.* Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins

in cancer. *Elife* **5**, (2016).

62. Chevalier, A. *et al.* Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017).
63. Langan, R. A. *et al.* De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
64. Lajoie, M. J. *et al.* Designed protein logic to target cells with precise combinations of surface antigens. *Science* **369**, 1637–1643 (2020).
65. Ng, A. H. *et al.* Modular and tunable biological feedback control using a de novo protein switch. *Nature* vol. 572 265–269 (2019).
66. Procko, E. *et al.* A computationally designed inhibitor of an Epstein-Barr viral Bcl-2 protein induces apoptosis in infected cells. *Cell* **157**, 1644–1656 (2014).
67. Correia, B. E. *et al.* Proof of principle for epitope-focused vaccine design. *Nature* vol. 507 201–206 (2014).
68. Joh, N. H. *et al.* De novo design of a transmembrane Zn2 -transporting four-helix bundle. *Science* vol. 346 1520–1524 (2014).
69. Boyken, S. E. *et al.* De novo design of tunable, pH-driven conformational changes. *Science* **364**, 658–664 (2019).
70. Dou, J. *et al.* De novo design of a fluorescence-activating β-barrel. *Nature* **561**, 485–491 (2018).
71. Silva, D.-A. *et al.* De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* **565**, 186 (2019).
72. Razvi, A. & Scholtz, J. M. Lessons in stability from thermophilic proteins. *Protein Sci.* **15**, 1569–1578 (2006).
73. Geiger-Schuller, K. *et al.* Extreme stability in de novo-designed repeat arrays is determined by unusually stable short-range interactions. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 7539–7544 (2018).
74. Huang, P.-S. *et al.* High thermodynamic stability of parametrically designed helical bundles. *Science* **346**, 481–485 (2014).
75. Silva, D.-A., Stewart, L., Lam, K.-H., Jin, R. & Baker, D. Structures and disulfide cross-linking of de novo designed therapeutic mini-proteins. *FEBS J.* **285**, 1783–1785 (2018).
76. Boyman, O. & Sprent, J. The role of interleukin-2 during homeostasis and activation of the immune system. *Nat. Rev. Immunol.* **12**, 180–190 (2012).
77. Siegel, J. P. & Puri, R. K. Interleukin-2 toxicity. *J. Clin. Oncol.* **9**, 694–704 (1991).
78. Krieg, C., Létourneau, S., Pantaleo, G. & Boyman, O. Improved IL-2 immunotherapy by selective stimulation of IL-2 receptors on lymphocytes and endothelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11906–11911 (2010).
79. Levin, A. M. *et al.* Exploiting a natural conformational switch to engineer an interleukin-2 'superkine'. *Nature* **484**, 529–533 (2012).
80. Johannsen, M. *et al.* The tumour-targeting human L19-IL2 immunocytokine: Preclinical safety studies, phase I clinical trial in patients with solid tumours and expansion into patients with advanced renal cell carcinoma. *European Journal of Cancer* vol. 46 2926–2935 (2010).
81. Klein, C. *et al.* Cergutuzumab amunaleukin (CEA-IL2v), a CEA-targeted IL-2 variant-based immunocytokine for combination cancer immunotherapy: Overcoming limitations of aldesleukin and conventional IL-2-based immunocytokines. *Oncoimmunology* **6**, e1277306 (2017).

Chapter 2

Targeted immunotherapy with a designed conditionally active IL-2 mimic

Abstract

Human IL-2 (IL-2) is a potent pleiotropic cytokine approved for treating melanoma and renal cell carcinoma [1], but its therapeutic potential is limited by preferential activation of CD25-positive cells and severe dose-limiting toxicities associated with systemic IL-2 activity [2]. The *de novo* designed IL-2 mimic Neoleukin-2/15 (Neo-2/15) does not bind CD25 and has superior therapeutic activity and lower toxicity than natural IL-2 in mouse tumor models [3]. Taking advantage of the robust folding of Neo-2/15, we designed conditionally active cytokine mimics consisting of two components that do not elicit IL-2 signaling individually, but fully regain activity when co-localized. The fragments can be independently targeted to cells expressing specific surface markers, and we demonstrate their use in trans-activating immune cells surrounding targeted tumor cells, and in cis-activating directly targeted immune cells. In trans-activation mode, tumor-antigen targeting of the two components enhanced antitumor activity and attenuated toxicity compared to systemic treatment with Neo-2/15 in a syngeneic murine melanoma model. In cis-activation mode, immune cell targeting of the two components selectively expanded CD8+ T cells in a syngeneic mouse melanoma model, and promoted CAR T cell function in a lymphoma xenograft model, enhancing antitumor efficacy in both cases. Conditionally active de novo cytokine mimetics provide a general route to overcoming long-standing limitations imposed by the severe side-effects of systemic cytokine therapies.

Main Text

Immunostimulatory cytokines are central regulators of the immune system and have shown great potential as immunotherapeutic agents [4,5]. However, their clinical use has been limited due to severe toxicity resulting from undesired systemic immune activation [2,5]. Fusions of cytokines to monoclonal antibodies or antibody fragments (immunocytokines) [6–8] have been developed with the aim of improving efficacy and reducing systemic toxicity by targeting cytokine activity. However, due to the high affinity of the cytokines for their cognate receptors (and the broad systemic distribution of these receptors), these approaches have shown limited success in mitigating toxicity and improving efficacy [9,10]. Several strategies have been explored to further limit systemic cytokine signaling: reducing the affinity of the cytokine moiety to improve targeting [9,11–14], cytokine fusions to inhibitory moieties that can be cleaved by tumor-associated proteases [15,16], and conditionally active versions of oligomeric cytokines that are activated by colocalizing the monomeric subunits [17,18]. However, for monomeric cytokines such as IL-2, the development of immunotherapeutics with effective "activity on-demand" has remained elusive.

Like most natural proteins, IL-2 has structural irregularities that translate into non-ideal biochemical properties, such as low solubility and poor mutational robustness [19,20], which hampers efforts to repurpose it as a conditionally active therapeutic agent. Recent advances in computational protein design [21] enabled the generation of highly stable functional proteins that fully recapitulate the activity of natural cytokines, but with improved biochemical and therapeutic properties [3,22]. Neoleukin-2/15 (Neo-2/15) is a *de novo* IL-2 mimic that reproduces the immunostimulatory function of IL-2 and IL-15, but is fully independent of CD25 and CD215 [3]. In preclinical cancer models, Neo-2/15 induces potent immunotherapeutic effects with reduced toxicity when compared to natural IL-2, but at high doses the systemic administration of Neoleukin-2/15 can still be toxic. We reasoned that the high stability and robust folding of Neo-2/15 could be leveraged to create a conditionally active cytokine, and set out to design two-component split versions of Neo-2/15 that are active only upon co-localization of the two disjointed fragments.

Development of functional Split Neo-2/15 variants

In Neo-2/15, helix H3 interacts with IL-2Rβ, H4 interacts with IL-2Rγ, and H1 interacts with both receptors (Fig. 1a). We aimed to separate one of these helices from the others so that receptor heterodimerization could only be achieved in the presence of both fragments. We tested three pairs of split versions for Neo-2/15: (i) between the helical elements H1 and H3-H2'-H4, (ii) between the helical elements H1-H3 and H2'-H4, or (iii) between the helical elements H1-H3-H2' and H4. In all three cases, much stronger IL-2 receptor binding and signaling was observed when the two fragments were combined than for either fragment alone (Fig. 1b,c, Extended Data Fig. 1). For the split pair H1 and H32'4, neither fragment alone showed any activity even at high concentrations (Extended Data Fig. 1a). We selected this pair for subsequent studies and the H1 and H32'4 split fragments are henceforth referred to as Neo2A and Neo2B, respectively. Further characterization of the fragments revealed that Neo2A is unfolded in solution, while Neo2B remains helical (Extended Data Fig. 2a). The two split fragments have low affinity for each other in absence of the IL-2 receptor subunits (Kd ~4.5 μM) but form a much higher affinity complex in the presence of the soluble IL-2Rβ and IL-2Rγ (Kd = 50.8 nM), a desirable property to ensure on-target reconstitution (Extended Data Fig. 2b-c).

Split Neo-2/15 is selectively reconstituted on target cells

We next sought to evaluate whether selective targeting of the Split Neo-2/15 components to the surface of cells could reconstitute activity at sub-micromolar concentrations through co-localization. We fused the fragments to designed ankyrin repeat proteins (DARPins) that target the extracellular domains of Her2 (G3 DARPin[23]) and the EGFR (E01 DARPin[24]). Intact Neo-2/15 and the split fragments thereof retained their function after either N- or C-terminal fusion to the targeting domains and irrespective of linker length (Extended Data Fig. 3a-e). The targeted split fragments were added to a mixture of four K562 cell lines engineered to express EGFR-iRFP, Her2-eGFP, both, or neither [25] (Figure 2a), and reconstitution of Neo-2/15 activity on the surface of target cells was measured by recruitment of soluble IL-2Rβγ complex (Fig. 2a). IL-2Rβγ was specifically recruited to Her2+ cells when both Neo2A and Neo2B split fragments were targeted to Her2 (αHer2-Neo2A + αHer2-Neo2B) with similar activity to Her2-targeted intact Neo-2/15 (αHer2-Neo2/15), demonstrating co-localization-dependent receptor binding (Fig. 2b, Extended Data Fig. 4a). IL-2Rβγ bound with high selectivity to the surface of double-positive Her2+/EGFR+ when one split fragment was targeted to Her2 and the other to EGFR (Fig. 2b, Extended Data Fig. 4b), demonstrating the feasibility of this approach to target two different antigens on the same cell surface. Similar selectivity was observed for other split pair variants (Extended Data Fig. 4c-d), but at high concentrations, their single fragments showed residual receptor binding capacity (Extended Data Fig. 4e).

Split Neo-2/15 trans-activates immune cells

To evaluate whether the targeted Split Neo-2/15 proteins could trans-activate immune cells when bound to target tumor cells, we first targeted the fragments to the double-positive Her2+/EGFR+ K562 cell line described above in the presence of an IL-2 responsive human natural killer cell line (YT-1) and assessed STAT5 phosphorylation as a readout for IL-2 signaling (Fig. 2c). Selective activation was observed for the targeted Neo2A and Neo2B pair (Fig. 2d, Extended Data Fig. 5a), but required a high ratio (20:1) of tumor cells to immune cells. We hypothesized that more potent trans-activation could occur for antigen-specific immune cells, since the immune synapse between the immune cell and the tumor cell would bring them into close proximity for an extended period of time, allowing the targeted cytokine to more efficiently dimerize the receptor subunits and activate the immune cell. To test this hypothesis, we targeted Neo2A and Neo2B to B16 melanoma cells overexpressing PD-L1 [26] by fusing them to the αPD-L1 nanobody B3 [27,28] and co-cultured them overnight with melanoma antigen-specific αTrp-1 CD8+ T cells (Fig. 2e) in the presence of an αCD28 antibody to provide co-stimulation. The Split Neo-2/15 fragments targeted to the surface of B16 cells efficiently trans-activated the antigen-specific T cells (Fig. 2f, Extended Data Fig. 6a). This activation was dependent on targeting: fusions to a control nanobody with irrelevant specificity (1B7, which recognizes a Toxoplasma gondii kinase[29]) did not activate T cells, and addition of an excess of soluble αPD-L1 nanobody to compete out cell surface binding considerably reduced trans-activation (Fig. 2g, Extended Data Fig. 6b).

Split Neo-2/15 reduces dose-limiting toxicity in murine models

We next sought to evaluate the safety profile and therapeutic activity of Split Neo-2/15. IL-2 immunotherapy induces toxicities such as pulmonary edema and weight loss in mice [2]. We previously showed that treatment with Neo-2/15 has significantly lower systemic toxicity than IL-2 in preclinical models [3]. We hypothesized that the conditional activity of Split Neo-2/15 could further expand the therapeutic window by preventing systemic activity of the individual fragments

until co-localized to the desired site. We treated healthy C57BL/6J mice with equimolar doses of Neo-2/15, targeted and untargeted Neo-2/15 fusion proteins (αPD-L1-Neo2/15 and Ctrl-Neo2/15), and targeted and untargeted Split Neo-2/15 fusion proteins (Fig. 3a) at a daily dose of 2.6 nmol (equivalent to 30 μg of Neo-2/15, higher than previously reported and anticipated to show moderate toxic effects [3]). Signs of toxicity were evident in mice treated with Neo-2/15 and Ctrl-Neo2/15 at 23 days post-treatment due to the high dose (Fig. 3a, Extended Data Fig. 7). The αPD-L1-Neo-2/15 fusion protein was found to be toxic just days into treatment, possibly due to accumulation by unwanted localization to PD-L1 expressing cells such as myeloid dendritic cells and brown adipocytes [27]. As anticipated, neither the untargeted nor the targeted Split Neo-2/15 fragment combinations showed any signs of toxicity at this dose.

Targeted Split Neo-2/15 induces potent antitumor effects in syngeneic melanoma models

We evaluated the antitumor effects of Split Neo-2/15 in syngeneic murine models of PD-L1-overexpressing B16 melanoma tumors [29]. First, we compared the therapeutic effect of intact Neo-2/15 fusion proteins (untargeted or targeted to PD-L1) with that of the Split Neo-2/15 fusion protein (untargeted or targeted to PD-L1). The therapeutic molecules were administered at the highest non-toxic dose, i.e., intact targeted Neo-2/15 controls (αPD-L1-Neo-2/15 and Ctrl-Neo-2/15) at 12 μg/mouse/day (430pmol), whereas the split fragments were dosed at a ~20-fold higher molar ratio (200 μg/mouse/day, 8 nmol). All proteins were administered intraperitoneally, except the Neo2B fusions, which were administered subcutaneously at the opposite flank of the tumor to prevent potential association of the split molecules before injection. The molecules were administered as single-agent therapeutics (Extended Data Fig. 8a) or in combination with the anti-melanoma antibody TA99 (Fig. 3b). Survival and tumor growth were evaluated for all groups. The antitumor efficacy of the PD-L1-targeted split fragments (αPD-L1-Neo2A + αPD-L1-Neo2B) was superior to that of the untargeted split fragments (Ctrl-Neo2A + Ctrl-Neo2B) as well as the targeted and untargeted intact Neo-2/15 fusion proteins, resulting in extended survival and complete tumor clearance in two mice (Fig. 3b). As expected, the individual fragments did not show any antitumor therapeutic activity. Consistent with these efficacy results, αPD-L1-Neo2A + αPD-L1-Neo2B therapy resulted in expansion and activation of CD8+ T cell expansion in peripheral blood (Extended Data Fig. 8c-d). Intact Neo-2/15 fusions (αPD-L1-Neo2/15 and Ctrl-Neo2/15) did not show a potent antitumor effect due to dose-limiting toxicity, as predicted by the safety studies presented in Fig. 3a and Extended Data Fig. 8b. Both intact Neo-2/15 fusions performed similarly regardless of targeting, in agreement with previous reports showing that efficacy of a targeted fully active IL-2-mimetic molecule is mainly mediated by systemic activity[9].

We next carried out a more extensive comparison of the antitumor efficacy of Neo-2/15 compared to untargeted and targeted Split Neo-2/15 (Fig. 3c). The αPD-L1 Split Neo2/15 molecules showed little toxicity and exhibited superior efficacy and safety compared to intact Neo-2/15 (Fig. 3d). The enhanced therapeutic effect was likely a result of increased on-tumor accumulation of active Neo-2/15 due to a higher tolerated dose and reduced systemic effects. In the case of mice treated with untargeted control Split Neo-2/15 at the high 8 nmol dose, we observed signs of systemic toxicity, further highlighting the benefit of targeting the conditionally-active cytokines to concentrate their activity at the tumor site. Remarkably, among all the treatment groups, the cohort treated with αPD-L1 Split Neo-2/15 was the only group in which some mice achieved complete tumor remission (Fig. 3e). We evaluated the durability of this effect by re-challenging the "cured" mice with new B16 tumors (Fig. 3f). Rechallenged mice showed improved survival (one mouse

again achieved complete remission) and had higher numbers of tumor-specific (anti-Trp1) CD8+ T cells compared to mice receiving a primary challenge, demonstrating immunologic memory. Overall, these results demonstrate the therapeutic potential of the targeted Split Neo-2/15 molecule, highlighting its ability to overcome dose-limiting toxicities often observed in the development of immunocytokines.

Specific cis-activation of immune cells with Split Neo-2/15 for immunotherapy

The results above illustrate the potential of Split Neo-2/15 for targeting solid tumors to locally expand immune cell populations (trans-activation of immune cells). Some immunotherapeutic strategies benefit from cytokine-mediated systemic amplification of specific immune cell subtypes to promote antitumor effects (cis-activation), for instance potentiating pre-existing cytotoxic CD8+ T cells [13,30] or Chimeric antigen receptor T (CAR-T) cells in adoptive cell therapy applications [31]. Whereas unwanted activation and expansion of CD25+ regulatory T cells (Tregs) has limited many IL-2 based immunotherapies [10,32], Neo-2/15 is CD25-independent and is thus not biased to preferentially activate CD25+ Tregs [3], which is one of its main advantages as a cancer therapeutic over IL-2. To specifically target CD8+ cytotoxic T cells, we fused the Split Neo-2/15 proteins to an αCD8 nanobody targeting domain [33] (Fig. 4a). First, we confirmed the in vitro capacity of the resulting αCD8 Split Neo-2/15 to selectively expand CD8+ T cells in cultures of splenocytes (containing both CD8+ and CD4+ T cells, Extended Data Fig. 9a-b). We then dosed αCD8 Split Neo-2/15, αCD8 intact Neo-2/15, and untargeted controls in healthy Foxp3-GFP mice and assessed the extent of selective expansion of CD8+ T cells over other T cell subtypes, such as Tregs (Fig. 4b, Extended Data Fig. 9c-d). We observed increased CD8+ frequencies in lymph node and spleen in animals treated with αCD8-Neo2/15 and αCD8 Split Neo-2/15. The cohort treated with targeted Split Neo-2/15 induced greater CD8+ T cell-specific proliferation compared to the cohort treated with targeted intact Neo-2/15, likely due to reduced off-target effects of the conditionally-active molecule. We then evaluated the antitumor efficacy of CD8-specific Split Neo-2/15 in the B16 mouse model of melanoma. Intact Neo-2/15 and split fragment fusion proteins were dosed as single agents (Extended Data Fig. 9e) and in combination with Ta99 (Fig. 4c). Consistent with our targeted activation studies, we observed delayed tumor growth and improved survival in mice treated with the αCD8 Split Neo-2/15 compared to both untreated mice and mice treated with untargeted split pair.

CAR-T therapies have achieved success in a subset of patients with B cell malignancies, but many patients do not respond to these treatments and additional challenges remain for their clinical application in solid tumors [34,35]. For instance, insufficient CAR-T cell expansion is correlated with treatment failure in hematological malignancies [36], and poor CAR-T cell accumulation at the tumor site is believed to lessen effectiveness in the treatment of epithelial cancers [37,38]. We hypothesized that highly-specific targeting of Split Neo-2/15 to the CAR-T cells would result in improved therapy through the selective expansion of CAR-T cells and reduced interaction with other immune cells that can lead to toxicity by immune overactivation. We engineered CD19 CAR-T cells that co-express truncated Her2 as a transduction marker for use as target for the split Neo-2/15 αHer2 fusion proteins described above (Fig. 4d). Split Neo-2/15 fragments fused to an EpCam binding DARPin (Ec1)[39] were used as an untargeted controls. The αHer2 Split Neo-2/15 molecules selectively activated IL-2 signaling leading to STAT5 phosphorylation in CAR transduced Her2+ cells, whereas αHer2 intact Neo-2/15 activated either transduced (Her2+) or untransduced (Her2-) cells (Fig. 4e). In a lymphoma xenograft model (Raji), co-treatment with CAR-T cells and targeted

split Neo2/15 exhibited improved tumor control and prolonged survival compared to CAR T-only treated control mice and compared to the untargeted split Neo2/15 (Fig. 4f, Extended Data Fig. 10). Taken together, these cis-activation studies illustrate the potential of our targeted conditional activation approach in cancer immunotherapy.

Discussion

Split Neo-2/15 is, to our knowledge, the first example of an IL-2-like molecule with designed conditional activity. The two components have no activity on their own, allowing for effective localization of the cytokine activity to the desired site of action, overcoming problems with residual systemic activity observed with intact immunocytokines fused to targeting domains [9]. Targeting the two components to the tumor to elicit trans-activation results in an increased therapeutic index and leads to complete remissions in syngeneic melanoma tumor models. Targeting the components to specific immune cells to elicit cis-activation enhances antitumor immunotherapeutic effects in syngeneic melanoma and lymphoma xenograft models in mice. Our approach can be readily extended to target other immune cells subsets of interest, such as NK cells to potentiate their innate anti-tumor immunity [40,41], or Tregs for treatment of autoimmune disorders [42,43]. Our approach also enables simultaneously targeting of two different cell surface antigens (i.e. AND logic[25,44]) that can be present on either the tumor cells (trans-activation) or the immune cells (cis-activation). More generally, this work highlights how the robust folding and high stability of *de novo* designed proteins can be exploited to create conditionally active signaling molecules that function robustly in animals, opening the door to a broad range of therapeutic and synthetic biology applications.

Figures and figure legends

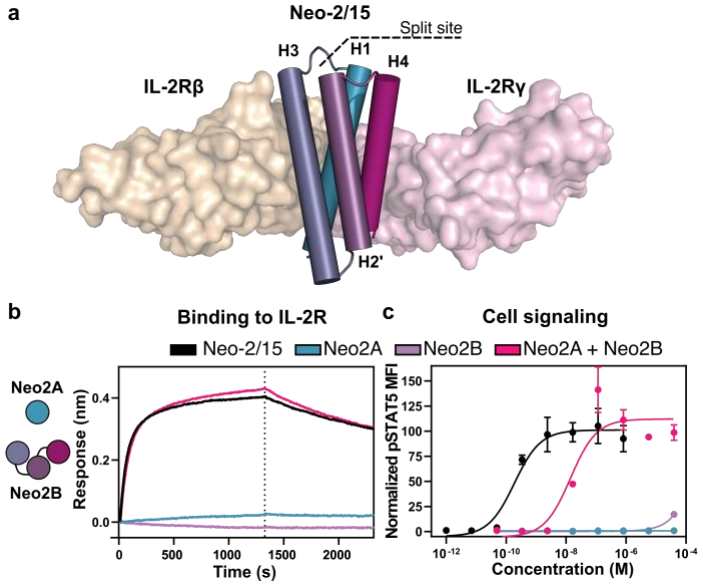


Figure 1. Neoleukin-2/15 can be split into two fragments that reconstitute activity when combined. a, Splitting strategy for Neo-2/15. Structure of Neo-2/15 (cylinder representation) and murine IL-2 Receptors β and γ (beige and pink surface representations respectively, PDB ID: 6dg5). Dashed line represents the selected split site between helix H1 (Neo2A), involved in the interface with both receptors, and helix H32'4 (Neo2B). **b**, Binding of Split Neo-2/15 fragments to the IL-2 Receptor. Biolayer interferometry binding assay illustrating the binding kinetics of intact Neo-2/15 (black line), Neo2A (blue), Neo2B (purple) and the combination of Neo2A + Neo2B (magenta) (1 μM concentration) binding to biotinylated human IL-2Rγ immobilized on an Octet streptavidin sensor in the presence of 250nM soluble hIL-2Rβ. Full titrations are provided in Extended Data Fig. 1a. **c**, Signaling of Split Neo-2/15 on human YT-1 cells. STAT5 phosphorylation in YT-1 cells following treatment with intact Neo-2/15 (black line), Neo2A (blue), Neo2B (purple), or the combination of Neo2A + Neo2B (magenta). Treatment with the split pair reconstitutes Neo-2/15 activity.

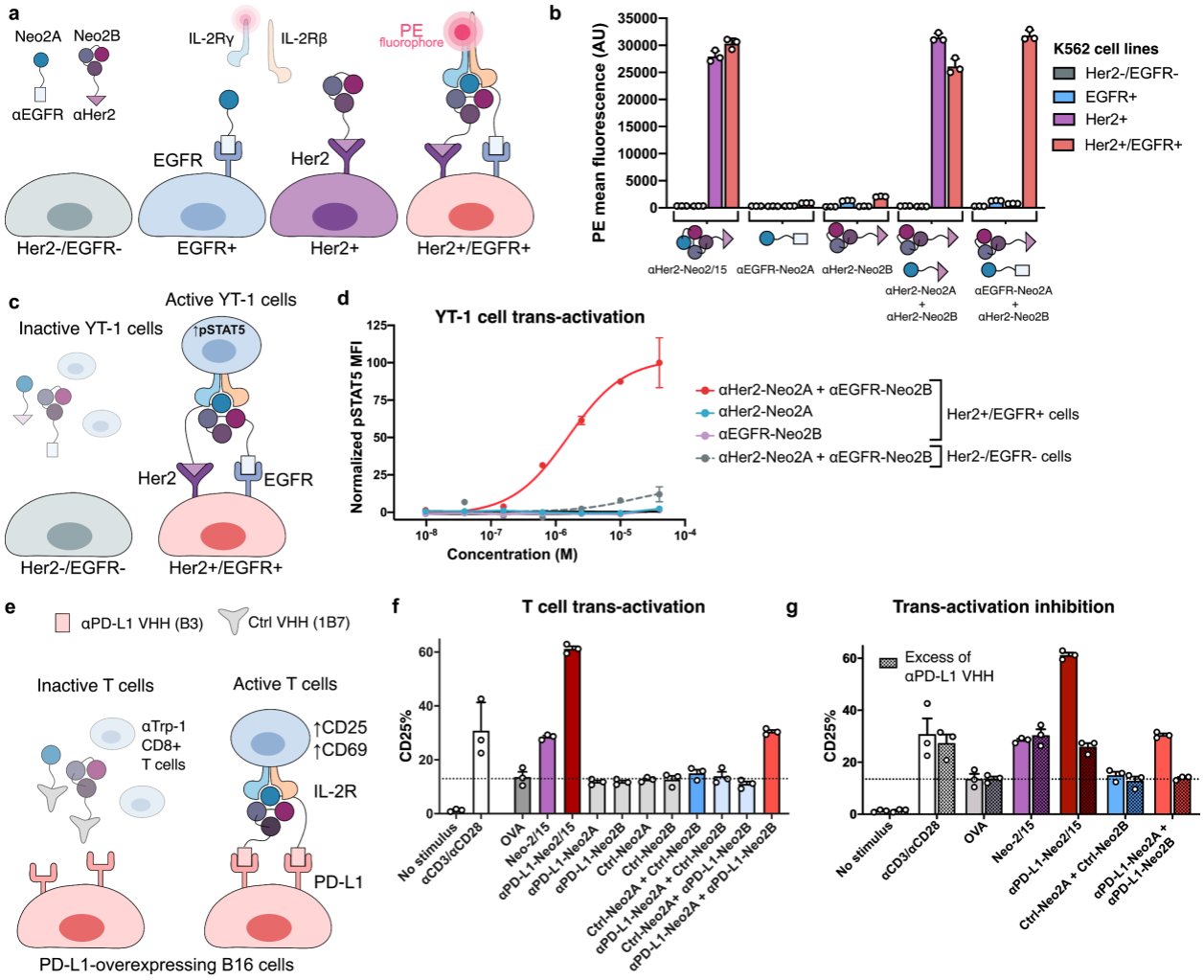


Figure 2. Trans-activation of immune cells through targeted reconstitution of Split Neo-2/15 on the surface of tumor cells. a, In vitro assay for reconstitution of Split Neo-2/15 binding activity on the surface of on-target K562 cells expressing Her2 and EGFR. Neo2A (H1) and Neo2B (H32'4) split protein variants were fused to anti-EGFR and anti-Her2 DARPin targeting domains, respectively. K562 cells (gray) and engineered K562 cell lines transduced with Her2-eGFP (purple), EGFR-iRFP (blue), or both Her2-eGFP and EGFR-iRFP (pink) were mixed in an equivalent ratio, incubated with an equimolar ratio of targeted Split Neo-2/15 proteins, washed, and then incubated with a mixture of: soluble hIL-2Rβ, soluble biotinylated hIL-2Rγ, and streptavidin-phycoerythrin fluorophore conjugate (PE). Reconstitution of Neo-2/15 binding activity on the cell surface was measured by recruitment of PE-labelled hIL-2Rβγ. Quantification of hIL-2Rbg binding to K562 cells incubated with the indicated Neo-2/15 and split Neo-2/15 fragment fusion proteins, as shown in (**a**). When the two split fragments were targeted to different surface markers (αEGFR-Neo2A + αHer2-Neo2B), selective reconstitution on the surface of double-positive Her2+/EGFR+ cells was observed. **c**, Cellular assay to detect trans-presentation of split Neo-2/15 on the surface of tumor cells to immune cells. Untransduced Her2-/EGFR- K562 cells (Off-target) or Her2+/EGFR+ co-expressing K562 cells (On-target) were co-cultured with human YT-1 NK cells in a 20:1 ratio (K562:YT-1) in the presence of αHer2-Neo2A and αEGFR-Neo2B fragment fusion proteins. YT-1 cell activation was analyzed by measuring STAT5 phosphorylation. **d**, STAT5 phosphorylation

of YT-1 cells from the experiment shown in (**c**). Strong signaling was observed for on-target Her2+/EGFR+ K562 cells incubated with αHer2-Neo2A + αEGFR-Neo2B, demonstrating trans-activation of immune cells from the surface of target-expressing cells. **e**, Cellular assay to detect trans-activation of antigen-specific αTrp-1 CD8+ T cells by split Neo-2/15 on the surface of B16 melanoma cells overexpressing PD-L1, allowing trans-activation of antigen-specific αTrp-1 CD8+ T cells in co-culture. Split Neo-2/15 proteins were fused to a nanobody (VHH) with irrelevant specificity (Ctrl) or an αPD-L1 nanobody (B3). **f**, Trans-activation of T cells in the co-culture assay described in (**e**), as measured by expression of the activation marker CD25. αCD3 and αCD28 antibody treatment was used as a positive control for T cell activation. OVA peptide was used to quantify basal CD25 expression levels in co-culture conditions. All samples were incubated with an αCD28 antibody to provide co-stimulation. **g**, B16 and CD8+ T cells were co-cultured in the presence of activating proteins, as described in (**e**). Addition of excess soluble αPD-L1 nanobody (VHH) to competitively inhibit binding of targeted split fusion proteins to the B16 cell surface resulted in attenuated T cell activation, confirming trans-activation of immune cells from the surface of tumor cells.

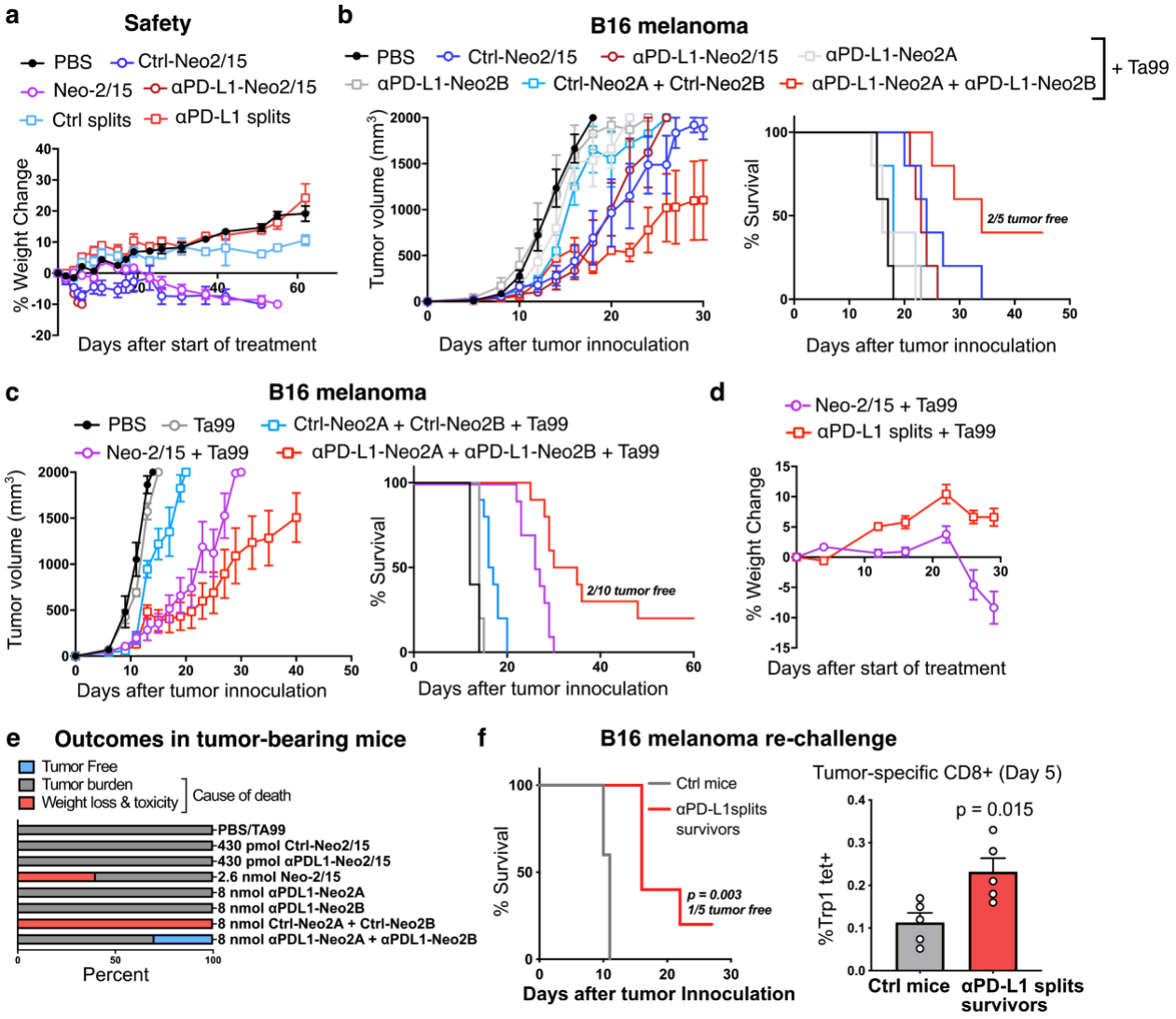


Figure 3. Targeting Split Neo-2/15 to tumors increases safety and enhances anti-tumor efficacy. a, Safety study in immunocompetent C57BL/6J mice (n=5/group) treated with targeted Neo-2/15 and targeted Split Neo-2/15. Mice were treated daily with equivalent doses of the indicated proteins (2.6 nmol, equivalent to 30 µg/mouse of Neo-2/15). Weight change and survival (Extended Data Fig. 7a) were monitored to evaluate toxicity of the molecules. **b,** Efficacy study in C57BL/6J mice (n=5/group) bearing B16 PD-L1-overexpressing melanoma cells in the flank**.** Starting on day 5 after tumor cell inoculation, mice were administered daily with therapeutic doses of αPD-L1-Neo2/15 and Ctrl-Neo2/15 at 430 nmol (12 µg/mouse) (intraperitoneally) or targeted split Neo-2/15 fragments were administered at 8 nmol (200 µg/mouse) (Neo2A fusions given intraperitoneally, and Neo2B fusions given subcutaneously at the opposite flank of the tumor). All groups were co-treated with Ta99, an αTrp1 melanoma-specific antibody administered bi-weekly starting on day 3 (150 µg/mouse). "Ctrl" represents fusion to an irrelevant nanobody as untargeted control (1B7[27]). **c,** Efficacy study in C57BL/6J mice bearing B16 PD-L1-overexpressing melanoma cells in the flank (n=5 for PBS and TA99 groups, n=10 for other groups)**.** Starting on day 5 after tumor cell inoculation, mice were administered daily with each test article at therapeutic

doses (Neo-2/15 at 2.6nmol (30 μg/mouse) and the split Neo-2/15 fragments at 8 nmol (~200 μg/mouse)). The indicated groups were co-treated with Ta99 mAb administered bi-weekly starting on day 3 (150 μg/mouse). **d,** Weight change in (**c**) to evaluate the toxicity of Neo-2/15 (2.6 nmol) and the αPD-L1 splits Neo-2/15 fusions (8 nmol) at full therapeutic doses. **e,** Summary of outcomes for efficacy studies shown in (**b**) and (**c**). **f,** Surviving αPD-L1 split Neo-2/15 fragment-treated mice were re-challenged alongside control mice receiving a primary challenge (n=5/group). CD8+ T cells from the treated mice were analyzed for recognition of tumor cells via presence of anti-Trp1-MHCI tetramer (right panel).

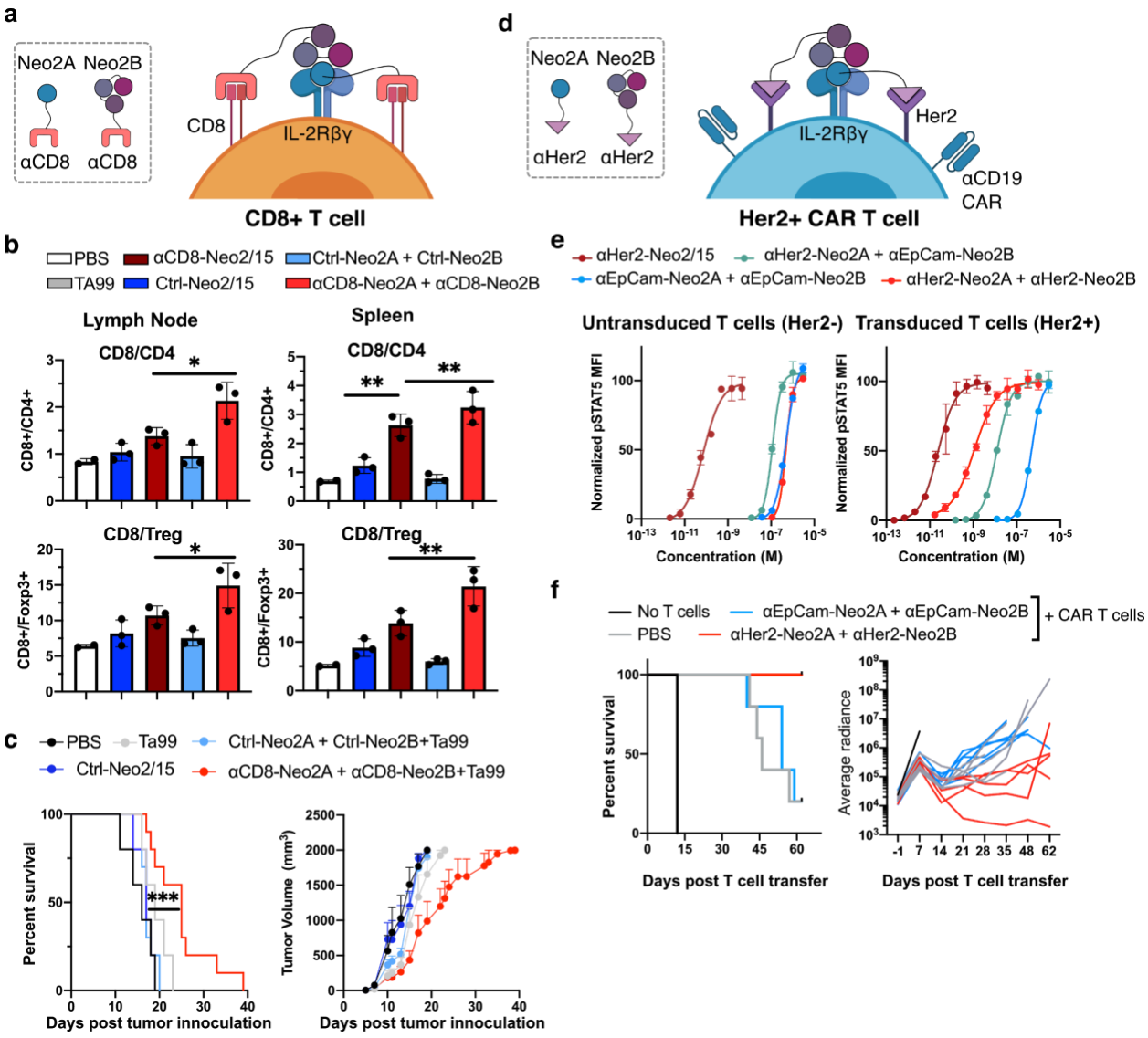


Figure 4. Targeting Split Neo-2/15 to specific immune cell subtypes enhances efficacy of immunotherapy. a, Depiction of the cis-activation mechanism of Split Neo-2/15 targeted to CD8 on the surface of T cells. **b**, αCD8 Split Neo-2/15 promotes specific CD8+ T cell proliferation in healthy mice. Proteins were administered daily to non-tumor bearing Foxp3-GFP mice for five days at 12 µg/mouse/day (500 pmol) for the intact Neo-2/15 fusions and 10 µg/mouse/day (500 pmol) for the Split Neo-2/15 fusions. On day 6, mice were euthanized, spleen and both inguinal lymph nodes were collected and cell populations were investigated by flow cytometry to evaluate specific expansion of CD8+ T cell populations. The Split Neo-2/15 fragments were also fused to an irrelevant targeted nanobody as untargeted control (Ctrl-Neo2A and Ctrl-Neo2B). * indicates P<0.05; ** indicates P<0.01, unpaired Student's t test. **c**, Efficacy study in C57BL/6J mice bearing B16 melanoma treated with αCD8 Split Neo-2/15 and Ta99. B16 melanoma tumors were inoculated subcutaneously on the left flank of B6 mice on day 0. Ta99 was administered biweekly starting on day 3 (150 µg/mouse) and all other proteins were administered daily starting on day 5. Neo2A splits were administered subcutaneously and Neo2B splits were administered

intraperitoneally at 500 pmol each. Tumor growth (left) and mouse survival (right) are shown (n=10 mice per group for split Neo-2/15 treated mice, n=5 for controls). * indicates P = 0.0303, Mantel−Cox log rank. **d,** Depiction of the cis-activation mechanism of Split Neo-2/15 targeted to Her2 expressed as a transduction marker on the surface of αCD19 CAR-T cells. **e**, STAT5 phosphorylation of untransduced Her2- primary CD8+ T-cells (left) or transduced Her2+ CD8+ T cells expressing the αCD19 CAR (right). An EpCam binding DARPin (Ec1) was fused to the split Neo-2/15 fragments as an untargeted control. **f**, Efficacy of αCD19 CAR-T cells and co-treatment with CAR-T-targeted split Neo-2/15 fragments in a lymphoma xenograft model. Both Her2-tageted and EpCam-targeted (control) split Neo-2/15 fragments were dosed at 7.5 mg/kg (n=5 mice per group).

References

1. Rosenberg, S. A. IL-2: the first effective immunotherapy for human cancer. *J. Immunol.* **192**, 5451–5458 (2014).
2. Siegel, J. P. & Puri, R. K. Interleukin-2 toxicity. *J. Clin. Oncol.* **9**, 694–704 (1991).
3. Silva, D.-A. *et al.* De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* **565**, 186–191 (2019).
4. Turner, M. D., Nedjai, B., Hurst, T. & Pennington, D. J. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochim. Biophys. Acta* **1843**, 2563–2582 (2014).
5. Waldmann, T. A. Cytokines in Cancer Immunotherapy. *Cold Spring Harb. Perspect. Biol.* **10**, (2018).
6. Hutmacher, C. & Neri, D. Antibody-cytokine fusion proteins: Biopharmaceuticals with immunomodulatory properties for cancer therapy. *Adv. Drug Deliv. Rev.* **141**, 67–91 (2019).
7. Klein, C. *et al.* Cergutuzumab amunaleukin (CEA-IL2v), a CEA-targeted IL-2 variant-based immunocytokine for combination cancer immunotherapy: Overcoming limitations of aldesleukin and conventional IL-2-based immunocytokines. *Oncoimmunology* **6**, e1277306 (2017).
8. Johannsen, M. *et al.* The tumour-targeting human L19-IL2 immunocytokine: preclinical safety studies, phase I clinical trial in patients with solid tumours and expansion into patients with advanced renal cell carcinoma. *Eur. J. Cancer* **46**, 2926–2935 (2010).
9. Tzeng, A., Kwan, B. H., Opel, C. F., Navaratna, T. & Wittrup, K. D. Antigen specificity can be irrelevant to immunocytokine efficacy and biodistribution. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3320–3325 (2015).
10. Tang, A. & Harding, F. The challenges and molecular approaches surrounding interleukin-2-based therapeutics in cancer. *Cytokine: X* **1**, 100001 (2019).
11. Garcin, G. *et al.* High efficiency cell-specific targeting of cytokine activity. *Nat. Commun.* **5**, 3016 (2014).
12. Gillies, S. D. *et al.* A low-toxicity IL-2-based immunocytokine retains antitumor activity despite its high degree of IL-2 receptor selectivity. *Clin. Cancer Res.* **17**, 3673–3685 (2011).
13. Sun, Z. *et al.* A next-generation tumor-targeting IL-2 preferentially promotes tumor-infiltrating CD8 T-cell response and effective tumor control. *Nat. Commun.* **10**, 3874 (2019).
14. Wüest, T. *et al.* TNF-Selectokine: a novel prodrug generated for tumor targeting and site-specific activation of tumor necrosis factor. *Oncogene* **21**, 4257–4265 (2002).
15. Puskas, J. *et al.* Development of an attenuated interleukin-2 fusion protein that can be activated by tumour-expressed proteases. *Immunology* **133**, 206–220 (2011).
16. Skrombolas, D., Sullivan, M. & Frelinger, J. G. Development of an Interleukin-12 Fusion Protein That Is Activated by Cleavage with Matrix Metalloproteinase 9. *J. Interferon Cytokine Res.* **39**, 233–245 (2019).
17. Venetz, D., Koovely, D., Weder, B. & Neri, D. Targeted Reconstitution of Cytokine Activity upon Antigen Binding using Split Cytokine Antibody Fusion Proteins. *J. Biol. Chem.* **291**, 18139–18147 (2016).
18. Mock, J. *et al.* An engineered 4-1BBL fusion protein with 'activity on demand'. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 31780–31788 (2020).
19. Taverna, D. M. & Goldstein, R. A. Why are proteins marginally stable? *Proteins* **46**, 105–109 (2002).

20. Baker, D. What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683 (2019).
21. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
22. Quijano-Rubio, A., Ulge, U. Y., Walkey, C. D. & Silva, D.-A. The advent of de novo proteins for cancer immunotherapy. *Curr. Opin. Chem. Biol.* (2020) doi:10.1016/j.cbpa.2020.02.002.
23. Zahnd, C., Pecorari, F., Straumann, N., Wyler, E. & Plückthun, A. Selection and characterization of Her2 binding-designed ankyrin repeat proteins. *J. Biol. Chem.* **281**, 35167–35175 (2006).
24. Boersma, Y. L., Chao, G., Steiner, D., Wittrup, K. D. & Plückthun, A. Bispecific designed ankyrin repeat proteins (DARPins) targeting epidermal growth factor receptor inhibit A431 cell proliferation and receptor recycling. *J. Biol. Chem.* **286**, 41273–41285 (2011).
25. Lajoie, M. J. *et al.* Designed protein logic to target cells with precise combinations of surface antigens. *Science* **369**, 1637–1643 (2020).
26. Dougan, S. K. *et al.* Transnuclear TRP1-specific CD8 T cells with high or low affinity TCRs show equivalent antitumor activity. *Cancer Immunol Res* **1**, 99–111 (2013).
27. Dougan, M. *et al.* Targeting Cytokine Therapy to the Pancreatic Tumor Microenvironment Using PD-L1-Specific VHHs. *Cancer Immunol Res* **6**, 389–401 (2018).
28. Ingram, J. R. *et al.* PD-L1 is an activation-independent marker of brown adipocytes. *Nat. Commun.* **8**, 647 (2017).
29. Xie, Y. J. *et al.* Nanobody-based CAR T cells that target the tumor microenvironment inhibit the growth of solid tumors in immunocompetent mice. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7624–7631 (2019).
30. Shen, S. *et al.* Engineered IL-21 Cytokine Muteins Fused to Anti-PD-1 Antibodies Can Improve CD8+ T Cell Function and Anti-tumor Immunity. *Front. Immunol.* **11**, 832 (2020).
31. Sockolosky, J. T. *et al.* Selective targeting of engineered T cells using orthogonal IL-2 cytokine-receptor complexes. *Science* **359**, 1037–1042 (2018).
32. Klatzmann, D. & Abbas, A. K. The promise of low-dose interleukin-2 therapy for autoimmune and inflammatory diseases. *Nat. Rev. Immunol.* **15**, 283–294 (2015).
33. Rashidian, M. *et al.* Predicting the response to CTLA-4 blockade by longitudinal noninvasive monitoring of CD8 T cells. *J. Exp. Med.* **214**, 2243–2255 (2017).
34. Salter, A. I., Pont, M. J. & Riddell, S. R. Chimeric antigen receptor-modified T cells: CD19 and the road beyond. *Blood* **131**, 2621–2629 (2018).
35. Martinez, M. & Moon, E. K. CAR T Cells for Solid Tumors: New Strategies for Finding, Infiltrating, and Surviving in the Tumor Microenvironment. *Front. Immunol.* **10**, (2019).
36. Porter, D. L. *et al.* Chimeric antigen receptor T cells persist and induce sustained remissions in relapsed refractory chronic lymphocytic leukemia. *Sci. Transl. Med.* **7**, 303ra139 (2015).
37. Ahmed, N. *et al.* Human Epidermal Growth Factor Receptor 2 (HER2) -Specific Chimeric Antigen Receptor-Modified T Cells for the Immunotherapy of HER2-Positive Sarcoma. *J. Clin. Oncol.* **33**, 1688–1696 (2015).

38. Kershaw, M. H. *et al.* A phase I study on adoptive immunotherapy using gene-modified T cells for ovarian cancer. *Clin. Cancer Res.* **12**, 6106–6115 (2006).
39. Stefan, N. *et al.* DARPins recognizing the tumor-associated antigen EpCAM selected by phage and ribosome display and engineered for multivalency. *J. Mol. Biol.* **413**, 826–843 (2011).
40. Hu, W., Wang, G., Huang, D., Sui, M. & Xu, Y. Cancer Immunotherapy Based on Natural Killer Cells: Current Progress and New Opportunities. *Front. Immunol.* **10**, 1205 (2019).
41. Ghasemi, R. *et al.* Selective targeting of IL-2 to NKG2D bearing cells for improved immunotherapy. *Nat. Commun.* **7**, 12878 (2016).
42. Raffin, C., Vo, L. T. & Bluestone, J. A. T reg cell-based therapies: challenges and perspectives. *Nat. Rev. Immunol.* **20**, 158–172 (2019).
43. Spangler, J. B. *et al.* Engineering a Single-Agent Cytokine/Antibody Fusion That Selectively Expands Regulatory T Cells for Autoimmune Disease Therapy. *J. Immunol.* **201**, 2094–2106 (2018).
44. Williams, J. Z. *et al.* Precise T cell recognition programs designed by transcriptionally linking multiple receptors. *Science* **370**, 1099–1104 (2020).

Authors and Acknowledgements

Authors: Alfredo Quijano-Rubio[1,2], Aladdin M. Bhuiyan[3,4,†], Huilin Yang[5,†], Isabel Leung[6,†], Elisa Bello[3,†], Lestat R. Ali[3], Kevin Zhangxu[4], Wentao Wang[7], Marc Lajoie[1], Rashmi Ravichandran[1], Yun-Huai Kuo[5], Stephanie K. Dougan[3,8], Stanley R. Riddell[6], Jamie B. Spangler*[,5,7], Michael Dougan*[,4], Daniel-Adriano Silva*[,1], David Baker*[,1,9]

Affiliations:
[1] Department of Biochemistry and Institute for Protein Design, University of Washington, Seattle, Washington 98195, USA
[2] Department of Bioengineering, University of Washington, Seattle, Washington 98195, USA
[3] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA, USA
[4] Department of Medicine, Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
[5] Department of Chemical & Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland, USA
[6] Fred Hutchinson Cancer Research Center, Clinical Research Division, Seattle, WA, USA
[7] Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, USA
[8] Department of Immunology, Harvard Medical School, Boston, MA, USA
[9] Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA

*Co-corresponding authors
[†]These authors contributed equally to this work

Author contributions

A.Q.-R, D.-A.S., J.B.S., M.D., and D.B. designed the research; A.Q.-R, D.-A.S., and D.B. designed the split proteins; A.Q.-R. characterized binding and stability. H.Y., W.W. and Y.-H.K. performed in vitro cell signaling and in vitro trans-activation assays with K562 cells. A.Q.-R. and M.J.L. performed in vitro cell targeting selectivity assays. A.M.B. optimized and performed in vitro B16 trans-activation assays; A.M.B. and L.R.A. performed in vivo safety studies, dose optimization and in vivo melanoma cancer model experiments; E.B. and K.Z. performed in vitro CD8 expansion studies and in vivo cis-activation studies in melanoma cancer models. I.L. performed in vitro and in vivo CAR-T cell experiments; A.Q.-R. and R.R. performed optimization and production of recombinant protein; S.K.D. and M.D. designed and coordinated research for in vivo melanoma cancer models; S.R.R. designed and supervised in vivo CAR-T cell experiments. A.Q.-R, D.-A.S., J.B.S., A.M.B., M.D., and D.B. wrote the manuscript; D.-A.S., J.B.S., M.D., and D.B. supervised and coordinated the overall research. All authors edited and accepted the manuscript.

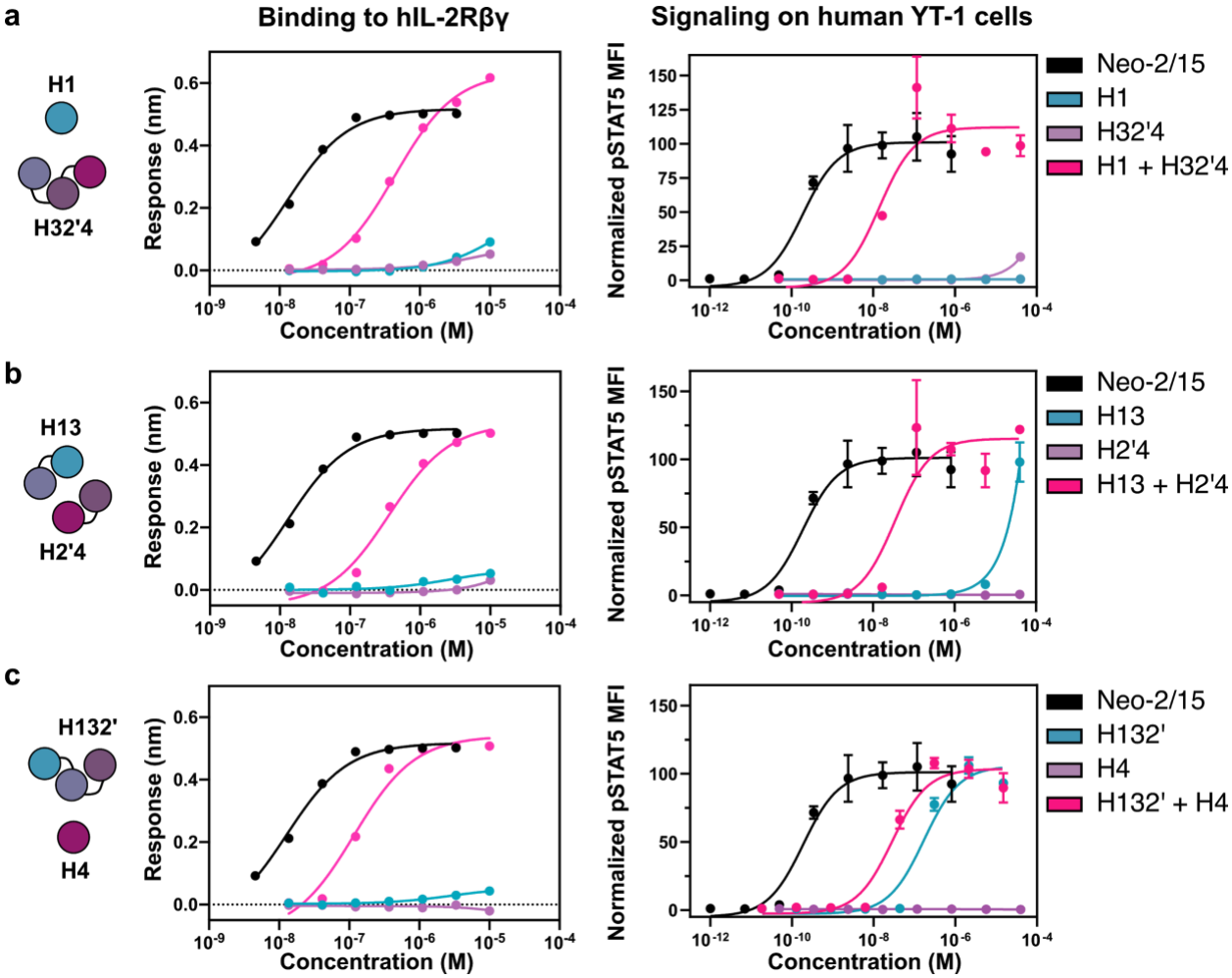
Funding

We acknowledge funding from the NIH Grant R01-CA240339-01, funding and support from the Washington Research Foundation, funding from HHMI (D.B.), The Audacious Project at the Institute for Protein Design (D.B., R.R). “la Caixa” Foundation (A.Q.-R., ID 100010434 under grant LCF/BQ/AN15/10380003), We acknowledge an Emerson Collective Cancer Research Award and a Maryland Innovation Initiative Phase I Project Award (J.B.S.). H.Y. is the recipient of a National Science Foundation Graduate Research Fellowship Program Award. S.K.D. was funded by the Melanoma Research Alliance, the Ludwig Center at Harvard, and the Claudia Adams Barr Foundation and is a Pew-Stewart Scholar in Cancer Research.

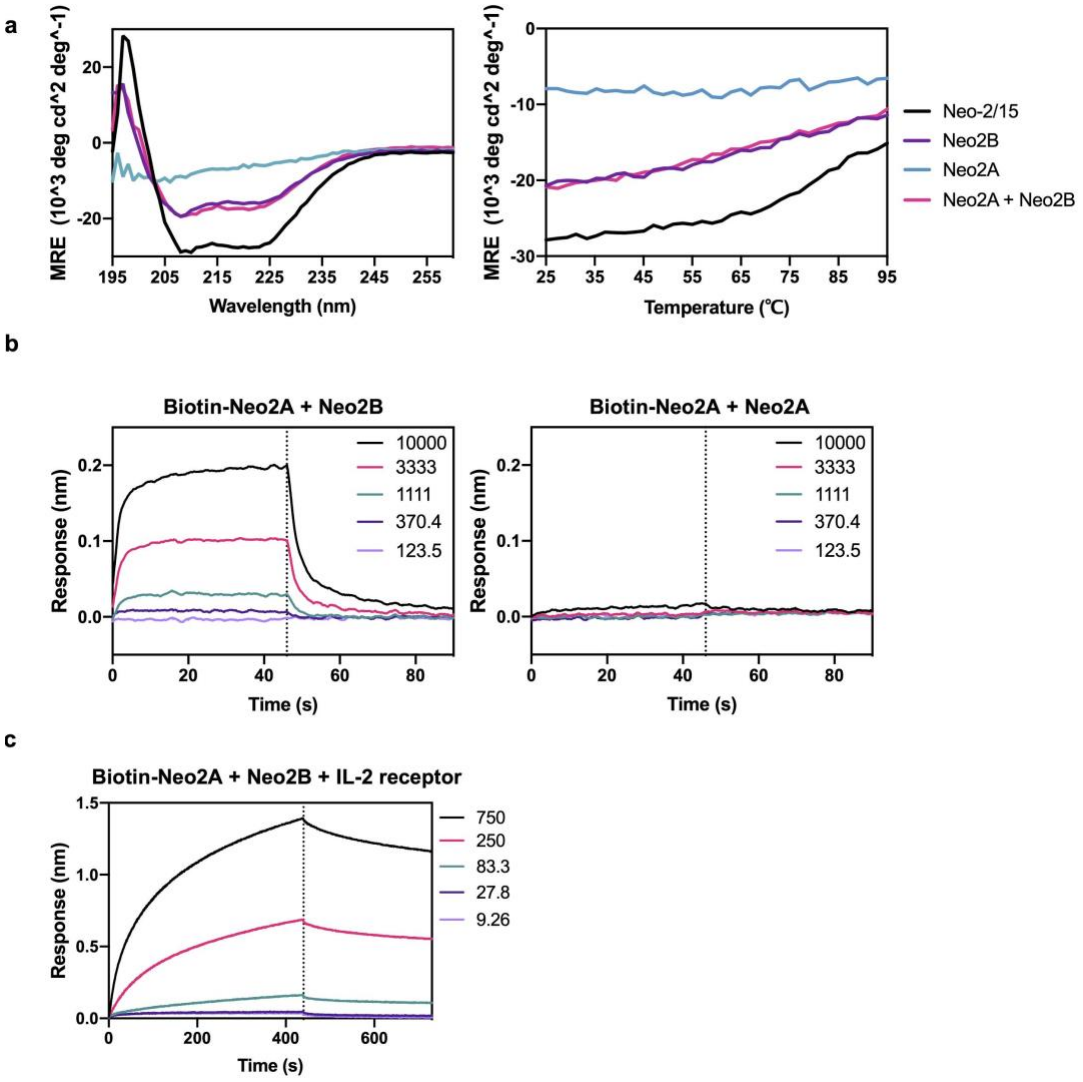
Extended Data Figures and Tables

Extended Data Table 1. Split Neoleukin-2/15 variants sequences

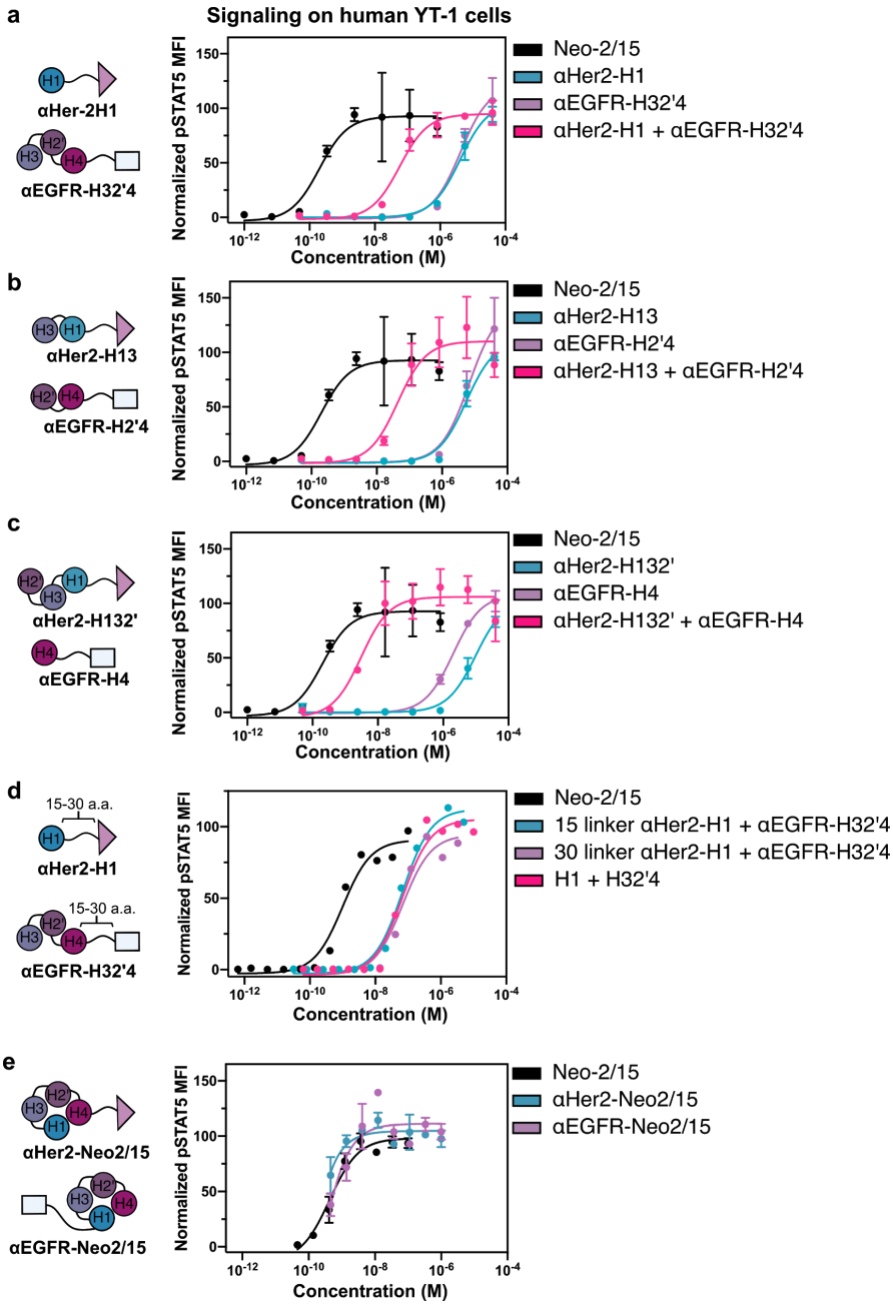
Split Neoleukin-2/15 variants	**Sequence**
H1	PKKKIQLHAEHALYDALMILNIVKTNS
H3-H2'-H4	TNSPPAEEKLEDYAFNFELILEEIARLFESGDQKDEAEKAKRMKEWM KRIKTTASEDEQEEMANAIITILQSWIFS
H1-H3-H2'	PKKKIQLHAEHALYDALMILNIVKTNSPPAEEKLEDYAFNFELILEEIA RLFESGDQKDEAEKAKRMKEWMKRIKTTAS
H4	TTASEDEQEEMANAIITILQSWIFS
H1-H3	PKKKIQLHAEHALYDALMILNIVKTNSPPAEEKLEDYAFNFELILEEIA RLFESGD
H2'-H4	DQKDEAEKAKRMKEWMKRIKTTASEDEQEEMANAIITILQSWIFS



Extended Data Figure 1. IL-2 receptor binding and immune cell signaling activity of Split Neo-2/15 pairs. Binding to immobilized IL-2Rβγ measured via biolayer interferometry (left) and STAT5 phosphorylation response (right) of YT-1 human NK cells for Neo-2/15 (black), one split fragment (blue), a second split fragment (purple) and the combination of the two split fragments (magenta). Results are shown for three split pairs: H1+H32'4 (**a**); H13+H2'4 (**b**); and H132'+H4 (**c**). The H1+H32'4 (Neo2A + Neo2B) pair demonstrates the optimal conditional activation profile. The calculated EC50 values for the STAT5 phosphorylation (**a-c**, right panels) are: Neo-2/15 0.187 nM; H1+H32'4 14 nM; H13+H2'4 34.6 nM; H132'+H4 28.3 nM; H1 >40 μM; H32'4 >40 μM; H13 >40 μM; H2'4 >40 μM; H132' 171 nM; H4 >40 μM. Experiments were performed in triplicate three times with similar results. Error bars represent s.d.

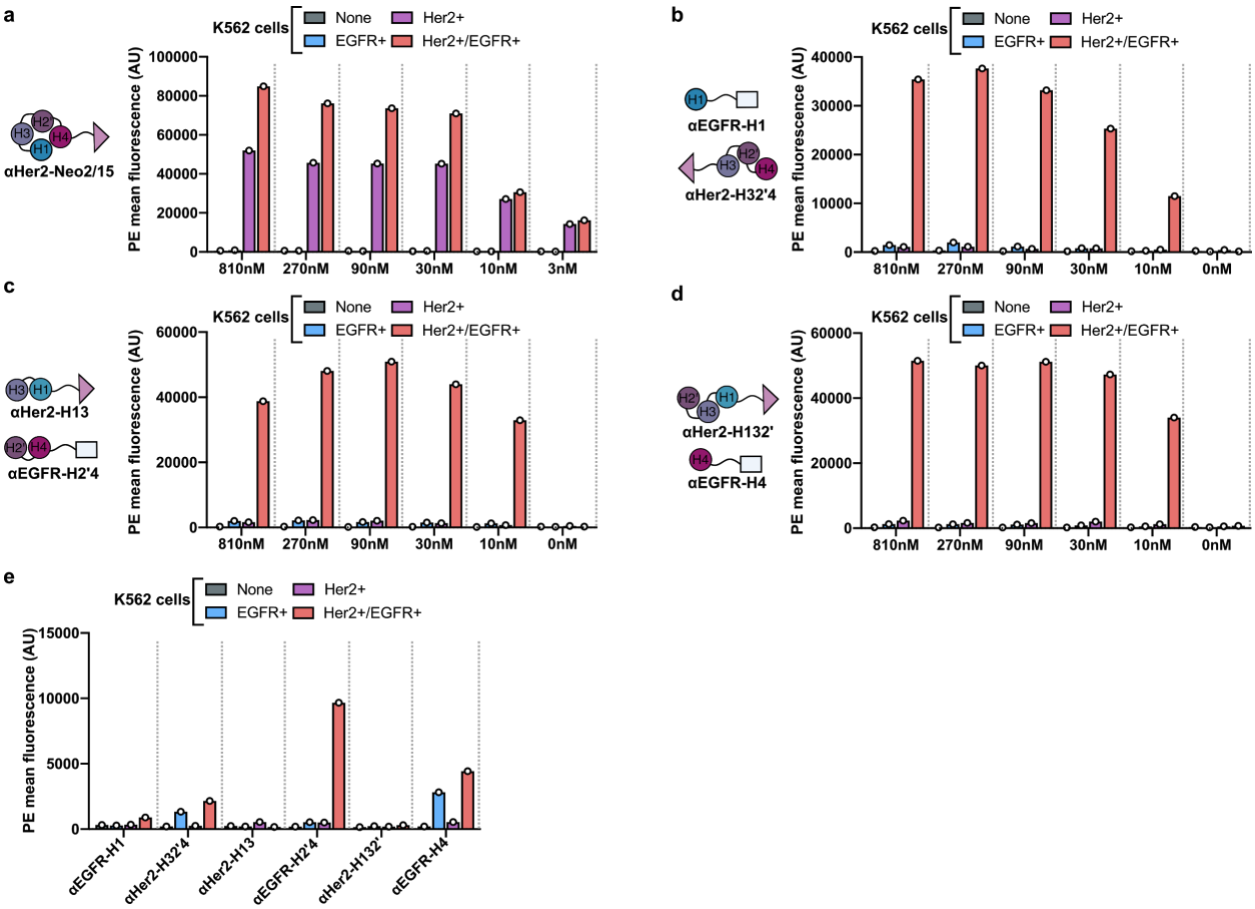


Extended Data Figure 2. Biophysical analyses of split Neo-2/15 fragments Neo2A and Neo2B. **a,** Left: Circular dichroism spectra of Neo-2/15 (black) and the Split Neo-2/15 fragments individually (blue and purple) or in combination (magenta). Neo2B shows negative ellipticity at 210 and 222 nm, indicating an alpha helical secondary structure. Neo2A does not show significant absorbance at that wavelength, indicating lack of structural organization. Right: Thermal melts of each split fragment, monitored by their signal at 222 nm during heating from 25ºC to 95ºC (heating rate ~2ºC/min). **b,** Biolayer interferometry binding analysis of Neo2A binding to Neo2B. Biotinylated-Neo2A was immobilized on streptavidin-coated tips and analyzed for binding in the presence of serial dilutions of Neo2B (left panel) or Neo2A (right panel). The calculated KD between Neo2A and Neo2B is 4.6 μM. Clored lines represent the concentration of analyte in nM. **c,** Biolayer interferometry binding analysis of Neo2A binding to Neo2B in the presence of soluble hIL-2Rβ and hIL-2Rγ. Biotin-Neo2A was bound to Streptavidin-coated tips and analyzed for binding with serially diluted Neo2B in presence of the soluble IL-2 receptor subunits hIL-2Rβ and hIL-2Rγ. The calculated KD for the complex is 50.8 nM. Colored lines represent equimolar concentration of Neo2B, hIL-2Rβ and hIL-2Rγ in nM.

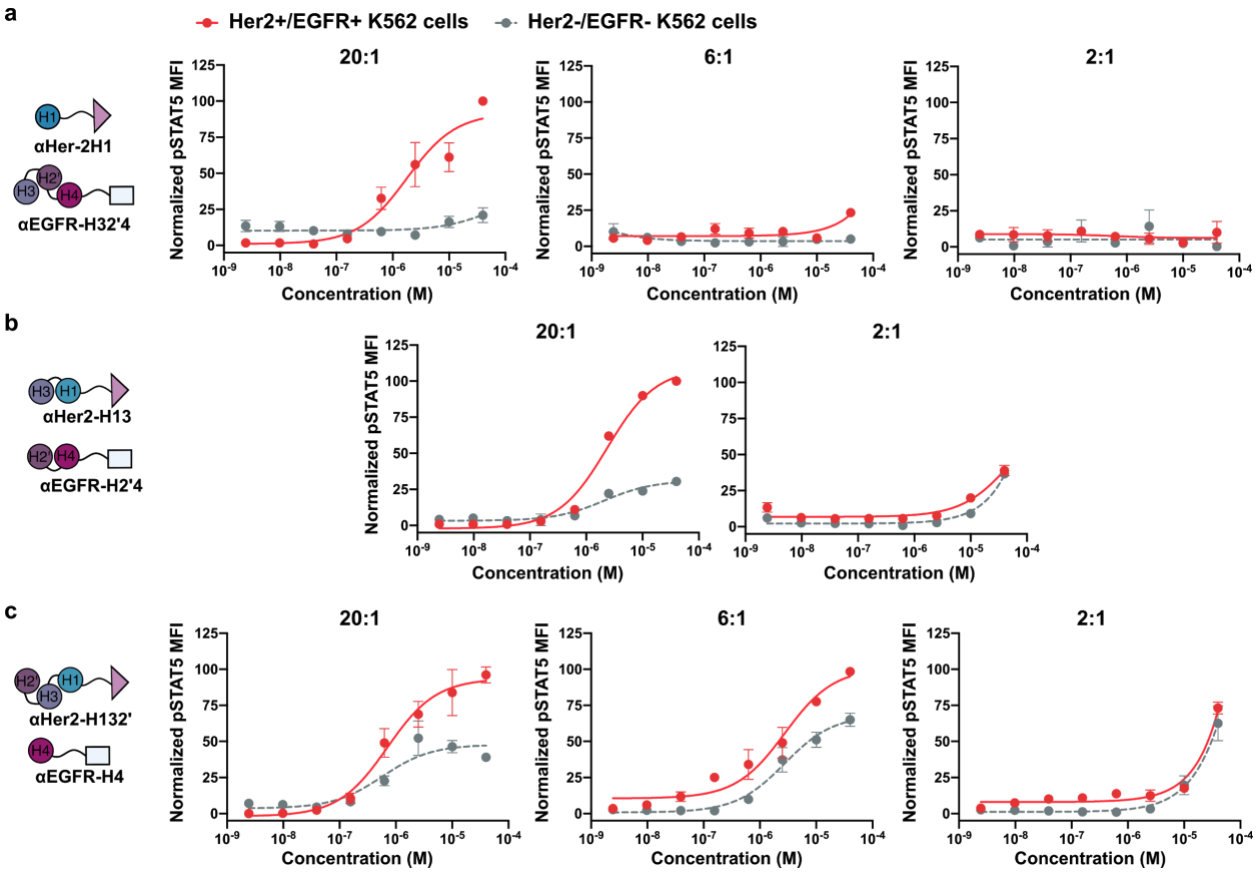


Extended Data Figure 3. Immune cell signaling activity of targeted Split Neo-2/15 pairs. **a-c,** STAT5 phosphorylation response in YT-1 human NK cells following treatment with intact Neo-2/15 or split Neo-2/15 fragments fused to Her2- or EGFR-targeted DARPins. Results are shown for three split pairs: H1+H32'4 (**a**); H13+H2'4 (**b**); and H132'+H4 (**c**). All Split Neo-2/15 variants remain functional after fusion to DARPin targeting domains. The calculated EC50 values are: (**a**) Neo-2/15 0.201 nM; αHer2-H1 4.09 μM; αEGFR-H132' 4.20 μM; αHer2-H1+αEGFR-H32'4 43.9 nM; (**b**) αHer2-H13 5.51 μM; αEGFR-H2'4 6.73 μM; αHer2-H13+αEGFR-H2'4 43.9 nM; (**c**) αHer2-H132' 9.87 μM; αEGFR-H4 2.03 μM; αHer2-H132'+αEGFR-H4 3.07 nM; **d,** Increasing length of the linker separating the Neo-2/15 split fragments from the DARPin targeting domain from 15 to 30 amino acids does not affect activity. The calculated EC50 values are: Neo-2/15 0.99 nM; 15 residue linker αHer2-H1+αEGFR-H32'4 68.3 nM; 30 residue linker αHer2-H1+αEGFR-

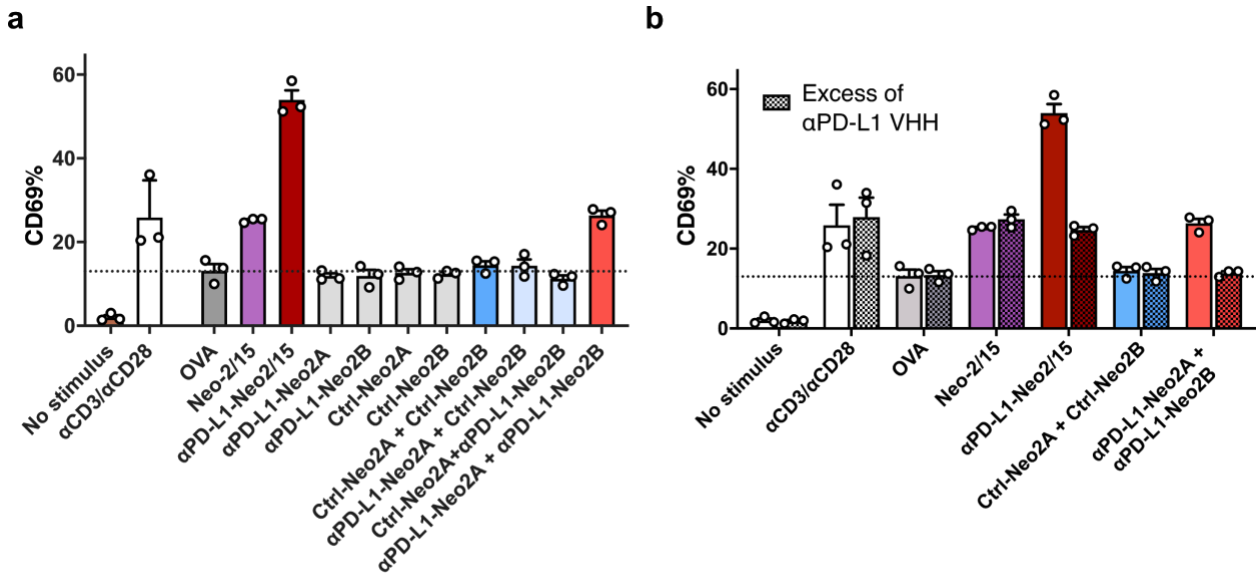
H32’4 68.1 nM; H1+H32’4 67.8 nM. **e,** Intact Neo-2/15 fused to Her2- or EGFR-targeted DARPins retains full activity, as measured by STAT5 phosphorylation response in YT-1 human NK cells. Experiments were performed in triplicate three times with similar results. Error bars represent s.d.



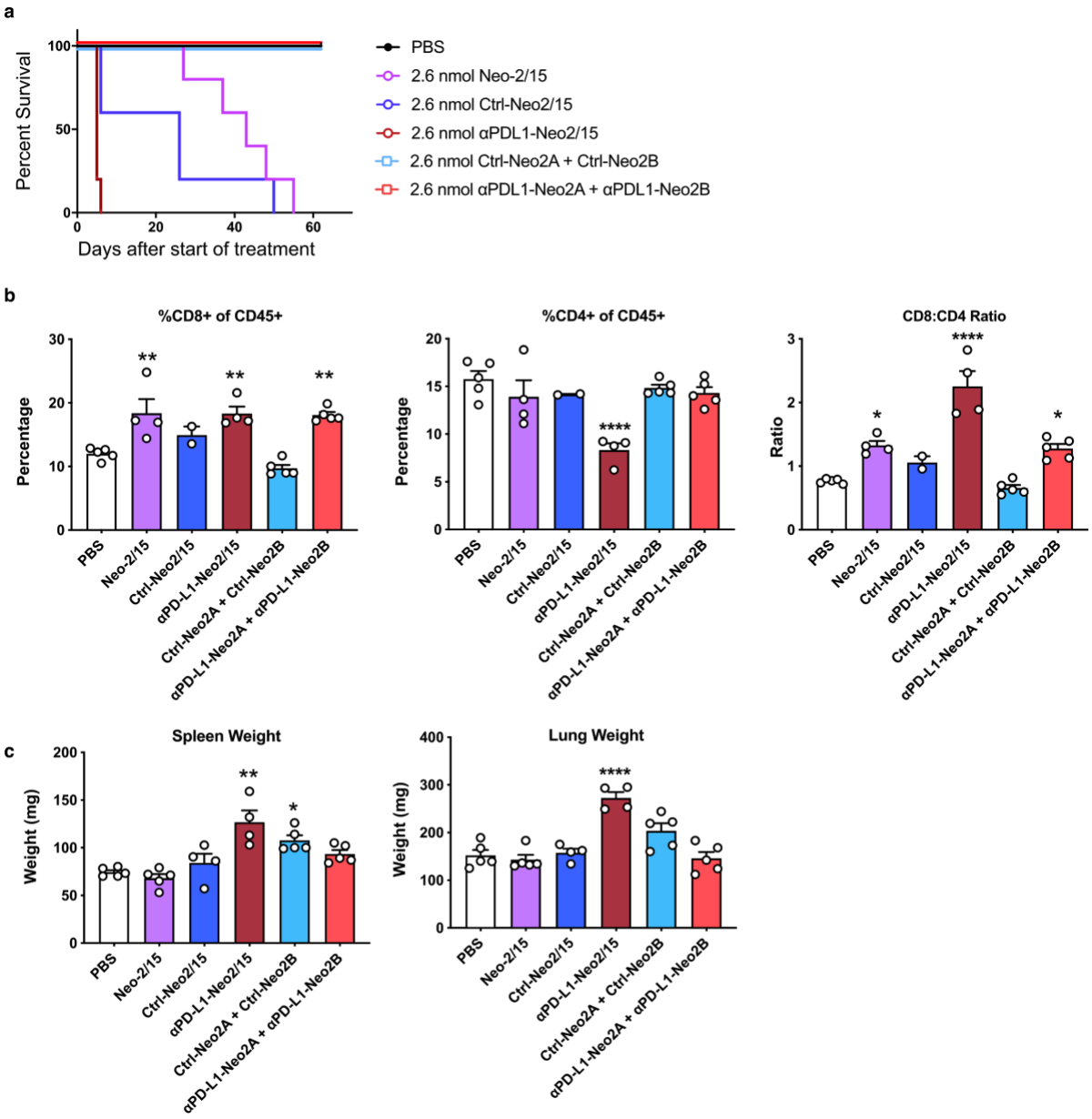
Extended Data Figure 4. Targeted reconstitution of Split Neo-2/15 targeted to the surface of transduced K562 cells. a-c, Dilution series of targeted Split Neo-2/15 variants were evaluated by flow cytometry using the experimental setup shown in Fig 2a. Reconstitution of IL-2 receptor binding activity is measured by recruitment of fluorescently-labelled hIL-2Rβγ (resulting from a mixture of IL-2Rβ, biotinylated IL-2Rγ, and PE-conjugated streptavidin). Data are shown for Her2-targeted intact Neo-2/15 (**a**) and three split pairs: H1+H32'4 (**b**); H13+H2'4 (**c**); and H132'+H4 (**d**). **e**, Activity of the individual split Neo-2/15 fragment fusion proteins (810 nM concentration) was tested to evaluate the potential for off-target activity. The dilution series were performed in a single replicate (n = 1).



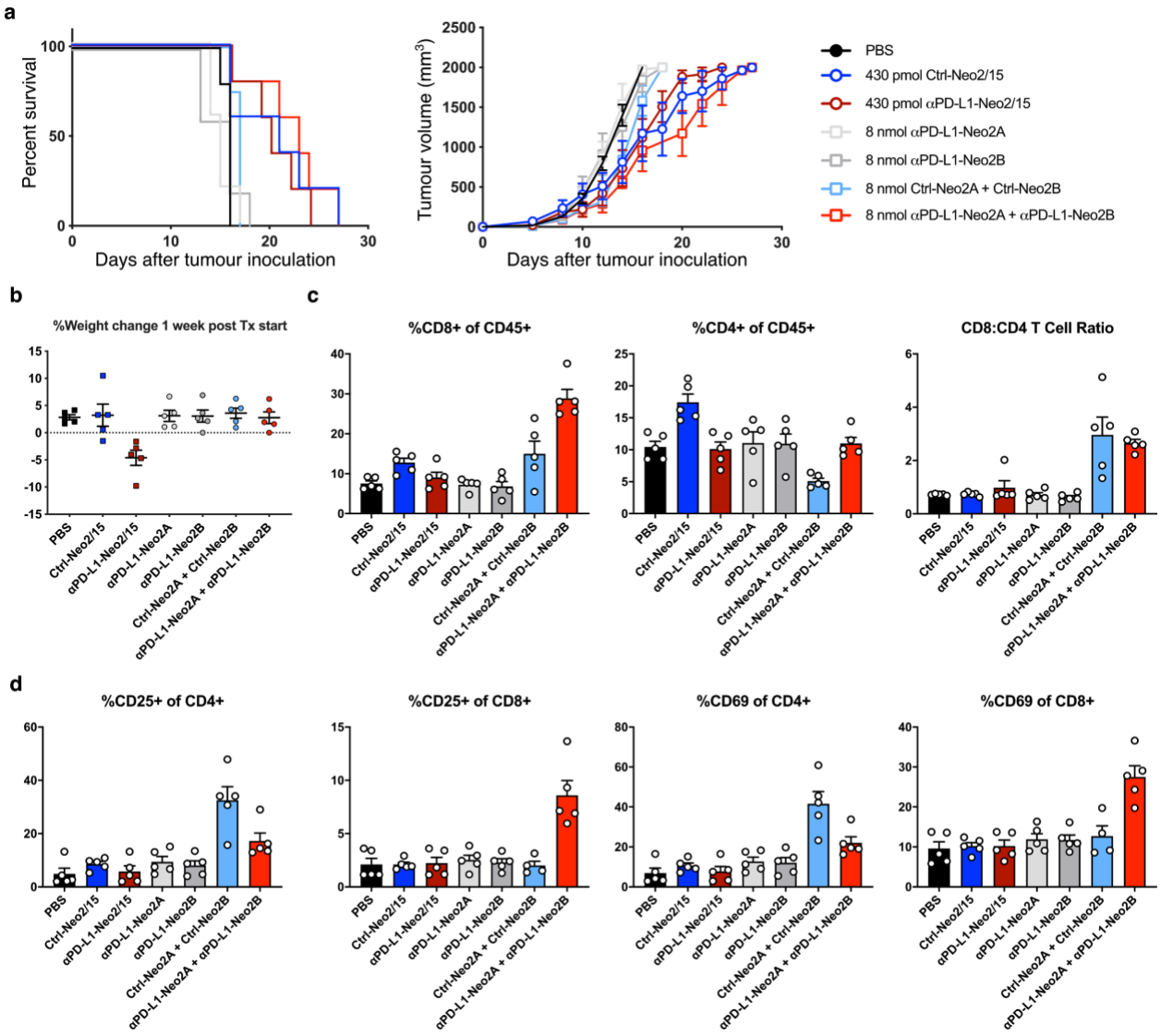
Extended Data Fig. 5. Trans-presentation of split-Neo-2/15 on the surface of K562 to YT-1 cells has limited potency in the absence of an immunological synapse. Untransduced Her2-/EGFR- K562 (off-target) cells or double-positive Her2+/EGFR+ K562 (on-target) cells were co-cultured with YT-1 human NK cells in varying K562:YT-1 cell ratios in the presence of αHer2-Neo2A and αEGFR-Neo2B or other split Neo-2/15 fragment pairs. STAT5 phosphorylation of immune cells was observed for high K562:YT-1 cell ratios, demonstrating trans-activation of immune cells from the surface of target-expressing cells. The experiments were performed in triplicate three times with similar results. Error bars represent s.d.



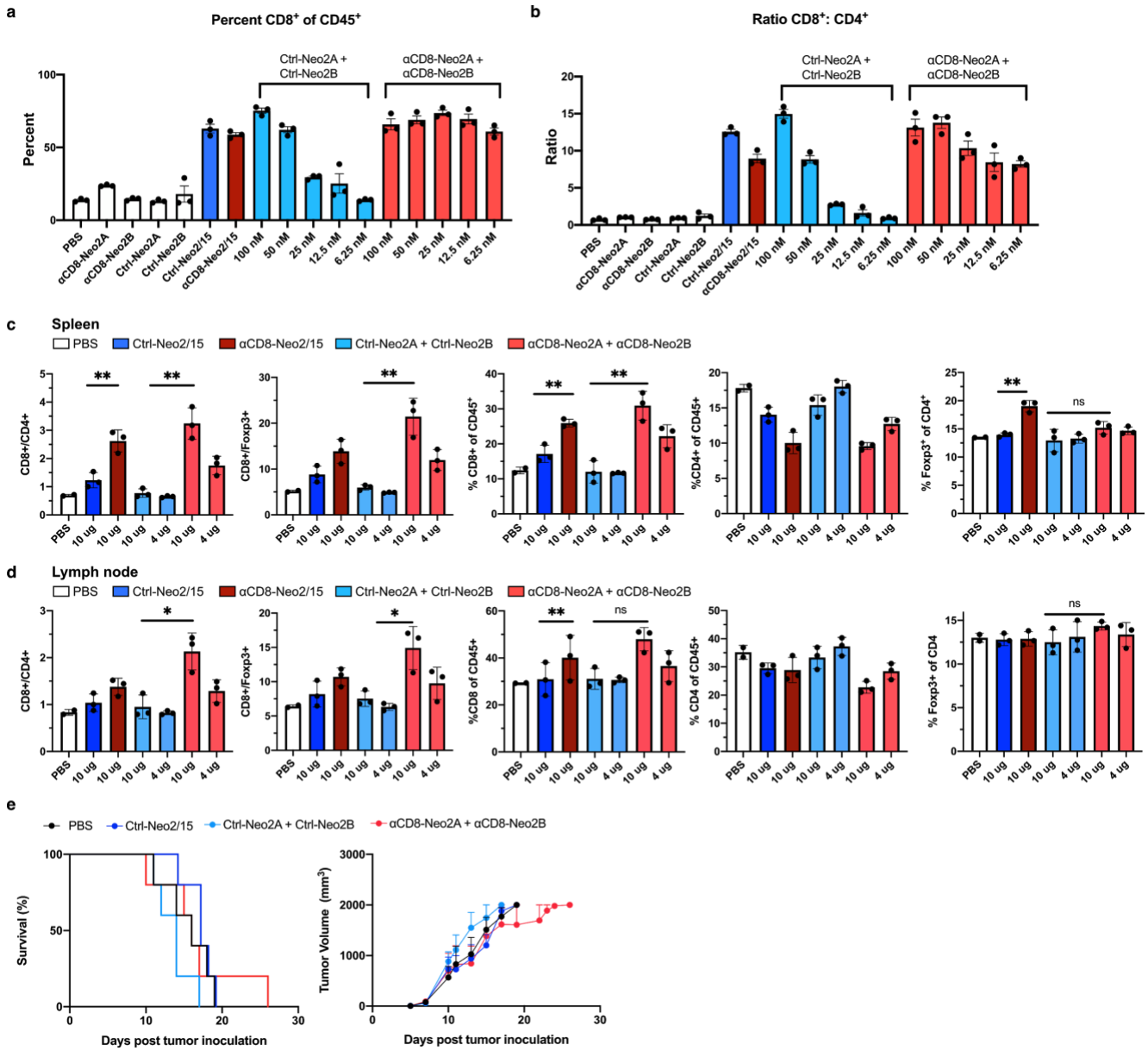
Extended Data Figure 6. CD69 upregulation data from CD8+ T cell trans-activation study shown in Fig. 2e. a, PD-L1-overexpressing B16 melanoma cells and αTrp-1 CD8+ T cells were co-cultured in the presence of the listed activating proteins at 1.0 μM concentrations. CD69 expression was measured to assess T cell activation. αCD3 and αCD28 antibodies were used as positive control for T cell activation. OVA peptide was used to quantify basal CD69 expression levels in co-culture conditions. All samples were incubated with an αCD28 antibody to provide a co-stimulatory signal. **b**, Addition of excess soluble αPD-L1 nanobody (VHH) to competitively inhibit binding of targeted split fusion proteins to the B16 cell surface reduced T cell activation, confirming trans-activation of immune cells from the surface of tumor cells. The experiments were performed in triplicate three times with similar results. Error bars represent s.d.



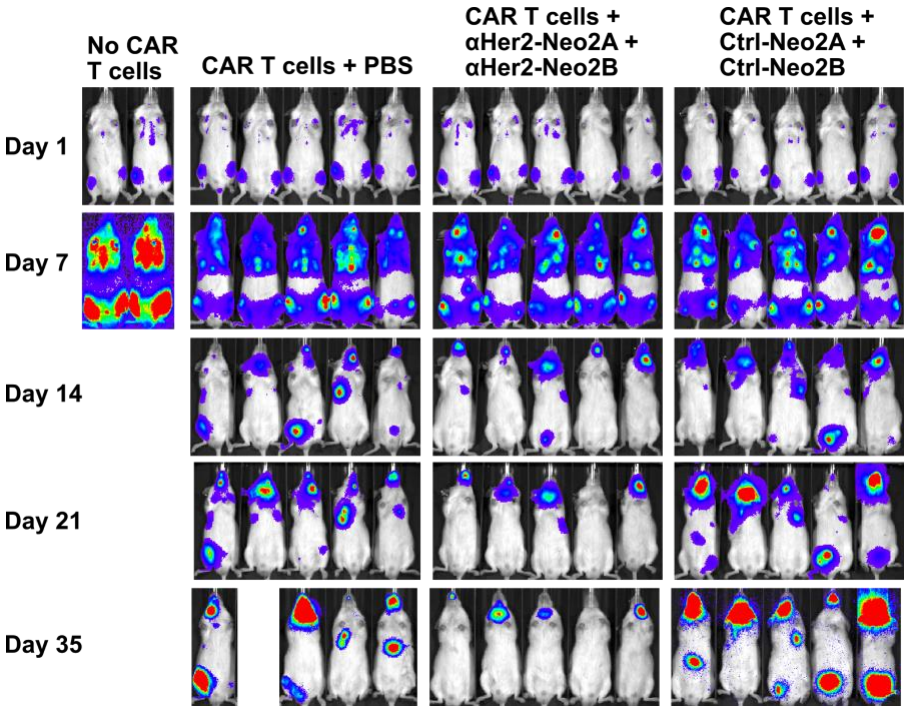
Extended Data Figure 7. Supplementary data for in vivo safety study presented in Figure 3a. Immunocompetent C57BL/6J mice (n=5/group) were treated with targeted Neo-2/15 and targeted Split Neo-2/15. Mice were treated daily with equivalent doses of the indicated proteins (2.6 nmol). Spleen and lungs were harvested upon euthanasia. Organs from mice in the Neo-2/15, Ctrl-Neo2/15, and αPD-L1-Neo2/15 cohorts were obtained and analyzed on different days, when the euthanasia criteria were met. Organs from the mice in the PBS and split cohorts were obtained on day 60, at the conclusion of the study. **a,** Survival curve**. b,** subjected to immunophenotyping by flow cytometry to quantify expansion of CD8+ T cells. **c,** The spleens and lungs were weighed to assess potential toxicity of treatment. Wet lung masses are reported. * indicates P<0.05; ** indicates P<0.01, **** indicates P<0.001 , unpaired Student's t-test. Error bars represent s.d.



Extended Data Figure 8. Supplementary data and peripheral blood analysis for the efficacy study presented in Figure 3b. C57BL/6J mice (n=5/group) bearing PD-L1-overexpressing B16 melanoma tumors in the flank were treated with the Neo-2/15 and Split Neo-2/15 fusion proteins as indicated**. a,** This efficacy experiment was carried out identically (and in parallel) to the efficacy experiment shown in Figure 3b. Data shown here were collected from mice that were not co-treated with Ta99. **b,** Weight change after one week of treatment for mice in the efficacy study shown in Figure 3b**. c,** Peripheral blood cell analysis of mice from the study shown in Figure 3b. Peripheral blood was collected at day 15 of treatment and analyzed by flow cytometry. Relative expansion of CD8+ T cells was measured to quantify *in vivo* activity of the Neo-2/15 and Split Neo-2/15 constructs. **d,** Expression of surface CD25 and CD69 on cells analyzed in (**c**) was measured to evaluate Split Neo-2/15-mediated T cell activation. Error bars represent s.d.



Extended Data Figure 9. CD8-targeted Split Neo-2/15 mediated T cell cis-activation. a-b, Splenocytes from WT B6 mice were cultured with the indicated Neo-2/15 and Split Neo-2/15 fusion proteins. After four days, cells were harvested and analyzed by flow cytometry. The change in the ratio of CD8:CD4 T cells was quantified to measure selective expansion of CD8+ T cells by the targeted constructs (**b)**. **c-d,** Healthy FoxP3-GFP mice were dosed daily with the indicated constructs. After 5 days, the mice were euthanized and their spleens and lymph nodes were harvested. The CD8:CD4 ratios in the spleen cells (**c**) and lymph node cells (**d**) were quantified by flow cytometry to assess the extent of selective expansion of CD8+ T cells in vivo. **e,** Anti-tumor efficacy of CD8-specific Split Neo-2/15 in a syngeneic mouse model of B16 melanoma. This efficacy experiment was carried out with the same conditions and in parallel to the study shown in Fig. 4c. Mice shown here were not co-treated with Ta99. * indicates P<0.05; ** indicates P<0.01, unpaired Student's t-test. Error bars represent s.d.



Extended Data Figure 10. Tumor growth following co-therapy with anti-CD19 CAR T cells and targeted Split Neo-2/15 in a lymphoma model. NSG mice inoculated with $0.5x10^6$ RAJI tumor cells were treated with $0.8x10^6$ anti-CD19 CAR-T cells 7 days post-tumor inoculation. αHer2-Neo2A + αHer2-Neo2B or control αEpCam-Neo2A + αEpCam-Neo2B were injected intraperitoneally at 7.5mg/kg daily from day 1 to 3, day 6 to 10 and day 13 to 15 after CAR-T cell injection. RAJI tumor cells were transduced with (ffLuc)-eGFP to assess tumor growth via bioluminescence imaging. The experiment was performed once.

Methods

Synthetic gene construction

The designed protein sequences were codon optimized for E. coli expression and ordered as synthetic genes in pET29b+ E. coli expression vectors (Genscript). The synthetic gene was inserted at the NdeI and XhoI sites of each vector, including an N-terminal hexahistidine tag followed by a TEV protease cleavage site and a stop codon was added at the C terminus. The VHH fusion proteins were also inserted in pCMV/R for mammalian cell expression cloned between NheI and AvrII sites and including an N-terminal signal peptide (MDSKGSSQKGSRLLLLLVVSNLLLPQGVLAGSDG) (Genscript).

Protein production and purification in E. coli

All DARPin protein fusions, H132' and H32'4 were expressed in the E. coli Lemo21(DE3) strain (NEB). The VHH fusion proteins were expressed in E. coli SHuffle T7 strain (NEB). Bacteria were transformed with a pET29b+ plasmid encoding the synthesized gene of interest. Cells were grown for 24 hours in LB media supplemented with kanamycin. Cells were inoculated at a 1:50 mL ratio in the Studier TBM-5052 autoinduction media supplemented with carbenicillin or kanamycin, grown at 37 °C for 2-4 hours, and then grown at 18 °C for an additional 18 h. Cells were harvested by centrifugation at 4000g at 4 °C for 15 min and resuspended in 30 ml lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 30 mM imidazole, 1 mM PMSF, 0.02 mg/mL DNAse). Cell resuspensions were lysed by sonication for 2.5 minutes (5 second cycles). Lysates were clarified by centrifugation at 24,000g at 4 °C for 20 min and passed through 2 ml of Ni-NTA nickel resin (Qiagen, 30250) pre-equilibrated with wash buffer, (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 30 mM imidazole). The resin was washed twice with 10 column volumes (CV) of wash buffer, and then eluted with 3 CV of elution buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM imidazole). The eluted proteins were concentrated using Ultra-15 Centrifugal Filter Units (Amicon) and further purified by using a SuperdexTM 75 Increase 10/300 GL (GE Healthcare) size exclusion column in Tris Buffered Saline (TBS; 25 mM Tris-HCl pH 8.0, 150 mM NaCl). Fractions containing monomeric protein were pooled, concentrated, the concentration was measured by absorbance at 280 nm (nanodrop), and snap-frozen in liquid nitrogen and stored at -80 °C. Proteins used in animal studies were further purified to remove endotoxin using NoEndoTM High Capacity Spin Columns (Protein Ark). The endotoxin levels were measured with Endosafe® LAL Cartridge, PTS201F 0.1EU/mL sensitivity (Charles River) in an Endosafe nexgen-MCS (Charles River).

Protein production in 293-6E cells

The nanobody (VHH) fusion proteins used in animal studies were expressed in 293-6E cells. Four liters of 293-6E cells were transfected using linear polyethylenimine (PEI). The cultures were harvested when the average viability was ≤90% (4-6 days). Protein expression was in the culture supernatant and total cell lysate was analyzed by reducing and non-reducing SDS-PAGE. The His-tagged proteins were purified from the clarified culture supernatant via Ni-NTA affinity chromatography. Then, the proteins were dialyzed overnight into PBS pH 7.4, concentrated using Ultra-15 Centrifugal Filter Units (Amicon) and analyzed by reducing and non-reducing SDS-PAGE and analytical SEC. Finally, the concentration was measured by absorbance at 280 nm (nanodrop) and aliquots were snap-frozen in liquid nitrogen and stored at -80 °C.

Peptide Synthesis

Peptides H1 (Neo2A), H13, H2'4 and H4 were synthesized to >85% purity by Genscript. For Biolayer interferometry assays measuring the affinity between Neo2A and Neo2B, the Neo2A peptide was biotinylated at the N-terminus and a flexible linker was added avoid steric hindrance (Biotin-GGGSGGGSPKKKIQLHAEHALYDALMILNIVKTNS)

Bio-Layer Interferometry

Binding data were collected in an Octet RED96 (ForteBio) and processed using ForteBio Data Analysis Software version 9.0.0.10. Biotinylated analyte human IL-2Rγ (Acro Biosystems Cat no. ILG-H85E8) was immobilized on streptavidin-coated biosensors (SA ForteBio) at 5 μg/ml in binding buffer (10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.05% surfactant P20, 0.5% non-fat dry milk) until reaching 0.5 nm . After loading the IL-2Rγ receptor onto the biosensor, baseline measurement was performed dipping the biosensors in binding buffer alone (60 seconds), then, the binding kinetics were monitored by dipping the biosensors in wells containing the target analyte protein (association step) and then dipping the sensors back into baseline/buffer (dissociation). For the association step, analyte proteins were diluted from concentrated stocks into binding buffer to the indicated final concentration. Human IL-2Rβ (Acro Biosystems Cat no. CD2-H5221) was added in solution at the indicated concentration. The experiments using Biotin-Neo2A were performed following the same method, except Biotin-Neo2A was loaded on the SA sensors and human IL-2Rγ-Fc (Acro Biosystems Cat no. ILG-H5256) was also added in solution during association.

Circular dichroism

Far-ultraviolet circular dichroism measurements were carried out with an AVIV spectrometer model 420 in PBS (pH 7.4) in a 1 mm path length cuvette with a protein concentration of ~0.20 mg/ml (unless otherwise mentioned in the text). Temperature melts were performed from 25 to 95 °C and monitored absorption signal at 222 nm (steps of 2 °C/min, 30 s of equilibration by step). Wavelength scans (195–260 nm) were collected at 25 °C and 95 °C, and again at 25 °C after fast refolding (~5 min).

Cell culture

YT-1 human NK and K562 human leukemia cell lines (including K562 cells transduced with EGFR-iRFP, Her2-eGFP, both, or neither [25]) were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 1% non-essential amino acids, 1 mM sodium pyruvate, 15 mM HEPES, and 1% penicillin/streptomycin (RPMI complete medium, Gibco). PD-L1-overexpressing B16 melanoma cells[29] were cultured in RPMI media supplemented with 10% FBS, 2 mM L-glutamine, 1% penicillin/streptomycin, 1% non-essential amino acids, 1 mM sodium pyruvate, (Gibco) and 0.1 mM β-mercaptoethanol. All cells were maintained at 37ºC in a humidified incubator with 5% CO_2. Murine IFNɣ was purchased from Peprotech and used at the indicated concentrations. Cells used for in vivo experiments had been passaged for less than 2 months, were negative for known mouse pathogens, and were implanted at >95% viability.

YT-1 cell STAT5 phosphorylation studies.

Approximately $2x10^5$ YT-1 cells were plated in each well of a 96-well plate and re-suspended in RPMI complete medium containing serial dilutions of targeted or untargeted intact or Split Neo-2/15 proteins. Cells were stimulated for 20 min at 37ºC and immediately fixed by addition of

formaldehyde to 1.5% and 10 min incubation at room temperature. After fixation, cells were permeabilized by resuspension in ice-cold 100% methanol for 30 min on ice. Fixed and permeabilized cells were washed twice with PBSA buffer (phosphate-buffered saline [PBS] pH 7.2 containing 0.1% bovine serum albumin) and then incubated with Alexa Fluor® 647-conjugated anti-STAT5 pY694 antibody (BD Biosciences) diluted 1:50 in PBSA buffer for 2 hr at room temperature. Cells were washed twice in PBSA buffer and Alexa Fluor® 647 was analyzed on a CytoFLEX flow cytometer (Beckman-Coulter). Dose-response curves were fitted to a logistic model and half-maximal effective concentrations (EC50s) were calculated using GraphPad Prism data analysis software after subtraction of the mean fluorescence intensity (MFI) of unstimulated cells and normalization to the maximum signal intensity. Experiments were conducted in triplicate and performed three times with similar results.

In vitro K562 cell targeting

The four K562 cell lines (Engineered K562 tumor cell lines transduced for expression of Her2+/eGFP+, EGFR+/iRFP+ or Her2+/eGFP+ and EGFR+/iRFP+ surface markers) were mixed in equivalent ratios. The cell mixtures were washed with flow buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1 mM MgCl2, 1 mM CaCl2 and 1% BSA) and aliquoted into V-bottom plates with 200,000 cells/well. Serially diluted Split Neo-2/15 fusion proteins were made from concentrated stocks, and then added to the cells in a 50 µl volume and incubated for 30 minutes. Subsequently, the cells were washed with 150 µl of flow buffer and incubated with a mixture of biotinylated human IL-2Rγ (Acro Biosystems Cat no. ILG-H85E8), human IL-2Rβ (Acro Biosystems Cat no. CD2-H5221) and a fluorescent streptavidin-phycoerythrin conjugate (SA-PE, Invitrogen) for 15 minutes. Data were acquired on a LSRII or FACSCelesta (BD Biosciences).

YT-1:K562 cell trans-activation assays

Approximately $2x10^5$ YT-1 cells and the indicated ratio (K562:YT-1=20:1, 6:1, or 2:1) (either untransduced Her2-/EGFR- K562 cells or transduced double-positive Her2+/EGFR+ K562 cells) were used in each well for trans-activation studies. K562 cells were plated in each well of a 96-well plate and re-suspended in RPMI complete medium containing serial dilutions of Split Neo-2/15 proteins. Cells were incubated for 30 min at 37ºC and then washed once with PBSA buffer. YT-1 cells were added to each well and co-cultured with K562 cells for 30 min at 37ºC. Cells were stained with BV421-conjugated anti-CD132 antibody (BD Biosciences) for 30 min at 4ºC and washed once with PBSA buffer. After the wash, immediately fixed by addition of formaldehyde to 1.5% and 10 min incubation at room temperature. After fixation, cells were permeabilized by resuspension in ice-cold 100% methanol for 30 min on ice. Fixed and permeabilized cells were washed twice with PBSA buffer (phosphate-buffered saline [PBS] pH 7.2 containing 0.1% bovine serum albumin) and then incubated with Alexa Fluor® 647-conjugated anti-STAT5 pY694 antibody (BD Biosciences) diluted 1:50 in PBSA buffer for 2 hr at room temperature. Alexa Fluor® 647 fluorescence on YT-1 cells (gated based on the CD132 expression) cells was measured on a CytoFLEX flow cytometer (Beckman-Coulter). Dose-response curves were fitted to a logistic model and half-maximal effective concentrations (EC50s) were calculated using GraphPad Prism data analysis software after subtraction of the mean fluorescence intensity (MFI) of unstimulated cells and normalization to the maximum signal intensity. Experiments were conducted in duplicate and performed twice with consistent results.

In vitro T cell:B16 trans-activation experiments

B16 cells overexpressing PD-L1 were treated overnight with 20 ng/mL IFNg, washed, coated with 1.0 μM desired protein (i.e., OVA, Neo-2/15, Ctrl-Neo2A, etc.) for 10 min, washed and resuspended repeatedly, and replated in a U-bottom plate. Mouse trp1-specific CD8 T cells (JAX Stock No: 030958) were isolated via a negative isolation kit (Stemcell Cat # 19853) and plated in the listed conditions either alone, or in co-culture with B16 cells at a 1:1 ratio for 1 day. To confirm trans-activation, 1.0 μM αPD-L1 or Ctrl VHHs were added at the start of co-culture. The T cells were then harvested 24-36 hours after and analyzed by flow cytometry for expression of activation markers CD25 (Biolegend Cat # 102018) and CD69 (Biolegend Cat # 104507). T cells were gated apart from B16 cells via FSC/SSC gating and CD8 staining (Biolegend Cat # 100728).

Melanoma in vivo experiments

All animal protocols were approved by the Dana-Farber Cancer Institute Committee on Animal Care (protocols 14-019 and 14-037) and are in compliance with the NIH/NCI ethical guidelines for tumor-bearing animals. At day 0, 6-8 week-old B6 mice (JAX Stock No: 000664) were inoculated with 80-90% confluent B16 F10 cells that overexpress PD-L1 (500k cells/mouse). Beginning day 3, mice were treated twice a week with 150 μg TA99 mAb. Beginning day 5, mice (N = 5 or 10 mice/group) were treated daily with the listed proteins (i.e. Neo-2/15,αPD-L11-Neo2A, etc.). The H32'4 (Neo2B) parts were injected subcutaneously on the non-tumor bearing flank, while the H1 parts (Neo2A) were injected intraperitoneally. Native Neo-2/15 and its conjugates with 1B7 (Ctrl) and B3 (αPD-L1) VHHs were also injected intraperitoneally. Mice were euthanized if they lost 10% body weight or if their tumors ulcerated or reached 2000 mm3 in volume.

In vitro X118 experiments

Splenocytes were harvested from a WT B6 mouse and plated together with Neo2 splits in a 96-well plate in RPMI complete. After four days, cells were harvested and stained for flow cytometry. X axis denotes final molar concentration of each individual Split Neo-2/15 construct. Negative control single split constructs wells were plated at 50nM.

In vivo expansion in Foxp3-GFP mice

Split Neo-2/15 fusion proteins (or intact Neo2 constructs) were administered daily to non-tumor bearing Foxp3-GFP mice (JAX Stock No: 006772) for five days. Neo2A fusions were injected subcutaneously in the scruff of the neck, while Neo2B fusions, PBS, and intact Neo-2/15 constructs were injected intraperitoneally. Intact Neo-2/15 fusions were dosed at 500 pmol (12 μg/mouse/day). Split Neo-2/15 fusions were dosed at 500pmol (10 μg/mouse/day) and 250 pmol (4 μg/mouse/day). On day 6, mice were euthanized, spleen and both inguinal lymph nodes collected and homogenized into single cell suspensions (and spleen resuspended briefly in ACK lysis buffer to lyse red blood cells), and cell populations investigated by flow cytometry.

CAR-T cell experiments

In vitro STAT5 phosphorylation assay: Primary human CD8 T cells were obtained from healthy donors with written informed consent for research protocols approved by the Institutional Review Board of the FHCRC. Peripheral blood mononuclear cells (PBMC) were isolated using lymphocyte separation medium (Corning). T cells were isolated using EasySep CD8 negative isolation kits (Stemcell Technologies). CAR-T cells were generated by stimulating thawed CD8

with anti-CD3/CD28 Dynabeads (Gibco) at three beads to one T cell ratio in medium supplemented with 50 IU/ml (3.1 ng/ml) IL-2. Next day, T cells were transduced with lentivirus encoding the anti-CD19 CAR by spinoculation for 90mins at 800g or left untransduced. Beads were removed five days post stimulation. CAR positive cells were FACS sorted by the Her2 transduction marker on day 8. On day 10, CAR-T cells or untransduced T cells were cultured in indicated concentrations of Neo2 variants in 96 well plates for 12mins. T cells were then methanol fixed, permeabilized and stained for pSTAT5 for flow cytometry. In vivo RAJI experiment: the FHCRC Institutional Animal Care and Use Committee approved all mouse experiments. Six-to-eight-week-old NSG mice were obtained from the Jackson Laboratory. RAJI tumor cells ($0.5x10^6$) transduced with (ffLuc)-eGFP were injected into the tail veins of NSG mice. CAR-T cells were produced as described above, except that CAR-T cells were expanded as described in Terakura et al, Blood 2012 post FACS sorting. Seven days after tumor injection, lentiviral transduced anti-CD19 Her2+ CART cells ($0.8x10^6$) were infused intravenously into mice. αHer2-Neo2A + αHer2-Neo2B or αEpCam-Neo2A + αEpCam-Neo2B at 7.5mg/ kg were injected into the peritoneum on days 1-3, day 6-10 and day 13-15 post T cell injection.

Chapter 3

De novo design of modular and tunable protein biosensors

As originally published in Quijano-Rubio, A., Yeh, H.-W., Park, J., Lee, H., Langan, R. A., Boyken, S. E., Lajoie, M. J., Cao, L., Chow, C. M., Miranda, M. C., Wi, J., Hong, H. J., Stewart, L., Oh, B.-H., & Baker, D. (2021). De novo design of modular and tunable protein biosensors. Nature, 591(7850), 482–487.

Abstract

Naturally occurring protein switches have been repurposed for developing novel biosensors and reporters for cellular and clinical applications[1], but the number of such switches is limited, and engineering them is often challenging as each is different. Here, we show that a very general class of protein-based biosensors can be created by inverting the flow of information through *de novo* designed protein switches in which binding of a peptide key triggers biological outputs of interest[2]. The designed sensors are modular molecular devices with a closed dark state and an open luminescent state; binding of the analyte of interest drives switching from the closed to the open state. Because the sensor is based purely on thermodynamic coupling of analyte binding to sensor activation, only one target binding domain is required, which simplifies sensor design and allows direct readout in solution. We demonstrate the modularity of this platform by creating biosensors that, with little optimization, sensitively detect the anti-apoptosis protein Bcl-2, the IgG1 Fc domain, the Her2 receptor, and Botulinum neurotoxin B, as well as biosensors for cardiac Troponin I and an anti-Hepatitis B virus (HBV) antibody that achieve the sub-nanomolar sensitivity necessary to detect clinically relevant concentrations of these molecules. Given the current need for diagnostic tools for tracking COVID-19[3], we used the approach to design sensors of antibodies against SARS-CoV-2 protein epitopes and of the receptor-binding domain (RBD) of the SARS-CoV-2 Spike protein. The latter, which incorporates a *de novo* designed RBD binder[4], has a limit of detection of 15 pM and a signal over background of over 50-fold. The modularity and sensitivity of the platform should enable the rapid construction of sensors for a wide range of analytes and highlights the power of *de novo* protein design to create multi-state protein systems with new and useful functions.

Main text

Protein-based biosensors play important roles in synthetic biology and clinical applications, but thus far, biosensor design has been mostly limited to reengineering natural proteins[1]. However, finding analyte-binding domains that undergo sufficient conformational changes is challenging, and even when available, extensive protein engineering efforts are generally required to effectively couple them to a reporter domain[5,6]. Hence it is desirable to construct modular biosensor platforms that can be easily repurposed to detect different protein targets of interest. Modular systems have been developed for detecting antibodies[7–9] and small molecules[10,11], but systems for detecting proteins with very different structures, sizes and oligomerization states using semisynthetic protein platforms[12–14] or based on calmodulin switches[15,16], usually require considerable screening to find potential candidates due to limited predictability[17].

A protein biosensor can be constructed from a system with two nearly isoenergetic states, the equilibrium between which is modulated by the analyte being sensed. Desirable properties in such a sensor are (i) the analyte triggered conformational change should be independent of the details of the analyte, so the same overall system can be used to sense many different targets, (ii) the system should be tunable so that analytes with different binding energies and at relevant concentrations can be detected over a large dynamic range, and (iii) the conformational change should be coupled to a sensitive output. We hypothesized that these attributes could be attained by inverting the information flow in *de novo* designed protein switches in which binding to a target protein of interest is controlled by the presence of a peptide actuator[2]. We developed a system consisting of two protein components: (a) a 'lucCage' comprising a cage domain and a latch domain containing a target binding motif and a split luciferase fragment (SmBiT 114[18]), and (b) a "lucKey", containing a key peptide which binds to the open state of lucCage and the complementary split luciferase fragment (LgBit 11S[18], Fig. 1a). lucCage has two states: a closed state in which the cage domain binds the latch and sterically occludes the binding motif from binding target and SmBiT from combining with LgBit to reconstitute luciferase activity, and an open state in which these binding interactions are not blocked, and lucKey can bind the cage domain. Association of lucKey with lucCage results in the reconstitution of luciferase activity (Fig. 1a, right). The thermodynamics of the system are tuned such that the binding free energy of lucKey to lucCage (ΔG_{CK}) is insufficient to overcome the free energy cost of lucCage opening (ΔG_{open}) in the absence of target ($\Delta G_{open} - \Delta G_{CK} >> 0$), but in the presence of the target, the additional binding free energy of the latch to the target (ΔG_{LT}) drives latch opening and luciferase reconstitution ($\Delta G_{open} - \Delta G_{CK} - \Delta G_{LT} << 0$) (Fig. 1b,c). This system satisfies properties (i) and (ii) above, as a wide range of binding activities can be caged, and since the switch is thermodynamically controlled, the lucKey and target binding energies can be adjusted to achieve activation at the relevant target concentrations. Because lucKey and lucCage are always the same, the system is modular: the same molecular association can be coupled to the binding of many different targets. Bioluminescence provides a rapid and sensitive readout of analyte driven lucCage-lucKey association, satisfying property (iii).

The states of this biosensor system are in thermodynamic equilibrium, with the tunable parameters ΔG_{open} and ΔG_{CK} governing the populations of the possible species, along with the free energy of association of the analyte to the binding domain ΔG_{LT} (Fig. 1b). We simulated the dependence of the sensor system on ΔG_{open} (Extended Data Fig. 1a), ΔG_{LT} (Extended Data Fig. 1b), and the concentration of analyte and the sensor components (Extended Data Fig. 1c,d). The sensitivity of analyte detection is a function of ΔG_{LT}, with a lower limit of roughly one-tenth the K_D for analyte

binding (Extended Data Fig. 1b). Above this lower limit, varying the concentration of lucCage and lucKey enables responding to different target concentration ranges (Extended Data Fig. 1c,d). Sensitivity can be further modulated by tuning the strength of the intramolecular cage-latch interaction and the intermolecular cage-key interaction (ΔG_{open} and ΔG_{CK}): for example too tight cage-latch interaction results in low signal in the presence of target, and too weak an interaction results in high background in the absence of target (Extended Data Fig. 1a,e). Our design strategy aims to find this balance through modulating ΔG_{open} and ΔG_{CK} by varying the length of the latch (and key) helix and by introducing either favorable hydrophobic or unfavorable buried polar interactions at the cage-latch/key interface[2] (Extended Data Fig. 1f,g).

Designing tunable lucCage sensors

To design sensors based on these principles, we developed a "GraftSwitchMover" Rosetta-based method to identify placements of target binding peptides within the latch such that the resulting protein is stable in the closed state and the interactions with the target are blocked (see Supplementary methods). As a first test, we grafted the SmBiT peptide and the Bim peptide in the closed state of the optimized asymmetric LOCKR switch described in Langan et al,[2] (Extended Data Fig. 2). SmBiT adopts a β-strand conformation within the luciferase holoenzyme, but we assumed that it could adopt a helical secondary structure in the context of the helical bundle scaffold, since secondary structure can be context dependent[19]. We sampled different threadings for the two peptide sequences across the latch, selected the lowest energy solutions (Extended Data Fig. 2a) and expressed twelve designs in *E. coli*. We mixed the designs with lucKey in a 1:1 ratio, then added Bcl-2, which binds with nanomolar affinity to Bim[20], and observed a rapid increase in luminescence (Extended Data Fig. 2b,f; we refer to the best of these as lucCageBim), showing that the LOCKR actuator[2] operated in reverse can function as a biosensor. The analyte detection range could be tuned by varying the concentration of the sensor (lucCage + lucKey) (Extended Data Fig. 2g) as anticipated in our model simulations (Extended Data Fig. 1c). lucCageBim has SmBiT at position 312 in the latch (SmBiT312; Extended Data Fig. 2d). The Cage with this placement ("lucCage") was used as the base scaffold for the biosensors described below.

lucCage sensors with miniprotein sensing domains

We next investigated the incorporation of a range of binding modalities for analytes of interest within lucCage by developing methods for computationally caging target-binding proteins, rather than peptides, in the closed state (See supplementary methods). As a test case, we caged the *de novo* designed Influenza A H1 hemagglutinin (HA)[21] binding protein HB1.9549.2 into a shortened version of the LOCKR switch[22] (sCage), optimized to improve stability and facilitate crystallization efforts (Fig. 2a). Two of five designs were functional, and bound HA in the presence but not the absence of key (Extended Data Fig. 3b). The crystal structure of the best design, sCageHA_267-1S, determined to 2.0 Å resolution (Table S1, PDB ID: 7CBC), showed that all HA-binding interface residues except one (F273) interact with the cage domain (blocking binding of the latch to the target) as intended by design (Fig. 2a, Extended Data Fig. 3a-c).

With this structural validation of the design concept, we next sought to develop sensors for Botulinum neurotoxin B (BoNT/B), the immunoglobulin Fc domain, and the Her2 receptor. To do so, we grafted a *de novo* designed binder for Botulinum neurotoxin (Bot.0671.2)[21], the C domain of the generic antibody binding protein Protein A[23], and a Her2-binding affibody [24], into lucCage. After screening a few designs for each target (Extended Data Fig. 4-5), we obtained highly sensitive lucCages (lucCageBot, lucCageProA, and lucCageHer2) that can detect BoNT/B (Fig.

2b, Extended Data Fig. 4), human IgG Fc domain (Fig. 2c, Extended Data Fig. 5a-d), and Her2 receptor (Fig. 2d; Extended Data Fig. 5e-h) respectively, demonstrating the modularity of the platform. The designed sensors respond within minutes upon adding the target, and their sensitivity can be tuned by changing the concentration of lucCage and lucKey (Fig. 2). With further development, these sensors could enable rapid and low-cost detection of botulinum neurotoxins in the food industry[25], and detection of serological levels of soluble Her2 (>15 ng/mL; within the detection range of lucCageHer2) associated with metastatic breast cancer[26].

lucCage sensor for cardiac troponin

We next designed sensors for cardiac troponin I (cTnI), which is the standard early diagnostic biomarker for acute myocardial infarction (AMI)[27]. We took advantage of the high-affinity interaction between cTnT, cTnC, and cTnI (Fig. 3a) and designed eleven biosensor candidates by inserting 6 truncated cTnT sequences at different latch positions (Extended Data Fig. 6a). The best candidate, lucCageTrop627, was able to detect cTnI but not at sufficiently low levels for clinical use because the rule-in and rule-out levels of cTnI assay for diagnosis of AMI in patients are in the low pM range[27]. As noted above, the limit of detection (LOD) of our sensor platform is about 0.1 x *K*d of the latch-target affinity (K_{LT}), we further increased the affinity of our sensor to cTnI by fusing cTnC to its terminus (Extended Data Fig. 6b-d). The resulting sensor, lucCageTrop, has a single-digit pM LOD suitable for quantification of clinical samples (Fig. 3b, Extended Data Fig. 6e,f).

lucCage sensors for anti-HBV and anti-SARS-CoV-2 antibodies

Detection of specific antibodies is important for monitoring the spread of a pathogen in a population[28], the success of vaccination[29], and levels of therapeutic antibodies[9]. To adapt our system for antibody serological analyses, we sought to incorporate linear epitopes recognized by the antibodies of interest into lucCage. We first developed a sensor for antibodies against the PreS1 domain of the hepatitis B surface protein L[30]. The best of 8 designs tested, lucCageHBV had a ~150% increase in luciferase activity upon addition of the anti-HBV antibody HzKR127-3.2[31] (Extended Data Fig. 7a-d). To further improve the dynamic range and LOD of lucCageHBV (~2 nM, Extended Data Fig. 7e), a second copy of the peptide was introduced at the end of the latch to increase latch affinity with the bivalent antibody (K_{LT}) (Fig. 3c,d). The resulting design, lucCageHBVα, had a LOD of 260 pM and a dynamic range of 225% (Fig. 3e; Extended Data Fig. 7g-i), with a luminescence intensity easily detectable with a camera (Extended Data Fig. 7j). Hence the platform is applicable to detecting specific antibodies with a LOD in the range for monitoring therapeutic antibodies[32].

We next sought to use the lucCageHBV sensor to detect hepatitis B surface antigen (HBsAg). Since our sensors are under thermodynamic control, we hypothesized that the pre-assembled sensor-antibody complex would re-equilibrate in the presence of the target HBsAg protein, PreS1, with antibody redistributing to bind free PreS1 instead of the epitope on lucCageHBV (Fig. 3f). Indeed, the luminescence of lucCageHBV plus HzKR127-3.2 mixture decreased shortly upon addition of the PreS1 domain (Fig. 3g); the sensitivity of this readout enabled quantification of PreS1 concentration in a clinically relevant range[33] (Fig. 3h, Extended Data Fig. 7f).
The COVID-19 pandemic has showcased the urgent need for diagnostics for both the SARS-CoV-2 virus and antiviral antibodies[3]. To design sensors for anti-SARS-CoV-2 antibodies, we first identified from the literature highly immunogenic linear epitopes in the SARS-CoV[34,35] and

SARS-CoV-2 proteomes[36] that are not present in "common" strains of *coronaviridae*. Among these, we focused on two epitopes in the Membrane (M) and Nucleocapsid (N) proteins found to be recognized by SARS and COVID-19[35,36] patient sera for which cross-reactive animal-derived antibodies are commercially available (see Methods). We designed sensors for each epitope and identified designs that specifically responded to anti-M and anti-N antibodies (Extended Data Fig. 8a,b). These sensors reached full signal in 2-5 minutes and had a ~50-70% dynamic range in response to low nanomolar amounts of antibodies (Fig. 4a,b, Extended Data Fig. 8c,d). For robust serological analysis, generation of an expanded set of more sensitive sensors spanning multiple SARS-CoV-2 epitopes recognized will be necessary.

lucCage sensors for SARS-CoV-2 Spike protein

To create sensors capable of detecting SARS-CoV-2 viral particles directly, we integrated a *de novo* designed picomolar affinity binder to the receptor-binding domain (RBD) of the SARS-CoV-2 Spike protein named LCB1[4] into the lucCage format (Fig. 4c). Of 13 candidates tested, the best, which we refer to as lucCageRBD, was able to detect both monomeric RBD and the full trimeric SARS-CoV2 spike protein[37] with 15 pM LOD and >1700% dynamic range for the RBD detection (Fig. 4c, Extended Data Fig. 9). We further increased the dynamic range of lucCageRBD to 5300% by using a short version of lucKey to tune K_{CK} (Extended data Fig. 10a-c).
To evaluate the ability of our sensor platform to function in complex biological matrices, we compared RBD detection by lucCageRBD in buffer, simulated nasal matrix[38], and human serum, and observed only a minor reduction in the latter two conditions (Fig. 4c). Following a suggestion by Maarten Merkx[39], we controlled for variation in absolute luminescence signal in spiked serum samples from four different donors and spiked simulated nasal matrix using a BRET internal reference[40] for internal calibration, and found that with such calibration the RBD could be accurately quantified without compromising sensor dynamic range (Extended Data Fig. 11). These results suggest that it should be feasible to implement the lucCage system for future point-of-care applications.

To test the specificity of the biosensors developed in this work, we measured the activation kinetics of each lucCage in response to all target proteins used in this work. Each sensor responded rapidly and sensitively to its cognate target, but not to any of the others (Fig. 4d). For the most part, the actual sensors (see Table S2 and Table S3) perform as predicted by the simple thermodynamic model; for example, experiments at varying key and sensor concentrations suggest little coupling between parameters. However, there is considerable variation between different sensors in the level of activation at saturating target concentrations or high lucKey concentrations, which for most is lower than that expected for the complete luciferase reconstitution predicted by the model (Extended Data Fig. 10d-g and Table S4). This may be a consequence of steric interference between target binding to the latch and luciferase reconstitution as the target binding motif and the luciferase SmBiT are adjacent to each other in the latch; such interference could be resolved by increasing the separation between the two in the switch. The potential of the lucCage system is illustrated by the high dynamic range (5300%) and picomolar sensitivity of the lucCageRBD sensor: the near optimal K_{open} value results in a very low background in the absence of target without compromising the extent of activation at low target concentrations.

Discussion

It is instructive to put our sensors in the context of the multiple protein-based biosensor platforms that have been developed over the years with considerable success (see Supplementary discussion, and Table S5). Our sensor platform is based on the thermodynamic coupling between defined closed and open states of the system, thus, its sensitivity depends on the free energy change upon the sensing domain binding to the target but not the specific geometry of the binding interaction (the semi-synthetic small molecule sensors[10,11] also have this property). This enables the incorporation of various binding modalities, including small peptides, globular miniproteins, antibody epitopes and *de novo* designed binders, to generate sensitive sensors for a wide range of protein targets with little or no optimization. For point of care (POC) applications, our system, like other luminescence based protein biosensor platforms[8], has the advantages of being homogeneous, no-wash, and a nearly instantaneous readout; the quantification of luminescence can be carried out with inexpensive and accessible devices such as a cell phone camera[8]. In hospital settings, the ability to modularly design sensors with identical readouts for diverse targets could enable quick readout of large numbers of different compounds using an array of hundreds of different sensors.
Up until recently, the focus of *de novo* protein design was on the design of proteins with new structures corresponding to single deep free energy minima; our results highlight the progress in the field which now enables more complex multistate systems to be readily generated. Our sensors, like other *de novo* designed proteins, are expressed at high levels in cells and are very stable, which should considerably facilitate their manufacturing and distribution. As highlighted by the outstanding performance of the lucCageRBD sensor, there is a strong synergy between the general "molecular device" architecture of our platform and *de novo* designed high-affinity miniprotein binders[4,21] (these *de novo* miniproteins also synergize with other platforms[41]). As the power of computational design continues to increase, it should become possible to detect an ever wider range of targets with higher sensitivity using lucCage sensors. Beyond biosensors, our results highlight the potential of *de novo* protein design to create more general solutions for current day challenges than can be achieved by repurposing native proteins that have evolved to solve completely different challenges.

Figures and figure legends

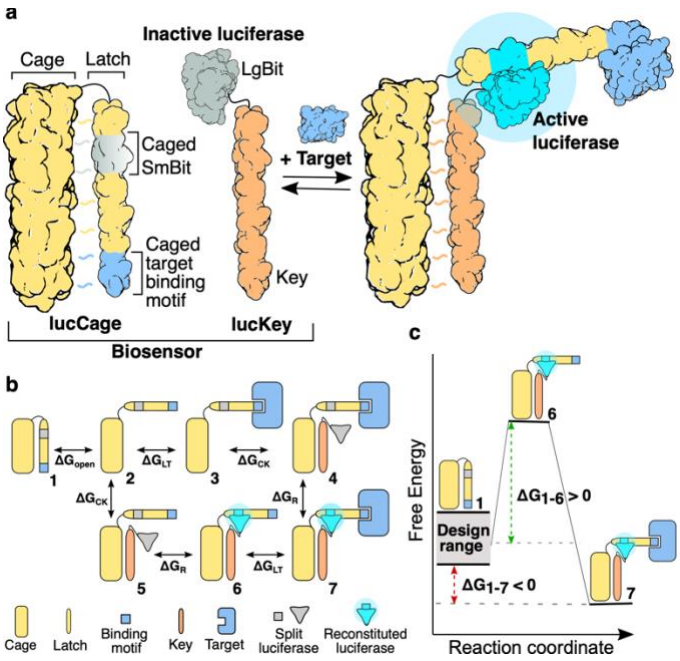


Fig. 1 ***De novo*** **design of multi-state biosensors. a**, Sensor schematic mechanism. The closed form of lucCage (left) can not bind to lucKey, thus preventing the split luciferase SmBiT fragment from interacting with LgBit. The open form (right) can bind both target and key, allowing the reconstitution of SmBiT and LgBiT for luciferase activity. **b,** Thermodynamics of biosensor activation. The free energy cost ΔG_{open} of the transition from closed cage (species 1) to open cage (species 2) disfavors association of key (species 5) and reconstitution of luciferase activity (species 6) in the absence of target. In the presence of the target, the combined free energies of target binding (2→3; ΔG_{LT}), key binding (3→4; ΔG_{CK}), and SmBiT-LgBiT association (4→7; ΔG_R) overcome the unfavorable ΔG_{open}, driving opening of the lucCage and reconstitution of luciferase activity. **c**, Thermodynamics of biosensor design. The designable parameters are ΔG_{open} and ΔG_{CK}; ΔG_R is the same for all targets, and ΔG_{LT} is pre-specified for each target. For sensitive but low background analyte detection, ΔG_{open} and ΔG_{CK} must be tuned such that the closed state (species 1) is substantially lower in free energy than the open state (species 6) in the absence of target, but higher in free energy than the open state in the presence of target (species 7).

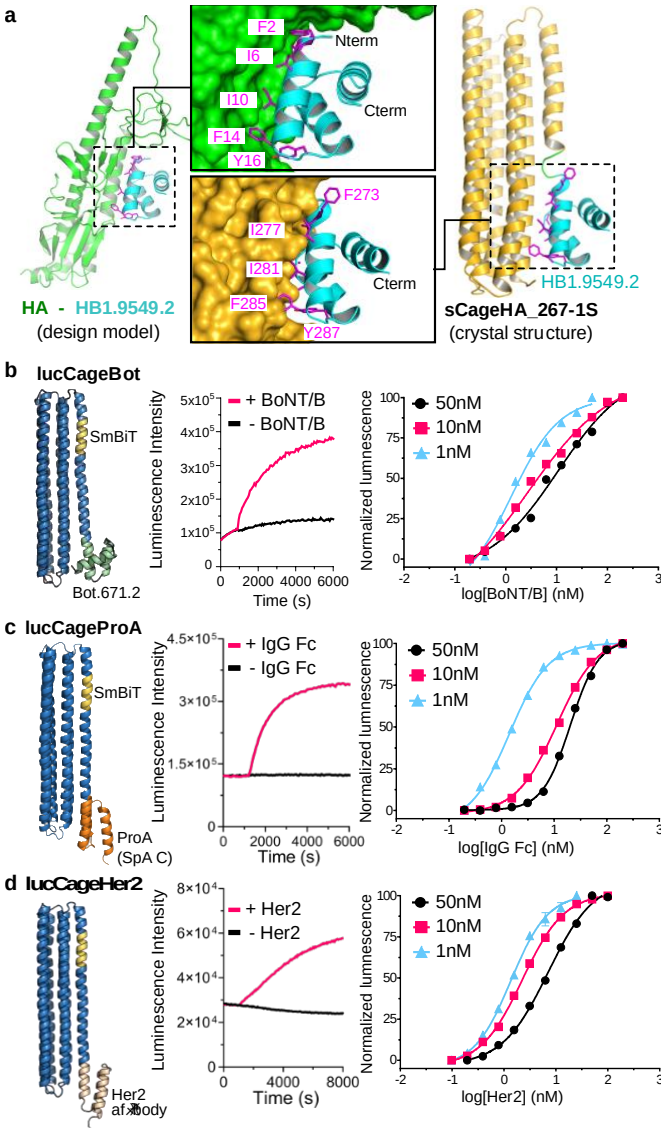


Fig. 2. Design and characterization of *de novo* biosensors incorporating small proteins as sensing domains. a, Structural validation of sCageHA_267-1S, caging small protein domains into a LOCKR switch. Left: design model of the *de novo* binder HB1.9549.2 (cyan ribbon) bound to the stem region of influenza hemagglutinin (HA, green ribbon)[21]. Right: crystal structure (PDB ID: 7CBC) of sCageHA_267_1S, comprising HB1.9549.2 (cyan) grafted into a shortened and stabilized version of the LOCKR switch[22] (sCage, yellow ribbon). Middle: All residues of HB1.9549.2 involved in binding to HA (magenta, top) except for F273 are buried in the closed state of the switch (bottom). The labels in magenta indicate the same set of amino acids in the two panels (e.g., F2 in the top panel corresponds to F273 in the lower panel). **b-d**, Functional characterization of lucCageBot, lucCageProA, and lucCageHer2. Left: structural models incorporate a *de novo* designed binder for BoNT/B (Bot.671.2) [21], the C domain of Protein A (SpaC)[23] or a Her2-binding affibody[24], respectively, into lucCage (blue ribbon) with caged SmBiT fragment (gold ribbon). Middle: Measurement of luminescence intensity upon addition of 50 nM of analyte (BoNT/B, IgG Fc, or Her2) to a mixture of 10 nM of each lucCage and 10 nM of lucKey. Right: detection over a wide range of analyte concentrations by changing the biosensor (lucCage + lucKey) concentration (colored lines). All experiments were performed in triplicate, representative data are shown, and data are presented as mean values +/- s.d.

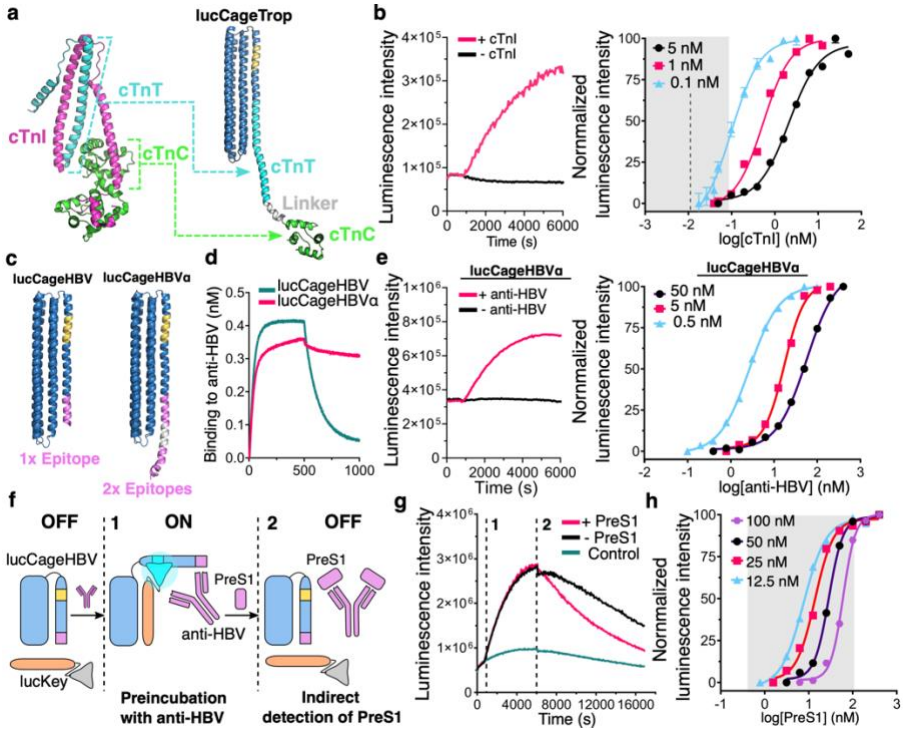


Fig. 3 Design and characterization of biosensors for cardiac troponin I and an anti-HBV antibody. a, Design of cardiac Troponin I sensor. Left: Structure of cardiac troponin (PDB ID: 4Y99); Troponin T, C and I (cTnT, cTnC and cTnI) are shown in cyan, green, and magenta, respectively. Right: Design model of lucCageTrop. **b**, Left: Kinetics of luminescence increase upon addition of 1 nM cTnI to 0.1 nM lucCageTrop+lucKey. Right: Wide detection range accessible by changing the concentration of the sensor components (colored lines). Grey area covers the cTnI concentration range relevant to the diagnosis of acute myocardial infarction (AMI) [27]; the dotted line indicates clinical AMI cut-off defined by W.H.O. (0.6 ng/mL, 25 pM). **c**, HBV sensor design models (gold, SmBiT; grey, linker; magenta, HBV PreS1 epitope). **d**, lucCageHBVα with two epitope copies shows higher affinity by biolayer interferometry for the anti-HBV antibody HzKR127-3.2 (*K*d=0.68 nM) than lucCageHBV (Kd=20 nM). **e**, Left: Kinetics of bioluminescence increase upon addition of 50 nM anti-HBV antibody to 1 nM lucCageHBVα+lucKey. Right: Sensitive anti-HBV antibody detection over a wide concentration range. **f**, Mechanism for PreS1 detection using lucCageHBV. **g**, Kinetics of bioluminescence following addition of the anti-HBV antibody ("1") and subsequently PreS1 ("2"), which decreases bioluminescence by competing with the sensor for the antibody. **h**, Detection of PreS1 can be achieved over the relevant post-HBV infection concentration levels (grey area) by varying the concentration of antibody (indicated by colored labels). All experiments were performed in triplicate, representative data are shown, and data are presented as mean values +/- s.d.

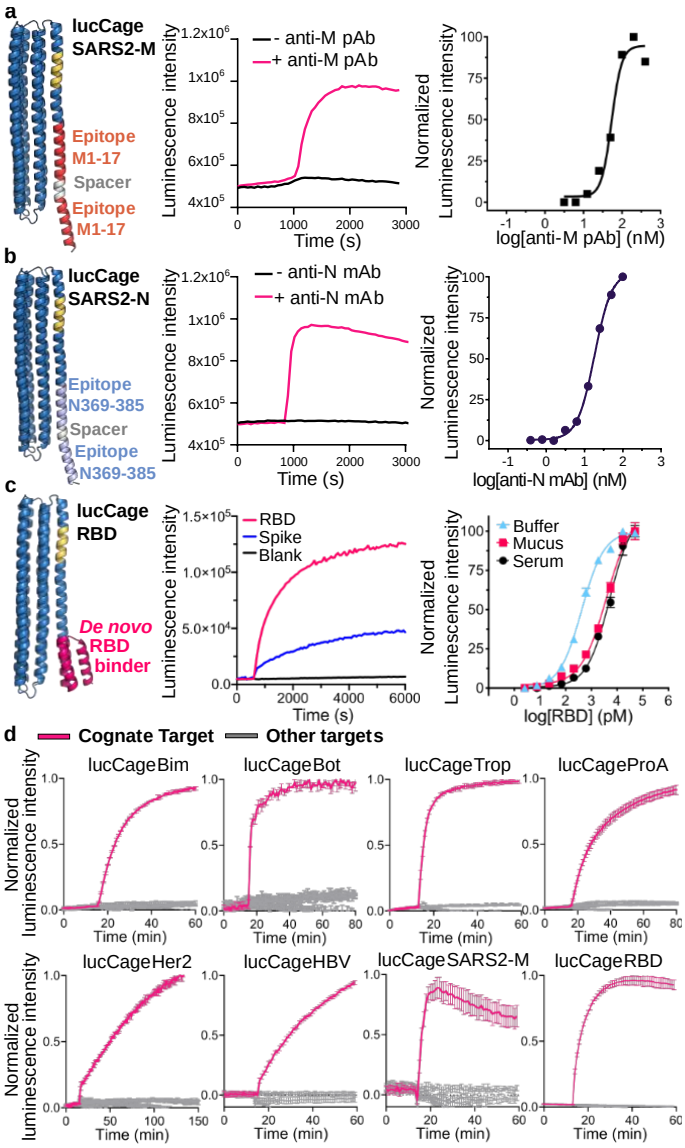


Fig. 4. Design of highly specific biosensors for the detection of anti-SARS-CoV-2 antibodies and SARS-CoV-2 viral proteins. a, Left panel: lucCageSARS2-M sensor incorporates two copies of the SARS-CoV-2 membrane protein 1-17 epitope (red) connected with a flexible spacer. Middle panel: kinetics of luminescence activation of 50 nM lucCageSARS2-M+lucKey upon addition of 100 nM anti-SARS-CoV-1-M rabbit polyclonal antibodies (ProSci, 3527) that cross-react with residues 1-17 of the SARS-CoV-2-M. Right panel: response of 5 nM lucCageSARS2-M+lucKey to varying concentrations of target anti-M pAb. **b,** Left panel: lucCageSARS2-N incorporates two copies of the SARS-CoV-2 nucleocapsid protein 369-382 epitope (light blue). Middle panel: kinetics of luminescence activation of 50 nM lucCageSARS2-N+lucKey upon addition of 100 nM anti-SARS-CoV-1-N mouse monoclonal antibody (clone 18F629.1), that recognizes the epitope. Right panel: response of 50 nM lucCageSARS2-N+lucKey to varying concentration of anti-N mAb. **c,** Left panel: lucCageRBD incorporates a *de novo* SARS-CoV-2 RBD binder[4] (LCB1, magenta). Middle panel: luminescence intensity upon addition of 16.7 nM SARS-Cov-2 RBD or trimeric spike protein to a mixture of 1 nM lucCageRBD+lucKey. Right panel: detection over a range of analyte concentrations in buffer, 10% synthetic nasal matrix[38] or

10% serum. **d,** Biosensor specificity. Each sensor at 1 nM was incubated with 50 nM of its cognate target (magenta lines) and the targets for the other biosensors (grey lines). Targets are Bcl-2, BoNT/B, human IgG Fc, Her2, cardiac Troponin I, anti-HBV antibody (HzKR127-3.2), anti-SARS-CoV-1-M polyclonal antibody and SARS-CoV-2 RBD. All experiments were performed in triplicate, representative data are shown, and data are presented as mean values +/- s.d.

References

1. Stein, V. & Alexandrov, K. Synthetic protein switches: design principles and applications. *Trends Biotechnol.* **33**, 101–110 (2015).
2. Langan, R. A. *et al.* De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
3. Udugama, B. *et al.* Diagnosing COVID-19: The Disease and Tools for Detection. *ACS Nano* vol. 14 3822–3835 (2020).
4. Cao, L. *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
5. Yeh, H.-W. & Ai, H.-W. Development and Applications of Bioluminescent and Chemiluminescent Reporters and Biosensors. *Annu. Rev. Anal. Chem.* **12**, 129–150 (2019).
6. Greenwald, E. C., Mehta, S. & Zhang, J. Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks. *Chem. Rev.* **118**, 11707–11794 (2018).
7. Banala, S., Aper, S. J. A., Schalk, W. & Merkx, M. Switchable reporter enzymes based on mutually exclusive domain interactions allow antibody detection directly in solution. *ACS Chem. Biol.* **8**, 2127–2132 (2013).
8. Arts, R. *et al.* Detection of Antibodies in Blood Plasma Using Bioluminescent Sensor Proteins and a Smartphone. *Anal. Chem.* **88**, 4525–4532 (2016).
9. van Rosmalen, M. *et al.* Dual-Color Bioluminescent Sensor Proteins for Therapeutic Drug Monitoring of Antitumor Antibodies. *Anal. Chem.* **90**, 3592–3599 (2018).
10. Yu, Q. *et al.* Semisynthetic sensor proteins enable metabolic assays at the point of care. *Science* **361**, 1122–1126 (2018).
11. Yu, Q. *et al.* A biosensor for measuring NAD levels at the point of care. *Nat Metab* **1**, 1219–1225 (2019).
12. Schena, A., Griss, R. & Johnsson, K. Modulating protein activity using tethered ligands with mutually exclusive binding sites. *Nat. Commun.* **6**, 7830 (2015).
13. Arts, R. *et al.* Semisynthetic Bioluminescent Sensor Proteins for Direct Detection of Antibodies and Small Molecules in Solution. *ACS Sens* **2**, 1730–1736 (2017).
14. Xue, L., Prifti, E. & Johnsson, K. A General Strategy for the Semisynthesis of Ratiometric Fluorescent Sensor Proteins with Increased Dynamic Range. *J. Am. Chem. Soc.* **138**, 5258–5261 (2016).
15. Guo, Z. *et al.* Generalizable Protein Biosensors Based on Synthetic Switch Modules. *J. Am. Chem. Soc.* **141**, 8128–8135 (2019).
16. Edwardraja, S. *et al.* Caged activators of artificial allosteric protein biosensors. *ACS Synth. Biol.* **9**, 1306–1314 (2020).
17. Ribeiro, L. F., Warren, T. D. & Ostermeier, M. Construction of Protein Switches by Domain Insertion and Directed Evolution. *Methods Mol. Biol.* **1596**, 43–55 (2017).
18. Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
19. Minor, D. L., Jr & Kim, P. S. Context-dependent secondary structure formation of a designed protein sequence. *Nature* **380**, 730–734 (1996).
20. Kale, J., Osterlund, E. J. & Andrews, D. W. BCL-2 family proteins: changing partners in the dance towards death. *Cell Death Differ.* **25**, 65–80 (2018).
21. Chevalier, A. *et al.* Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017).
22. Lajoie, M. J. *et al.* Designed protein logic to target cells with precise combinations of surface antigens. *Science* **369**, 1637–1643 (2020).
23. Deis, L. N. *et al.* Suppression of conformational heterogeneity at a protein-protein interface. *Proc. Natl.*

Acad. Sci. U. S. A. **112**, 9028–9033 (2015).
24. Eigenbrot, C., Ultsch, M., Dubnovitsky, A., Abrahmsén, L. & Härd, T. Structural basis for high-affinity HER2 receptor binding by an engineered protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15039–15044 (2010).
25. Hobbs, R. J., Thomas, C. A., Halliwell, J. & Gwenin, C. D. Rapid Detection of Botulinum Neurotoxins-A Review. *Toxins* **11**, (2019).
26. Perrier, A., Gligorov, J., Lefèvre, G. & Boissan, M. The extracellular domain of Her2 in serum as a biomarker of breast cancer. *Lab. Invest.* **98**, 696–707 (2018).
27. Rubini Gimenez, M. *et al.* One-hour rule-in and rule-out of acute myocardial infarction using high-sensitivity cardiac troponin I. *Am. J. Med.* **128**, 861–870.e4 (2015).
28. Collins, M. H. Serologic Tools and Strategies to Support Intervention Trials to Combat Zika Virus Infection and Disease. *Trop Med Infect Dis* **4**, (2019).
29. Pondé, R. A. de A. Expression and detection of anti-HBs antibodies after hepatitis B virus infection or vaccination in the context of protective immunity. *Arch. Virol.* **164**, 2645–2658 (2019).
30. Chi, S.-W. *et al.* Broadly neutralizing anti-hepatitis B virus antibody reveals a complementarity determining region H3 lid-opening mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9230–9235 (2007).
31. Kim, J. H. *et al.* Enhanced humanization and affinity maturation of neutralizing anti-hepatitis B virus preS1 antibody based on antigen-antibody complex structure. *FEBS Lett.* **589**, 193–200 (2015).
32. Ovacik, M. & Lin, K. Tutorial on Monoclonal Antibody Pharmacokinetics and Its Considerations in Early Development. *Clin. Transl. Sci.* **11**, 540–552 (2018).
33. Locarnini, S. & Bowden, S. Hepatitis B surface antigen quantification: Not what it seems on the surface. *Hepatology* **56**, 411–414 (2012).
34. Chow, S. C. S. *et al.* Specific epitopes of the structural and hypothetical proteins elicit variable humoral responses in SARS patients. *J. Clin. Pathol.* **59**, 468–476 (2006).
35. He, Y., Zhou, Y., Siddiqui, P., Niu, J. & Jiang, S. Identification of immunodominant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus. *J. Clin. Microbiol.* **43**, 3718–3726 (2005).
36. Wang, H. *et al.* SARS-CoV-2 proteome microarray for mapping COVID-19 antibody interactions at amino acid resolution. *ACS Central Science* **6**, 2238–2249 (2020).
37. Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
38. Panpradist, N. *et al.* Swab sample transfer for point-of-care diagnostics: characterization of swab types and manual agitation methods. *PLoS One* **9**, e105786 (2014).
39. Ni, Y. *et al.* RAPPID: a platform of ratiometric bioluminescent sensors for homogeneous immunoassays. doi:10.1101/2020.10.31.363044.
40. Yeh, H.-W. *et al.* Red-shifted luciferase-luciferin pairs for enhanced bioluminescence imaging. *Nat. Methods* **14**, 971–974 (2017).
41. Yang, C. *et al.* Bottom-up de novo design of functional proteins with complex structural features. *Nat. Chem. Biol.* (2021) doi:10.1038/s41589-020-00699-x.

Methods

Design of the sensor system: lucCage and lucKey

The low affinity SmBiT 114 (VTGYRLFEEIL)[18] was grafted into the latch of the asymmetric LOCKR switch described in Langan et al, 2019[2] using GraftSwitchMover, a RosettaScripts-based protein design algorithm (See Supplementary methods for details). The grafting sampling range was assigned between residues 300-330. The resulting designs were energy-minimized, visually inspected and selected for subsequent gene synthesis, protein production and biochemical analyses. The best SmBit position on the latch was experimentally determined to be an insertion at residue 312, as described in Extended Data Fig. 2. This design was named lucCage. lucKey was assembled by genetically fusing the LgBit of NanoLuc[18] to the key peptide described in Langan et al, 2019. All protein sequences are listed in Table S6.

Computational grafting of sensing domains into lucCage

Peptides and epitopes: The amino acid sequence for each sensing domain was grafted using Rosettascripts[42] GraftSwitchMover into all α-helical registers between residues 325-359 of lucCage. In the cases where the desired sequence to be inserted exceeded the length of the lucCage latch, we made use of Rosetta Remodel [43] to model the C-terminus extension of lucCage (See Supplementary methods for details). The resulting lucCages were energy-minimized using Rosetta fast relax[44], visually inspected and typically less than ten designs were selected for subsequent protein production and biochemical characterization.

Protein domains: the main secondary structure element segment forming the interface of the binding protein domain with the target was identified. The amino acid sequence was extracted and grafted into lucCage using the GraftSwitchMover or Rosetta Remodel as described above. Then, we used MergePDBMover and Pymol 2.0 to align, model and visualize the full-length binding domain in the context of the switch (See Supplementary methods for details). The designs were energy-minimized using Rosetta fast relax and visually inspected for selection.

Synthetic gene construction

The designed protein sequences were codon optimized for *E. coli* expression and ordered as synthetic genes in pET21b+ or pET29b+ *E. coli* expression vectors. The synthetic gene was inserted at the NdeI and XhoI sites of each vector, including an N-terminal hexahistidine tag followed by a TEV protease cleavage site and a stop codon was added at the C terminus.

General procedures for bacterial protein production and purification

The *E. coli* Lemo21(DE3) strain (NEB) was transformed with a pET21b+ or pET29b+ plasmid encoding the synthesized gene of interest. Cells were grown for 24 hours in LB media supplemented with carbenicillin or kanamycin. Cells were inoculated at a 1:50 mL ratio in the Studier TBM-5052 autoinduction media supplemented with carbenicillin or kanamycin, grown at 37 °C for 2-4 hours, and then grown at 18 °C for an additional 18 h. Cells were harvested by centrifugation at 4000*g* at 4 °C for 15 min and resuspended in 30 ml lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 30 mM imidazole, 1 mM PMSF, 0.02 mg/mL DNAse). Cell resuspensions were lysed by sonication for 2.5 minutes (5 second cycles). Lysates were clarified by centrifugation at 24,000*g* at 4 °C for 20 min and passed through 2 ml of Ni-NTA nickel resin (Qiagen, 30250) pre-equilibrated with wash buffer, (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 30 mM imidazole). The resin was washed twice with 10 column volumes (CV) of wash buffer, and then eluted with 3 CV of elution buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM imidazole). The eluted proteins were concentrated using Ultra-15 Centrifugal Filter Units (Amicon) and further purified by using a Superdex^{TM} 75 Increase 10/300 GL (GE Healthcare) size exclusion column in Tris Buffered Saline (TBS; 25 mM Tris-HCl pH 8.0, 150 mM NaCl). Fractions containing monomeric protein were pooled, concentrated, and snap-frozen in liquid nitrogen and stored at -80 °C.

In vitro bioluminescence characterization

A Synergy Neo2 Microplate Reader (BioTek) was used for all in vitro bioluminescence measurements. Assays were performed in 1:1=DPBS (with calcium, Gibco):Nano-Glo (Promega) assay buffer for cTnI sensors while 1:1=HBS-EP (GE Healthcare Life Sciences) :Nano-Glo assay buffer was used for other sensors. 10X lucCage, 10X lucKey, and 10X target proteins of desired concentrations were first prepared from stock solutions. For each well of a white opaque 96-well plate, 10 µL of 10X lucCage, 10 µL of 10X lucKey, and 20 µL of buffer were mixed to reach the indicated concentration and ratio. The lucCage and lucKey components were incubated for 60 minutes at RT to enable pre-equilibration. The plate was centrifuged at 1000 × g for 1 min and incubated at RT for additional 10 min. Then, 50 µL of 50X diluted furimazine (Nano-Glo luciferase assay reagent, Promega) was added to each well. For assays containing serum or simulated nasal matrix (110mM NaCl, 1% w/v mucin, 10µg/mL human genomic DNA[38]), buffer composition was replaced by the biological matrix. Bioluminescence measurements in the absence of target were taken every 1 min post-injection (0.1 s integration and 10 s shaking during intervals). After ~15 min, 10 µL of serially diluted 10X target protein plus a blank was injected and

bioluminescence kinetic acquisition continued for a total of 2 h. To derive EC_{50} values from the bioluminescence-to-analyte plot, the top three peak bioluminescence intensities at individual analyte concentrations were averaged, subtracted from blank, and used to fit the sigmoidal 4PL curve. To calculate the limit of detection (LOD), the linear region of bioluminescence responses of sensors to its analyte was extracted and a linear regression curve was obtained. It was used to derive the standard deviation (s.d.) of the response and the slope of the calibration curve (S). The LOD was determined as 3×(s.d./S).

Detection of spiked RBD in human serum specimens

Serum specimens were derived from excess plasma or sera from adults (>18 yo) of both genders kindly provided by the Director of the Clinical Chemistry Division, the hospital of University Washington. All anonymized donor specimens were provided de-identified. Since the donors consented to have their excess specimens be used for other experimental studies, they could be transferred to our study without additional consent. All samples were passed through 0.22 μm filters before use. 10 μL of 10X serial diluted monomeric RBD (167-0.69nM), 5 μL of 20X lucCage (20nM), 5 μL of 20X lucKey (20nM), 5 μL of 20X Antares2 (2nM), and 10, 20, 25, or 50 μL of human donor serum or simulated nasal matrix were mixed with 1:1=HBS:Nano-Glo assay buffer to reach a total volume of 75 μL. The plate was centrifuged at 1000 × g for 1 min. Then, 25 μL of 25X diluted furimazine in buffer was added to each well. Bioluminescence signals were recorded from both 470/40 nm and 590/35 nm channels every 1 min for a total of 1h. Ratio at each time point was calculated by the equation described in Extended Data Figure 11b. Monomeric SARS-CoV-2 RBD was expressed and purified as described elsewhere[45].

Biolayer interferometry (BLI)

Protein-protein interactions were measured by using an Octet® RED96 System (ForteBio) using streptavidin-coated biosensors (ForteBio). Each well contained 200 μL of solution, and the assay buffer was HBS-EP+ Buffer (GE Healthcare Life Sciences, 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% v/v Surfactant P20) + 0.5% Non-fat dry milk blotting grade blocker (BioRad). The biosensor tips were loaded with analyte peptide/protein at 20 μg/mL for 300 s (threshold of 0.5 nm response), incubated in HBS-EP+ Buffer for 60 s to acquire the baseline measurement, dipped into the solution containing Cage and/or Key for 600 s (association step) and dipped into the HBS-EP+ Buffer for 600 s (dissociation steps). The binding data were analyzed with the ForteBio Data Analysis Software version 9.0.0.10.

Design and characterization of lucCageBim

The Bim peptide sequence (EIWIAQELRRIGDEFNAYYA) was threaded into the lucCage scaffold as described in the "Design of sensing domains into lucCage" section. The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. The bioluminescence detection signal was measured for each design lucCage at 20 nM mixed with lucKey at 20 nM, in the presence or absence of target Bcl-2 protein at 200nM. Recombinant Bcl-2 was produced as described somewhere else[46].

Design and characterization of lucCageHer2, lucCageProA, lucCageBot and lucCageRBD

The main binding motifs of the Bot.0671.2 de novo binder, *S. aureus* Protein A domain C (SpaC), the Her2 affibody and the *de novo* RBD binder LCB1 were threaded into lucCage as described in the "Design of sensing domains into lucCage" section (See Table S3 and Table S6 for sequences). The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. The designs were screened by measuring bioluminescence signal for each design lucCage at 20 nM mixed with lucKey at 20 nM, in the presence or absence of 200 nM target protein. The target proteins used were: Botulinum Neurotoxin B HcB expressed as previously described[47], human IgG1 Fc-HisTag (AcroBiosystems, Cat. No. IG1-H5225) and human Her2-HisTag (AcroBiosystems, Cat. No. HE2-H5225). Monomeric SARS-CoV-2 RBD and the trimeric SARS-CoV-2 spike protein (Hexapro pre-stabilized version[37]) were expressed and purified as described previously[45].

Design and characterization of lucCageTrop

The cardiac Troponin T (cTnT) binding motif sequence was truncated into fragments of different length (see Extended Data Fig. 6) and threaded into the lucCage scaffold as described in the "Design of sensing domains into lucCage" section. The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. The designs were screened by measuring bioluminescence signal for each design lucCage at 20 nM mixed with lucKey at 20 nM in the presence or absence of 100 nM cardiac Troponin I (Genscript, Cat. No. Z03320-50). Subsequently, lucCageTrop, an improved version by fusion to cardiac Troponin C (cTnC), was created by genetically fusing the following sequence to the C terminus of lucCageTrop627.

Design and characterization of lucCageHBV and lucCageHBVα

The binding motif (GANSNNPDWDFN) of the PreS1 domain was threaded into the lucCage scaffold at every position after residues 336 using the Rosetta GraftSwitchMover. Following the

Rosetta FastRelax protocol, eight designs were selected for protein production. The designs were screened by measuring bioluminescence signal for each design lucCage (20 nM) and lucKey (20 nM) in the presence or absence of the anti-HVB antibody HzKR127-3.2 (100 nM) to select lucCageHBV. Subsequently, lucCageHBVα was constructed by genetically fusing a sequence containing a second antigenic motif (GGSGGGSSGFGANSNNPDWDFNPN) to lucCageHBV.

Design and characterization of lucCageSARS2-M and lucCageSARS2-N

Antigenic epitopes of the SARS-CoV-2 membrane protein (a.a. 1-31, 1-17 and 8-24) and the nucleocapsid protein (a.a. 368-388 and 369-382) were computationally grafted into lucCage as described in the "Design of sensing domains into lucCage" section. The selected designs were expressed in *E. coli*, purified and characterized for luminescence activation. All designs at 50nM were mixed with 50nM lucKey and experimentally screened for an increase in luminescence in the presence of rabbit anti-SARS-CoV Membrane polyclonal antibodies (ProSci, Cat. No.: 3527) at 100nM or mouse anti-SARS-CoV Nucleocapsid monoclonal antibody (clone 18F629.1, NovusBio Cat. No. NBP2-24745) at 100 nM.

Design and characterization of sCageHA variants

HB1.9549.2 was embedded into the parental six-helix bundle for sCage design at different positions along the latch helix of the scaffold. To promote more favorable intramolecular interactions, three consecutive residues on the latch were intentionally substituted with glycine to allow for conformational freedom. The five designs were produced in *E. coli*. Biolayer interferometry analysis was performed with purified Cages (1 μM) and biotinylated Influenza A H1 hemagglutinin (HA)[21] loaded onto streptavidin-coated biosensor tips (ForteBio) in the presence or absence of the key (2 μM) using an Octet instrument (ForteBio).

Production and purification of HzKR127-3.2

The synthetic V_H and V_L DNA fragments were subcloned into the pdCMV-dhfrC-cA10A3 plasmid containing the human Cγ1 and C*κ* DNA sequences. The vector was introduced into HEK 293F cells using Lipofectamine (Invitrogen), and the cells were grown in FreeStyle 293 (GIBCO) in 5% CO_2 in a 37 °C humidified incubator. The culture supernatant was loaded onto a protein A-sepharose column (Millipore), and the bound antibody was eluted by the addition of 0.2 M glycine–HCl (pH 2.7), followed by immediate neutralization with 1 M Tris–HCl (pH 8.0). The solution was dialyzed against 10 mM HEPES-NaOH (pH 7.4), and the purity of the protein was analyzed by SDS-PAGE.

Production and purification of the PreS1 domain

The DNA fragment encoding the PreS1 domain (residues 1-56) was cloned into the pGEX-2T (GE Healthcare) plasmid, and the protein was produced in the *E. coli* BL21(DE3) strain (NEB) at 18 ^{o}C as a fusion protein with glutathion-S-transferase (GST) at the N-terminus. The cell lysates were prepared in a buffer solution (25 mM Tris-HCl pH 8.0, 300 mM NaCl), and clarified supernatant was loaded onto GSTBind™ Resin (Novagen). The GST-PreS1 domain was eluted with the same buffer containing additional 10 mM reduced glutathione, further purified using a Superdex^{TM} 75 Increase 10/300 GL (GE Healthcare) size exclusion column, and concentrated to 34 μM.

Production of SCageHA_267-1S and its variants

sCageHA_267-1S and sCageHA_267-1S(E99Y/T144Y) were expressed at 18 °C in the *E. coli* LEMO21(DE3) strain (NEB) as a fusion protein containing a $(His)_{10}$-tagged cysteine protease domain (CPD) derived from *Vibrio cholerae* [48] at the C-terminus. The protein was purified using HisPur^{TM} nickel resin (Thermo), a HiTrap Q anion exchange column (GE Healthcare) and a HiLoad 26/60 Superdex 75 gel filtration column (GE Healthcare). For Selenomethionine (SelMet)-labeling, an I30M mutation was introduced additionally to generate a sCageHA_267-1S(E99Y/T144Y/I30M) variant. This protein was expressed in the *E. coli* B834 (DE3) RIL strain (Novagen) in the minimal media containing SeMet, and purified according to the same procedure for purifying the other variants.

Crystallization and structure determination of sCageHA_267-1S

Two point mutations (Glu99Tyr and Thr144Tyr) were introduced in an attempt to induce favorable crystal packing interactions. Good-quality single crystals of sCageHA_267-1S(E99Y/T144Y/I30M) were obtained in a hanging-drop vapor-diffusion setting by micro-seeding in a solution containing 11% (v/v) ethanol, 0.25 M NaCl, 0.1 M TrisHCl (pH 8.5). The crystals required strict maintenance of the temperature at 25 °C. For cryoprotection, the crystals were soaked briefly in the crystallization solution supplemented with 15% *2,3*-butanediol and flash-cooled in the liquid nitrogen. A single-wavelength anomalous dispersion (SAD) data set was collected at the Se absorption peak and processed with *HKL2000*[49]. Se positions and initial electron density map were calculated using the AutoSol module in *PHENIX* [50]. The model building and structure refinement were performed by using *COOT* [51] and *PHENIX*.

Methods References

42. Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6**, e20161 (2011).

43. Huang, P.-S. *et al.* RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS One* **6**, e24109 (2011).

44. Khatib, F. *et al.* Algorithm discovery by protein folding game players. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18949–18953 (2011).

45. Walls, A. C. *et al.* Elicitation of potent neutralizing antibody responses by designed protein nanoparticle vaccines for SARS-CoV-2. *Cell* **183**, P1367–1382.E17 (2020).

46. Berger, S. *et al.* Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins in cancer. *Elife* **5**, (2016).

47. Jin, R., Rummel, A., Binz, T. & Brunger, A. T. Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity. *Nature* **444**, 1092–1095 (2006).

48. Shen, A. *et al.* Mechanistic and structural insights into the proteolytic activation of Vibrio cholerae MARTX toxin. *Nat. Chem. Biol.* **5**, 469–478 (2009).

49. Otwinowski, Z. & Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

50. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861–877 (2019).

51. Potterton, L. *et al.* Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2288–2294 (2004).

Authors and Acknowledgments

Authors: Alfredo Quijano-Rubio[†,1,2], Hsien-Wei Yeh[†,1], Jooyoung Park[1], Hansol Lee[3], Robert A. Langan[1], Scott E. Boyken[1], Marc J. Lajoie[1], Longxing Cao[1], Cameron M. Chow[1], Marcos C. Miranda[1], Jimin Wi[4], Hyo Jeong Hong[4], Lance Stewart[1], Byung-Ha Oh*[1,3], David Baker*[1,5]

Affiliations:
[1]Department of Biochemistry and Institute for Protein Design, University of Washington, Seattle, Washington 98195, USA
[2]Department of Bioengineering, University of Washington, Seattle, Washington 98195, USA
[3]Department of Biological Sciences, KAIST Institute for the Biocentury, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea
[4]Department of Systems Immunology, College of Biomedical Science, Kangwon National University, Chuncheon 200-701, Republic of Korea
[5]Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA

*Co-corresponding authors
†These authors contributed equally to this work

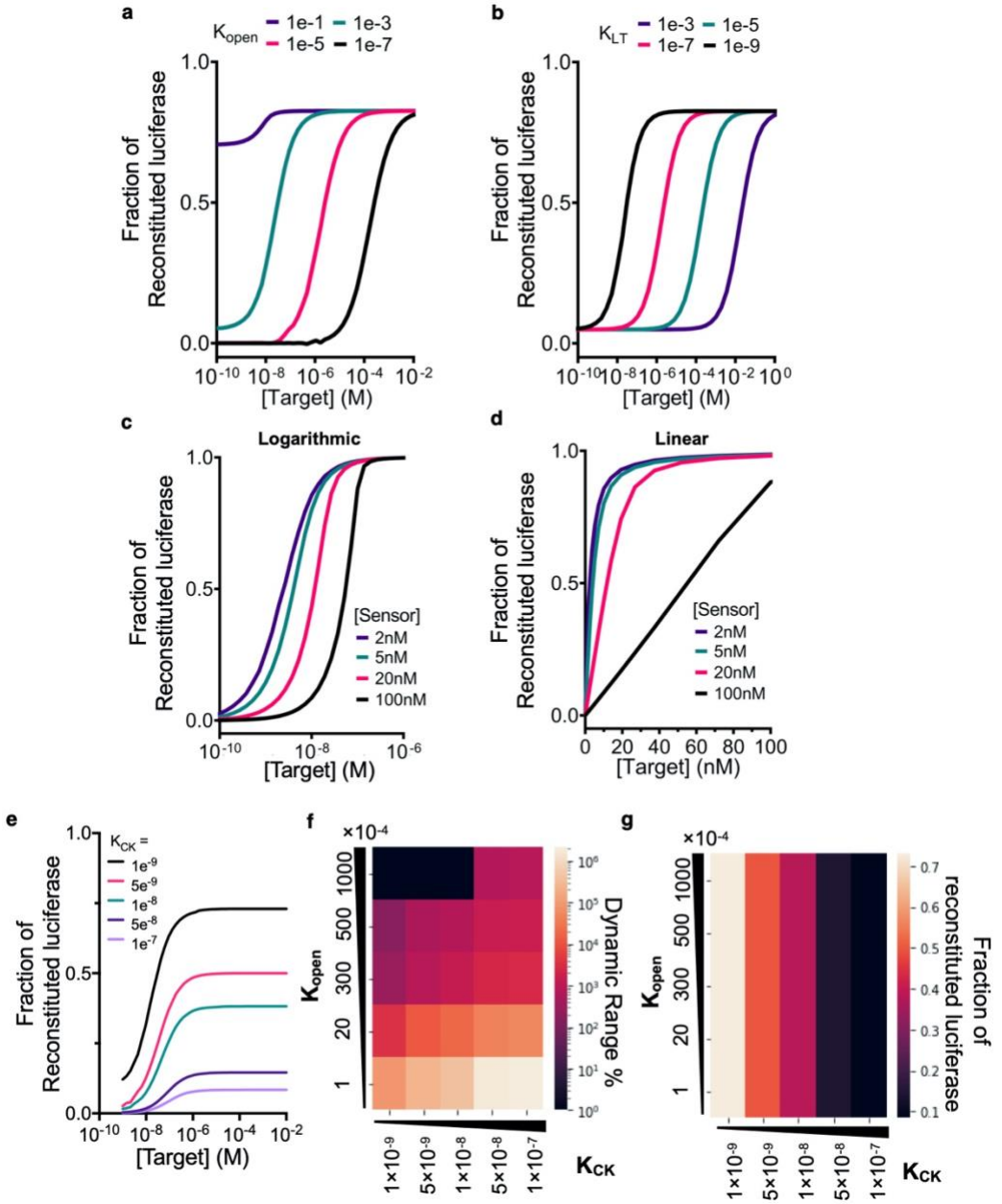
Author contributions:
D.B. directed the work. D.B., A.Q.-R., H.-W.Y., B.-H.O., J.P. and L.S. designed and further conceptualized the research. A.Q.-R., J.P., B.-H.O., and H.-W.Y. performed the computational design of the sensors. A.Q.-R., J.P. and B.-H.O. explored the early sensor designs. H.-W.Y., and A.Q.-R. optimized the design of sensors to improve performance and conducted the experimental characterizations. B.H.O. directed, and H.L. performed the crystallographic work. H.-W.Y. and R.A.L wrote the thermodynamic model and performed the simulations. R.A.L wrote GraftSwitchMover. S.E.B. and M.J.L. designed the parental cage and key protein scaffolds. L.C. designed the RBD binder LCB1. C.M.C., M.C.M., J.W. and H.J.H. performed protein purification. D.B, B.-H.O., A.Q.-R. and H.-W.Y. wrote the original draft. All authors reviewed and accepted the manuscript.

Funding: We acknowledge funding from HHMI (D.B.), the LG Yonam Foundation (B.-H.O.), the BK21 PLUS project of Korea (H.L.), the United World Antiviral Research Network (UWARN) one of the Centers Researching Emerging Infectious Diseases "CREIDs", NIAID 1 U01 AI151698-01 (D.B., L.S., and H-.W.Y.), The Audacious Project at the Institute for Protein Design (D.B., H-.W.Y., C.M.C., and M.C.M.), Eric and Wendy Schmidt by recommendation of the Schmidt Futures (A.Q-.R. and H-.W.Y.), the Washington Research Foundation (J.P. and M.J.L.), the Nordstrom Barrier Institute for Protein Design Directors Fund (R.A.L.), The Open Philanthropy Project Improving Protein Design Fund (D.B. and S.E.B.), the gift support from Gree Real Estate (A.Q.-R.), "la Caixa" Foundation (A.Q.-R., ID 100010434 under grant LCF/BQ/AN15/10380003), Support 1U19AG065156-01 (D.B.), and Air Force Office of Scientific Research FA9550-18-1-0297 (D.B.).

We thank Dr. Mark Wener for collecting de-identified human sera specimens, Dr. Wesley C. Van Voorhis for advice and support with the anti-SARS-CoV-2 antibody sensors, Nuttada Panpradist and Dr. Barry Lutz for providing simulated nasal matrix, Stephanie Berger for sharing the Bcl-2 protein target, Daniel Adriano Silva Manzano for providing Botulinum Neurotoxin B, Alex Kang

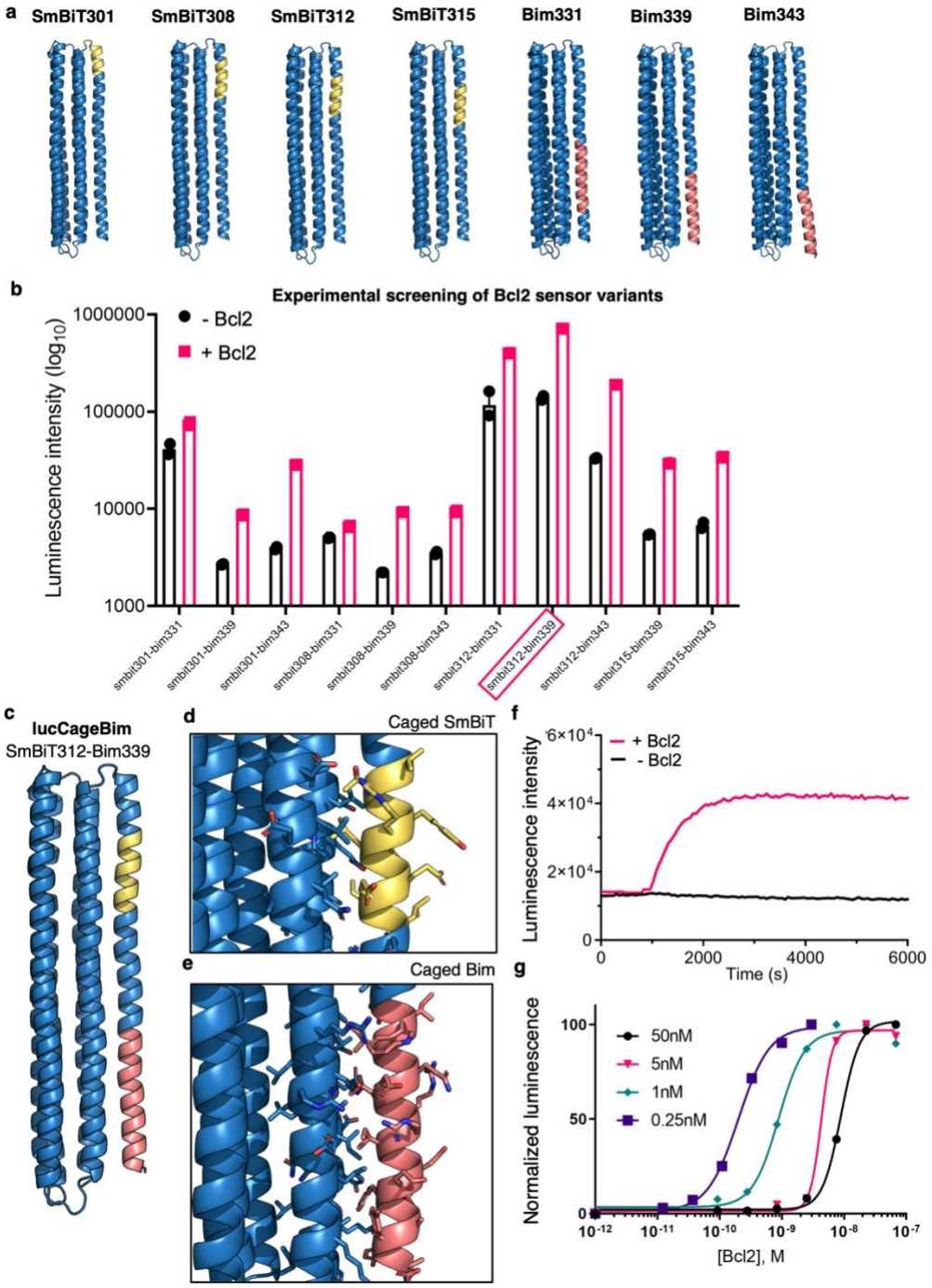
for setting up screening crystal trays, and Lauren Carter, Brooke Fiala and the Institute for Protein Design for providing SARS-CoV-2 RBD and Spike protein. We thank Dr. Maarten Merkx for suggesting the internal BRET referencing to control for sample to sample fluctuations in luciferase activity. The X-ray data were collected on the Beamline 5C at the Pohang Accelerator Laboratory, Korea. All protein structure and model images were generated using PyMOL 2.0.

Extended Data figures legends



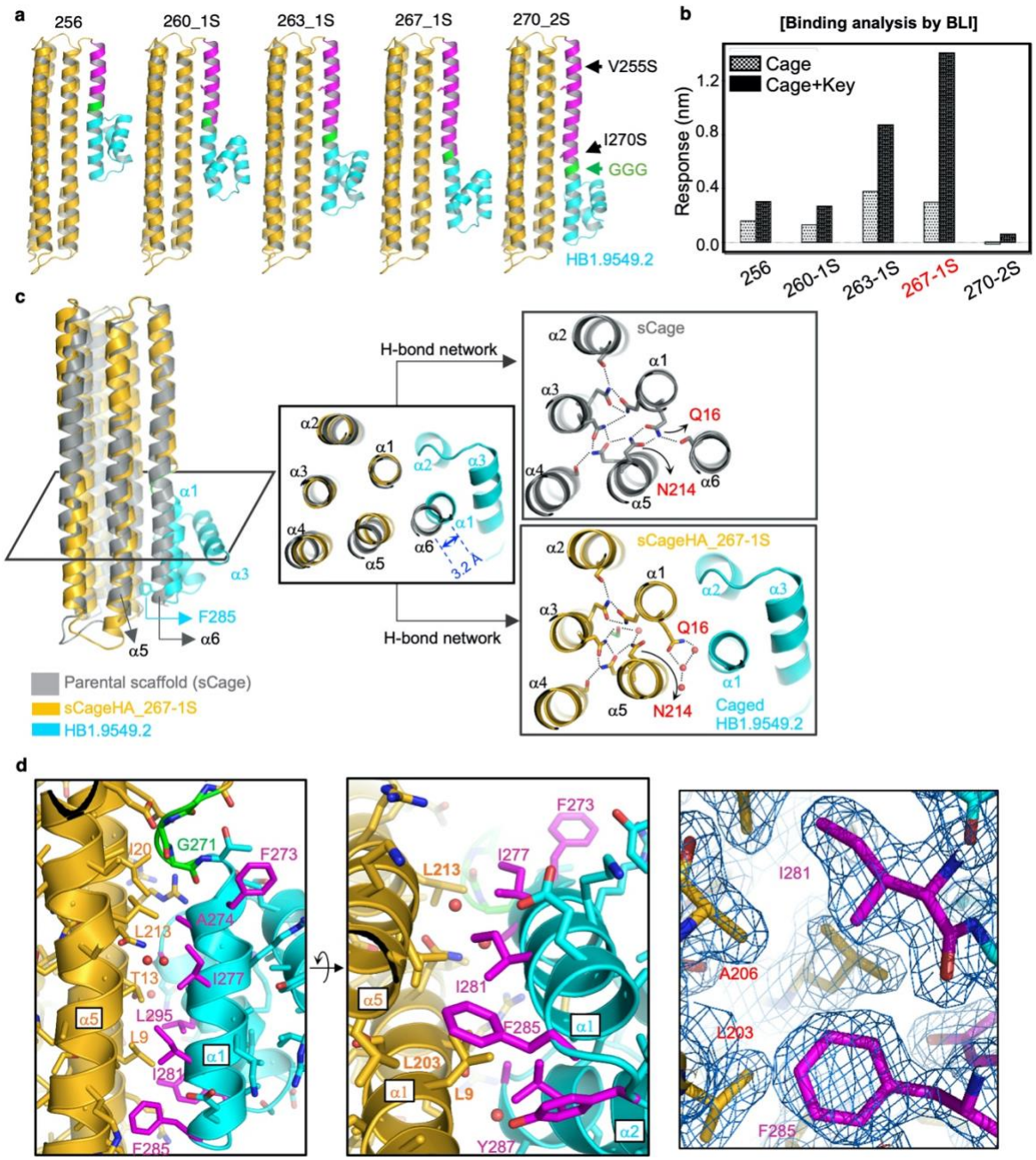
Extended Data Figure 1. Numerical simulations of the sensor thermodynamic equilibria showing the tunability of the lucCage platform to optimize sensitivity and dynamic range. Numerical simulations of the coupled equilibria shown in **Fig. 1b** for different values of (**a**) K_{open}, (**b**) K_{LT}, (**c**)(**d**) $[lucKey]_{tot}$ and $[lucCage]_{tot}$ and (**e**) K_{open}. K_{LT} and K_{CK} were set to 1×10^{-3}, 10^{-9} M, and 10^{-8} M respectively, and the concentration of the sensor components to 10:100 nM ($[lucCage]_{tot}$:$[lucKey]_{tot}$) except where explicitly indicated. **a**, Increasing ΔG_{open} (smaller K_{open}) shifts the sensor response to higher analyte concentrations. **b**, The sensor LOD is approximately $0.1 \times K_{LT;}$ the driving force for opening the switch becomes too weak below this concentration. **c-d**, The effective target detection range can be tuned by changing the concentrations of the two sensor components. Simulation results shown in a logarithmic scale (**c**) or linear scale (**d**) for target concentration illustrate that the steepness of the response depends on the ratio of the sensor

concentration to the K_D of the binding interaction (K_{LT}). **e**, K_{CK} values affect both species responsible for background and signal (species 6 and 7 in Fig. 1b, respectively), leading to different sensor dynamic ranges. **f-g**, Simulations with various K_{open} and K_{CK} values. Too large K_{open} value and strong lucCage-lucKey interaction (K_{CK}) increase the formation of the species 6 (in Fig. 1b). **f,** A heatmap representing the calculated sensor dynamic range according to the K_{open} and K_{CK} values. K_{open} exerts a predominant effect on the dynamic range, while K_{CK} provides an additional one-order of tunability. **g**, A heatmap showing the fraction of reconstituted luciferase (sensitivity) at saturating target concentration, indicating a trade-off of K_{CK} tuning.



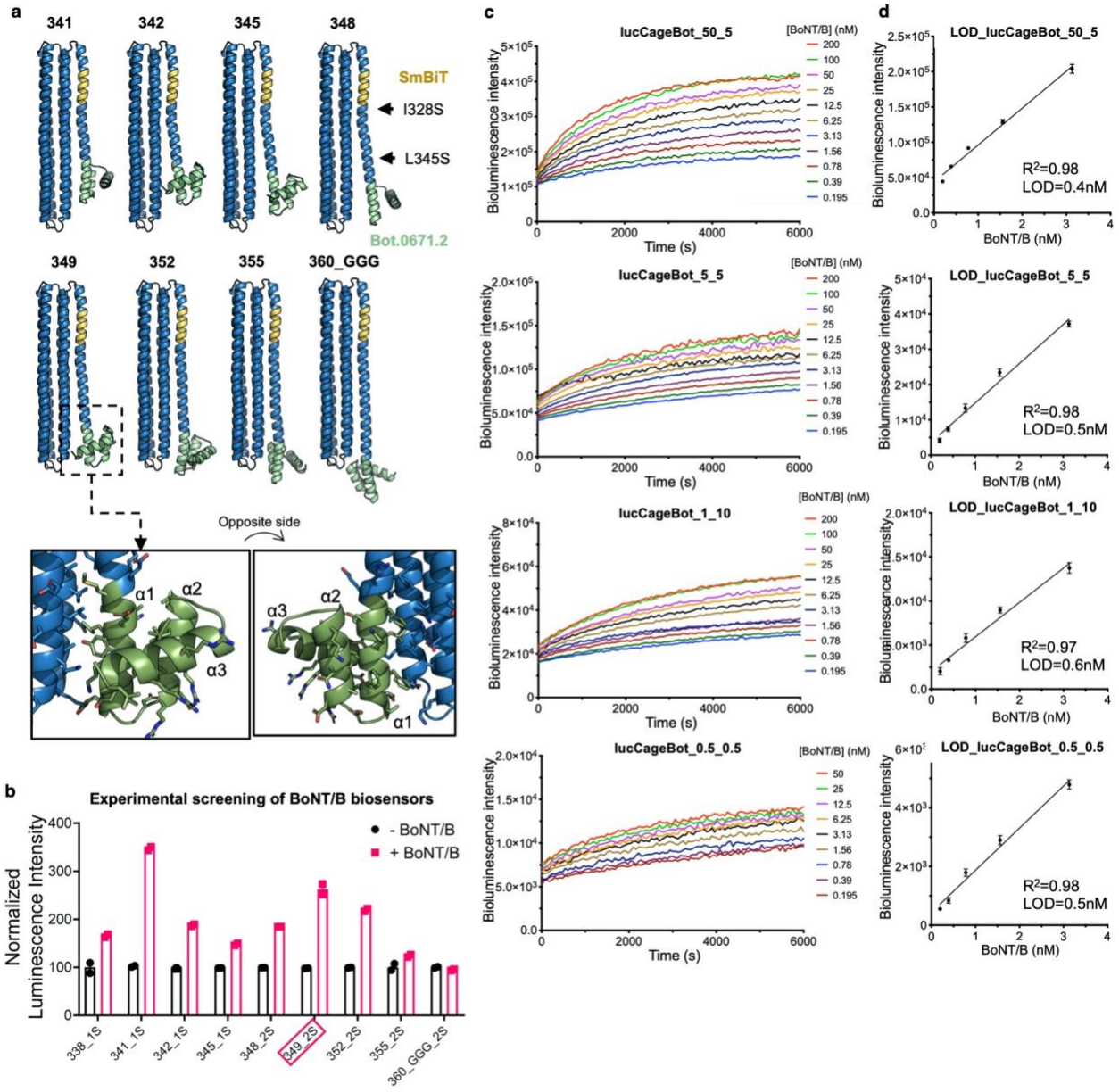
Extended Data Figure 2. Determination of the optimal SmBiT position in lucCage and characterization of lucCageBim, a Bcl-2 biosensor. a, Protein models showing the different threading positions of SmBiT (gold) and the Bim peptide (salmon) on the latch helix of the *de novo* LOCKR switch (blue). **b**, Experimental screening of 11 *de novo* Bcl-2 sensors. Eleven variants were generated by combining the SmBiT and Bim positions in (a) and characterized by activation of their luminescence upon addition of Bcl-2. Luminescence measurements were performed with each design (20 nM) and lucKey (20 nM) in the presence or absence of Bcl-2 (200 nM). SmBiT312-Bim339 (hence referred to as lucCageBim) was selected for posterior

characterization due to its higher brightness, dynamic range and stability. **c-g**, Characterization of lucCageBim. **c**, Structural design model in ribbon representation. **d**, close up view showing the predicted interface of SmBiT (gold) and Cage (blue). **e**, close up view showing the predicted interface of Bim (salmon) and Cage (blue). **f**, Kinetic luminescence measurements upon addition of Bcl-2 (200 nM) to a mix of lucCageBim (20 nM) and lucKey (20 nM). **g**, Tunable sensitivity of lucCageBim to Bcl-2 by changing the concentrations of sensor (lucCageBim and lucKey) components (colored curves).



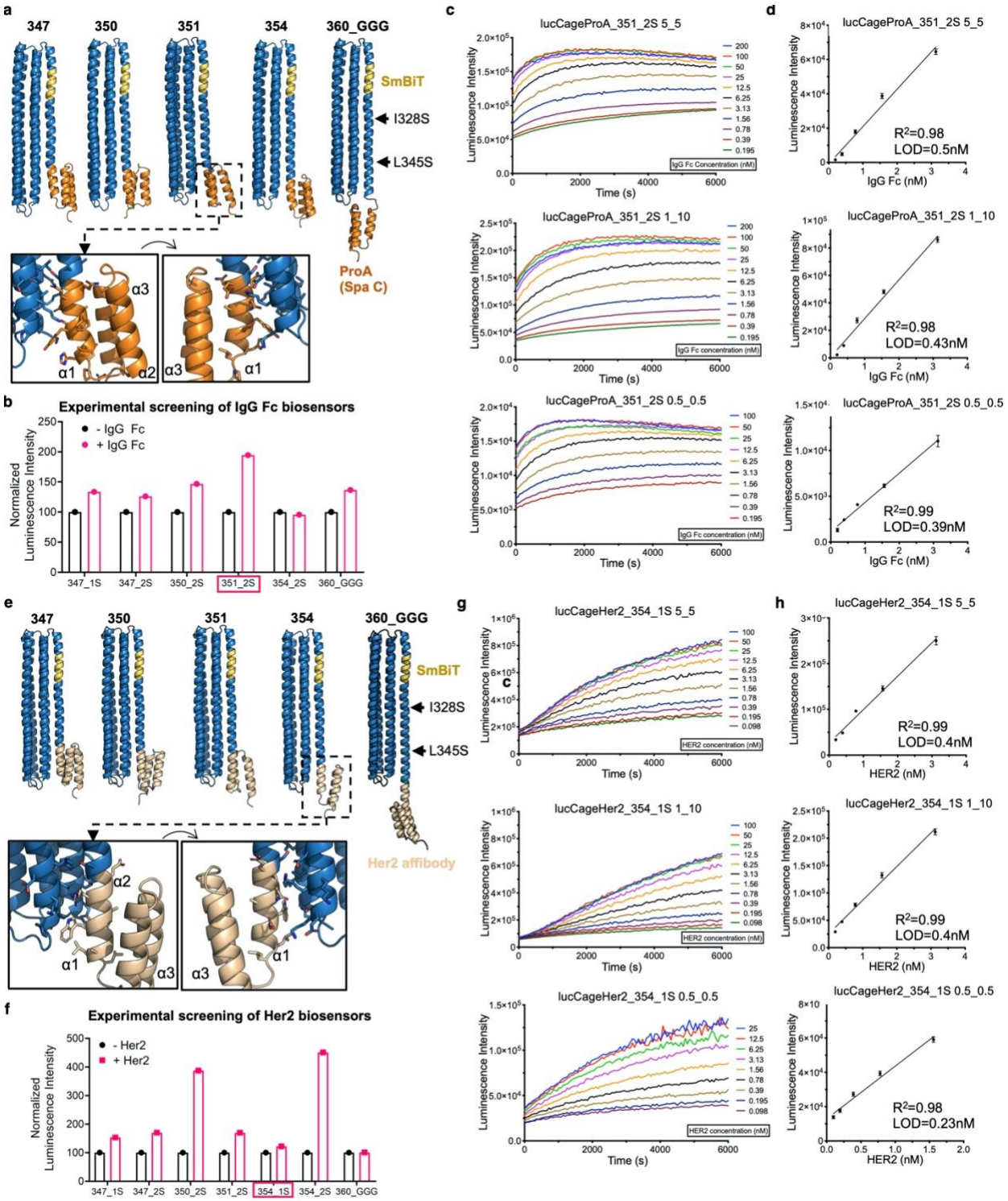
Extended Data Figure 3. Functional screening of sCageHA designs and crystal structure of sCageHA_267-1S. a, Structural models of sCageHA designs with the embedded *de novo* binder HB1.9549.2. The HB1.9549.2 protein (cyan) was grafted into a parental six-helix bundle (sCage, yellow) at different positions along the latch helix (magenta) including three consecutive glycine residues (green). The black arrows indicate the additionally introduced single V255S (1S) or double V255S/I270S (2S) mutation(s) on the latch. **b**, Experimental validation of five sCageHA designs binding to HA in the presence or absence of the key by biolayer interferometry. The concentration of the sCages and the key were 1 μM and 2 μM, respectively. Each experiment was performed once. sCageHA_267-1S exhibited the highest fold of activation. **c**, Structural comparison showing the flexible nature of sCage to enable caging of HB1.9549.2. The structural model of sCage (grey) and the crystal structure of sCageHA_267-1S (gold) are superposed, and a

narrow section (black box) is shown in an orthogonal view for detail. The N-terminal helix of HB1.9549.2 is displaced from the latch helix (α6) by 3.2 Å (middle panel) with a concomitant displacement of α5 and partial disruption of a hydrogen-bond network involving Q16 and N214 of sCage (right panels). **d**, A close up view of the intramolecular interactions of sCageHA_267-1S. The HA-binding residues are highlighted in magenta. Both the N-terminal helix (cyan α1) and the following helix (cyan α2) of HB1.9549.2 interact with the cage. The intramolecular interactions are all hydrophobic. The bulky hydrophobic side chain of F285 tightly abuts against the backbone atoms of α5 of sCage, which is unlikely to happen without a bending of α5. Unfavorable interactions are also found: F273 is solvent-exposed, and the Y287 hydroxyl group is buried in the apolar environment. The rightmost panel shows the quality of the electron density map.



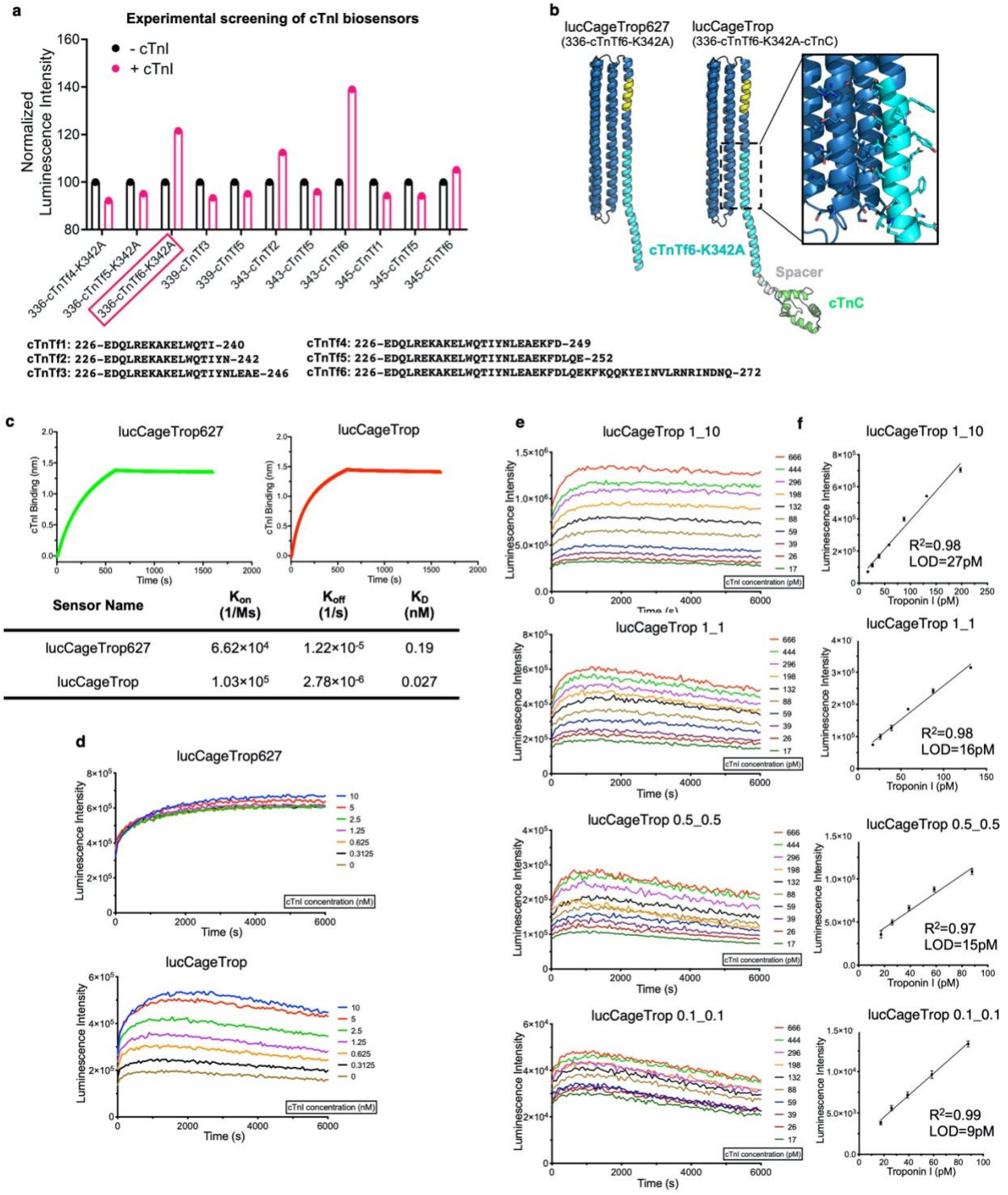
Extended Data Figure 4. Design and characterization of a Botulinum neurotoxin B sensor. **a**, Structural models of the botulinum neurotoxin B (BoNT/B) sensor designs showing the different threading positions of Bot.0671.2 (green, PDB ID: 5VID) on the latch of lucCage (blue). The SmBiT peptide is shown in gold ribbon representation. I328S and L345S indicate mutations introduced to tune the latch-cage interface (1S=I328S, 2S=I328S/L345S) [2], and "GGG" indicates the presence of three consecutive glycine residues between the latch and the grafted protein. The black box shows a close-up view of the interface of Cage (blue) and Bot.0671.2 (green) in the 349_2S design. **b**, Experimental screening of 9 *de novo* BoNT/B sensors. Luminescence measurements were performed for each design (20 nM) and lucKey (20 nM) in the presence or absence of the BoNT/B protein (200 nM). The luminescence values for each design were normalized to 100 in the absence of BoNT/B. Design 349_2S was selected as the best candidate due to high sensitivity and stability, and was named lucCageBot. **c**, Determination of lucCagerBot sensitivity. Bioluminescence was measured over 6000 s in the presence of serially diluted BoNT/B

protein. From top to bottom - lucCageBot:lucKey concentration (nM) = 50:5, 5:5, 1:10, 0.5:0.5. **d**, Limit of detection (LOD) calculations for the sensor at different concentrations. From top to bottom - lucCageBot:lucKey concentration (nM) = 50:5, 5:5, 1:10, 0.5:0.5. All experiments were performed in triplicate, representative data are shown, and data are presented as mean values +/- s.d.



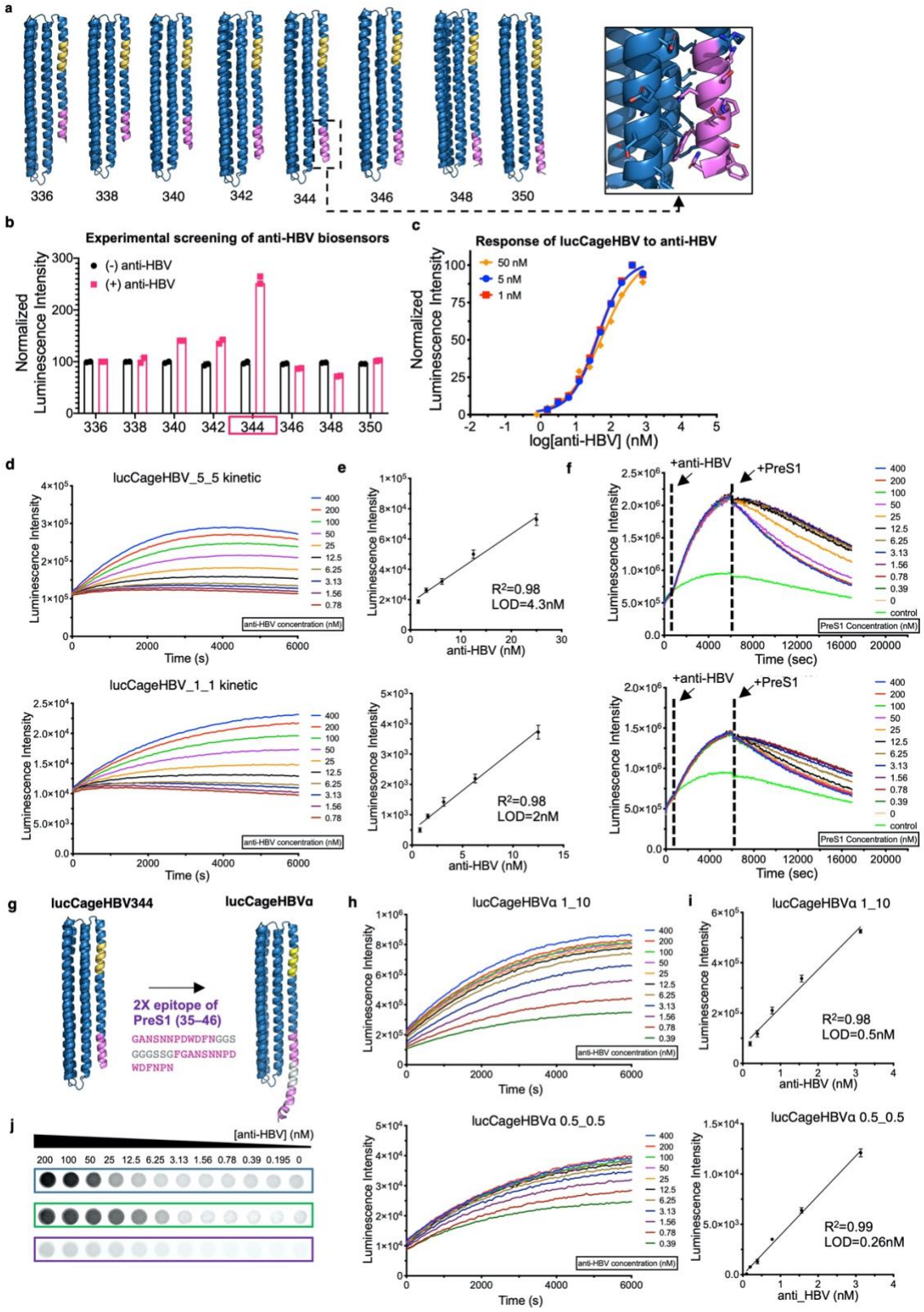
Extended Data Figure 5. Design and characterization of an Fc domain sensor (lucCageProA) and a Her2 sensor (lucCageHer2). a, Structural models of the Fc sensor designs showing the different threading positions of the *S. aureus* Protein A domain C (orange, PDB ID: 4WWI) on the latch of lucCage (blue). The SmBit peptide is shown in gold ribbon representation. I328S and L345S indicate mutations introduced to tune the latch-cage interface, (1S=I328S, 2S=I328S/L345S) [2], and "GGG" indicates the presence of three consecutive glycine residues

between the latch and the grafted protein. **b**, Experimental screening of 6 *de novo* Fc domain sensors. Luminescence measurements were performed for each design (20 nM) and lucKey (20 nM) in the presence or absence of recombinant human IgG1 Fc (200 nM). The luminescence values were normalized to 100 in the absence of Fc. Design 351_2S was selected as the best candidate due to high sensitivity and stability, and was named lucCageProA. This experiment was performed using single replicates in two independent instances, representative data are shown. **c**, Determination of lucCageProA's sensitivity. Bioluminescence was measured over 6000 s in the presence of serially diluted Fc protein. From top to bottom - lucCageBot:lucKey concentration (nM) = 5:5, 1:10, 0.5:0.5. **d**, LOD calculations for the sensor at different concentrations. From top to bottom - lucCageBot:lucKey concentration (nM) = 5:5, 1:10, 0.5:0.5. **e**, Structural models of the Her2 sensor designs showing the different threading positions of the Her2 affibody protein (PDB ID: 3MZW, beige) on the latch of lucCage (blue). The SmBiT peptide is shown in gold ribbon representation. I328S and L345S indicate mutations introduced to tune the latch-cage interface, (1S=I328S, 2S=I328S/L345S) [2], and "GGG" indicates the presence of three consecutive glycine residues between the latch and the grafted protein. The black boxes show a close-up view of the interface of Cage (blue) and the Her2 affibody (beige) in the 354_2S design. **f**, Experimental screening of 7 *de novo* Her2 sensors. Luminescence measurements were taken for each design (20 nM) and lucKey (20 nM) in the presence or absence of the ectodomain of Her2 (200 nM). The luminescence values were normalized to 100 in the absence of Her2 ectodomain. This experiment was performed using single replicates in two independent instances, representative data are shown. Design 354_2S was selected as the best candidate due to high sensitivity and stability, and was named lucCageHer2. **g**, Determination of lucCagerHer2's sensitivity. Bioluminescence was measured over 6000 s in the presence of serially diluted Her2 ectodomain protein. From top to bottom - lucCageBot:lucKey concentration (nM) = 5:5, 1:10, 0.5:0.5. **h**, Limit of detection (LOD) calculations for the sensor at different concentrations. From top to bottom - lucCageBot:lucKey concentration (nM) = 5:5, 1:10, 0.5:0.5. All experiments were performed in triplicate unless specifically indicated, representative data are shown, and data are presented as mean values +/- s.d.



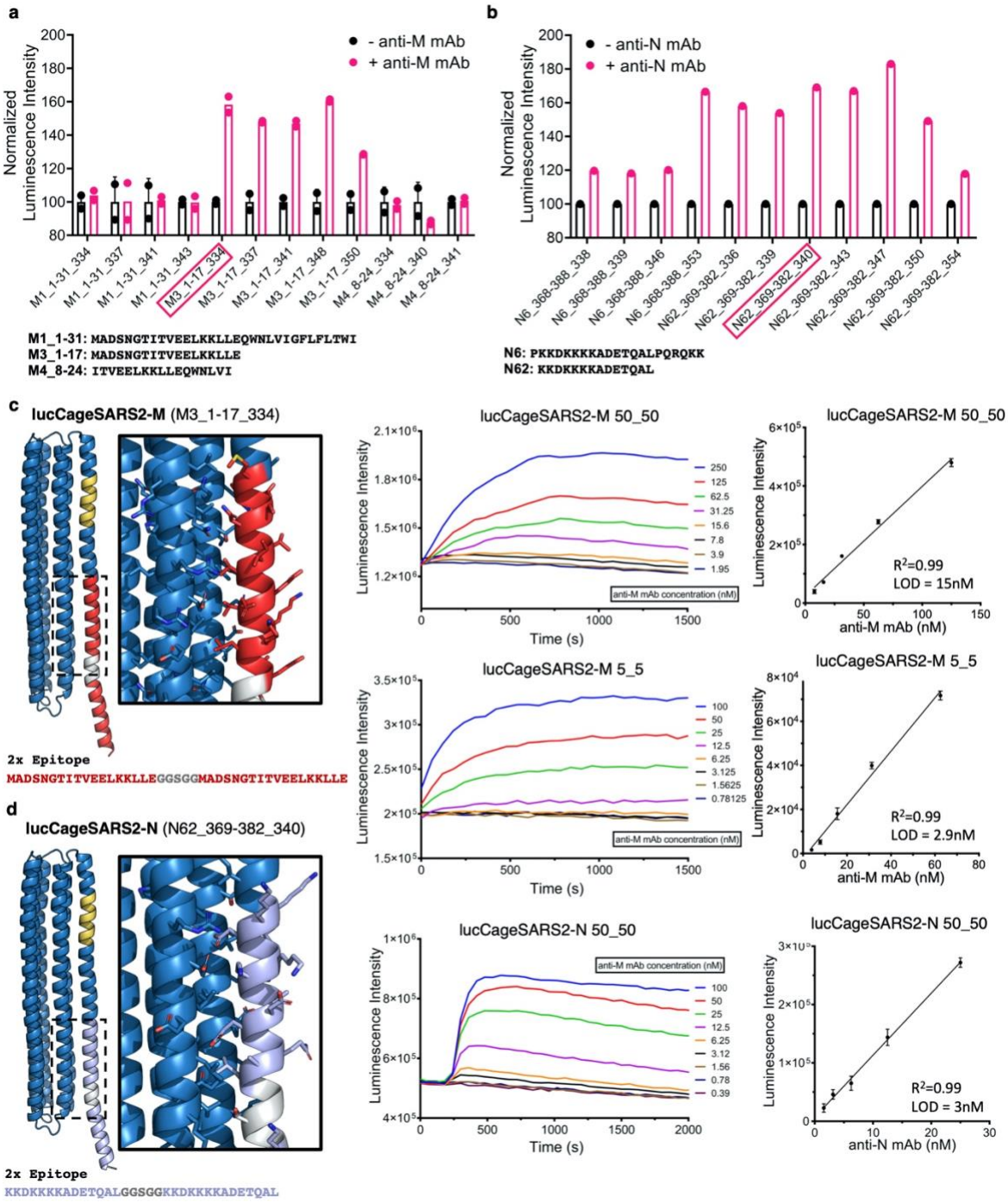
Extended Data Figure 6. Design, selection, and engineering of lucCageTrop for cardiac Troponin I detection. a, Experimental screening of designed sensors for cardiac Troponin I (cTnI). Fragments of cardiac Troponin T, namely cTnTf1-f6, were computationally grafted into lucCage at different positions of the latch. All designs were produced in *E. coli* and experimentally screened at 20 nM and 20 nM lucKey for an increase in luminescence in the presence of cTnI (100 nM). The luminescence values were normalized to 100 in the absence of cTnI. This experiment

was performed using single replicates in two independent instances, representative data are shown. Design 336-cTnTf6-K342A was selected as the best candidate (named lucCageTrop627) based on its sensitivity, activation fold-change, and stability. **b**, Models of lucCageTrop627 and lucCageTrop, an improved version by fusion of cardiac Troponin C (cTnC) at the C-terminus of lucCageTrop627. The models are shown in ribbon representation comprising SmBit (gold) a fragment of cTnT (cyan, PDB ID: 4Y99), and cTnC (green, PDB ID: 4Y99). The black box shows a close-up view of the interface of Cage (blue) and cTnT (cyan) in the lucCageTrop design. **c**, The binding affinity of lucCageTrop627 and lucCageTrop to cTnI was measured by biolayer interferometry. lucCageTrop showed 7-fold higher affinity to cTnI than lucCageTrop627. **d**, Comparison of bioluminescence kinetics between lucCageTrop627 (top) and lucCageTrop (bottom) in the presence of serially diluted cTnI. Higher binding affinity leads to improved dynamic range and sensitivity of the sensor. **e**, Determination of lucCageTrop's sensitivity. Bioluminescence was measured over 6000 s in the presence of serially diluted cTnI. From top to bottom - lucCageTrop:lucKey concentration (nM) = 1:10, 1:1, 0.5:0.5, 0.1:0.1. **f**, LOD calculations for the sensor at different concentrations. From top to bottom - lucCageTrop:lucKey concentration (nM) = 1:10, 1:1, 0.5:0.5, 0.1:0.1. All experiments were performed in triplicate unless otherwise indicated, representative data are shown, and data are presented as mean values +/- s.d.



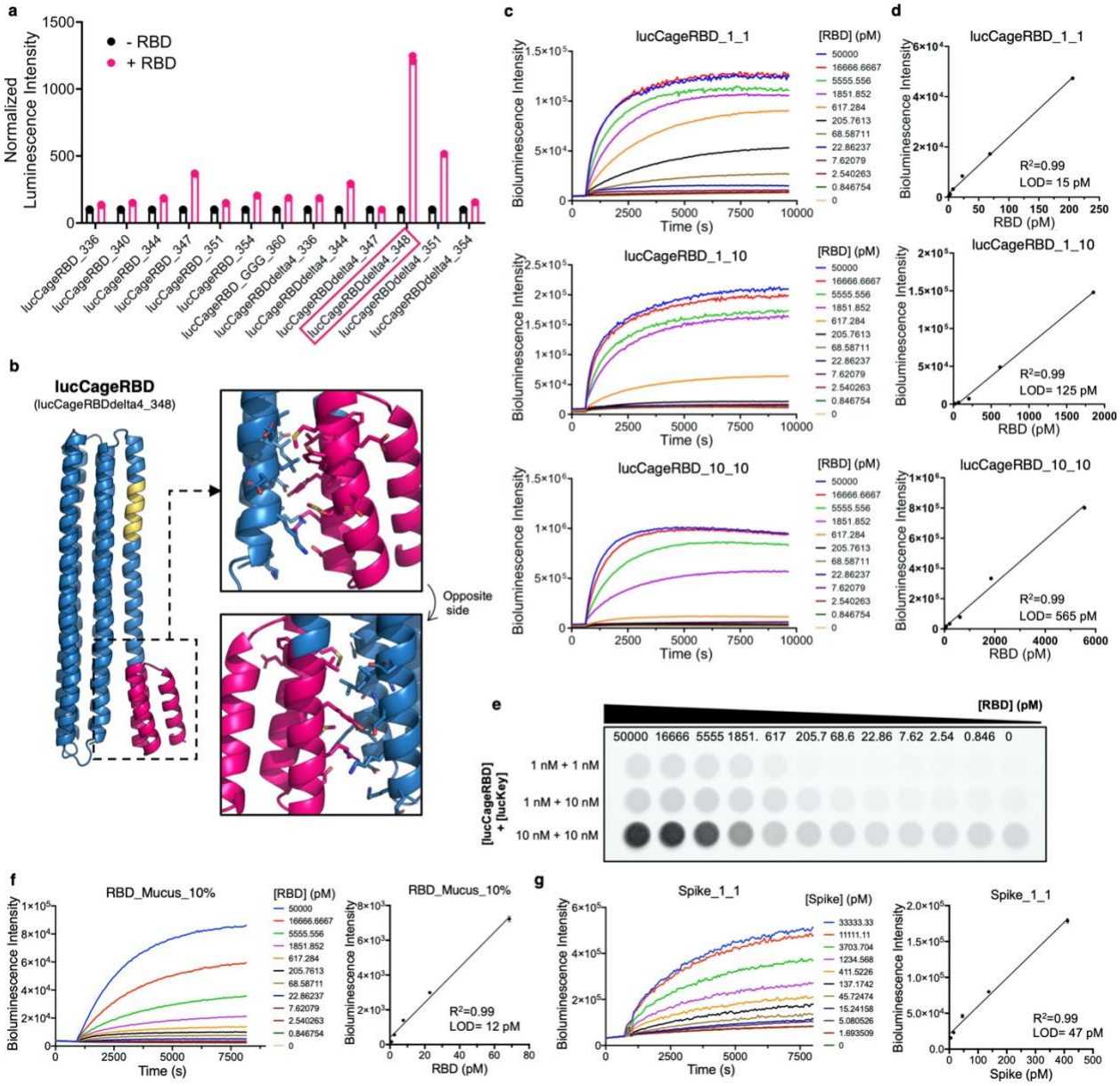
Extended Data Figure 7. Design and characterization of an anti-HBV antibody sensor.
a, The energy-minimized models of lucCage designs are shown with the threaded segments of SmBiT (gold) and the antigenic motif of PreS1 (magenta). The right box shows a close-up view

of the cage-motif interface of the HBV344 design. **b**, Experimental screening of all designs performed by monitoring the luminescence of each lucCage (20 nM) and lucKey (20 nM) in the presence or absence of the anti-HBV antibody HzKR127-3.2 (100 nM). The luminescence values were normalized to 100 in the absence of anti-HBV. This experiment was performed in duplicate in two independent instances, representative data are shown. The design HBV344 was selected due to its better performance and was named lucCageHBV. **c,d,** Determination of lucCageHBV sensitivity. Bioluminescence was measured over 6000 s in the presence of serially diluted HzKR127-3.2. From top to bottom - lucCageHBV:lucKey concentration (nM) = 5:5, 1:1. The maximum values of the curves in **d**, are used to obtain the curves in **c**. **e**, LOD calculations for the sensor at different concentrations. From top to bottom - lucCageHBV:lucKey concentration (nM) = 5:5, 1:1. **f**, Detection of PreS1 by competition of lucCageHBV344 and HzKR127-3.2 shown in Fig. 3f. Luminescence kinetics after the addition of the antibody (anti-HBV, first arrow). From top to bottom - anti-HBV antibody concentrations = 50, 12.5 nM. At 6000 s, different concentrations of the PreS1 domain were injected into the wells, and the decreased luminescence signals were used to detect PreS1. **g,** Design of lucCageHBVα for improved detection of an anti-HBV antibody. The structural model of lucCageHBVα is shown with a close-up detail of the predicted interface between the PreS1 epitope (magenta) and lucCage (blue). The design comprises two copies of the epitope PreS1 (a.a. 35-46), spaced by a flexible linker (grey) to enable bivalent interaction with the antibody. The SmBit peptide is shown in gold. **h**, Determination of lucCageHBVα detection sensitivity to the presence of the antibody HzKR127-3.2 (anti-HBV). Bioluminescence was measured over 6000 s in the presence of serially diluted HzKR127-3.2. From top to bottom - lucCageHBVα:lucKey concentration (nM) = 1:10, 0.5:0.5. **i**, The linear region of a calibration curve was used to determine the LOD and the dynamic range of antibody detection. **j**, Bioluminescence images acquired with a BioRad ChemiDoc imaging system. From top to bottom, lucCageHBVα:lucKey concentration (nM) = 50:5, 5:5, 1:10. Changes in bioluminescence intensity levels were detected as a function of the concentration of HzKR127-3.2. All experiments were performed in triplicate unless specifically indicated, and representative data are presented as mean values +/- s.d.



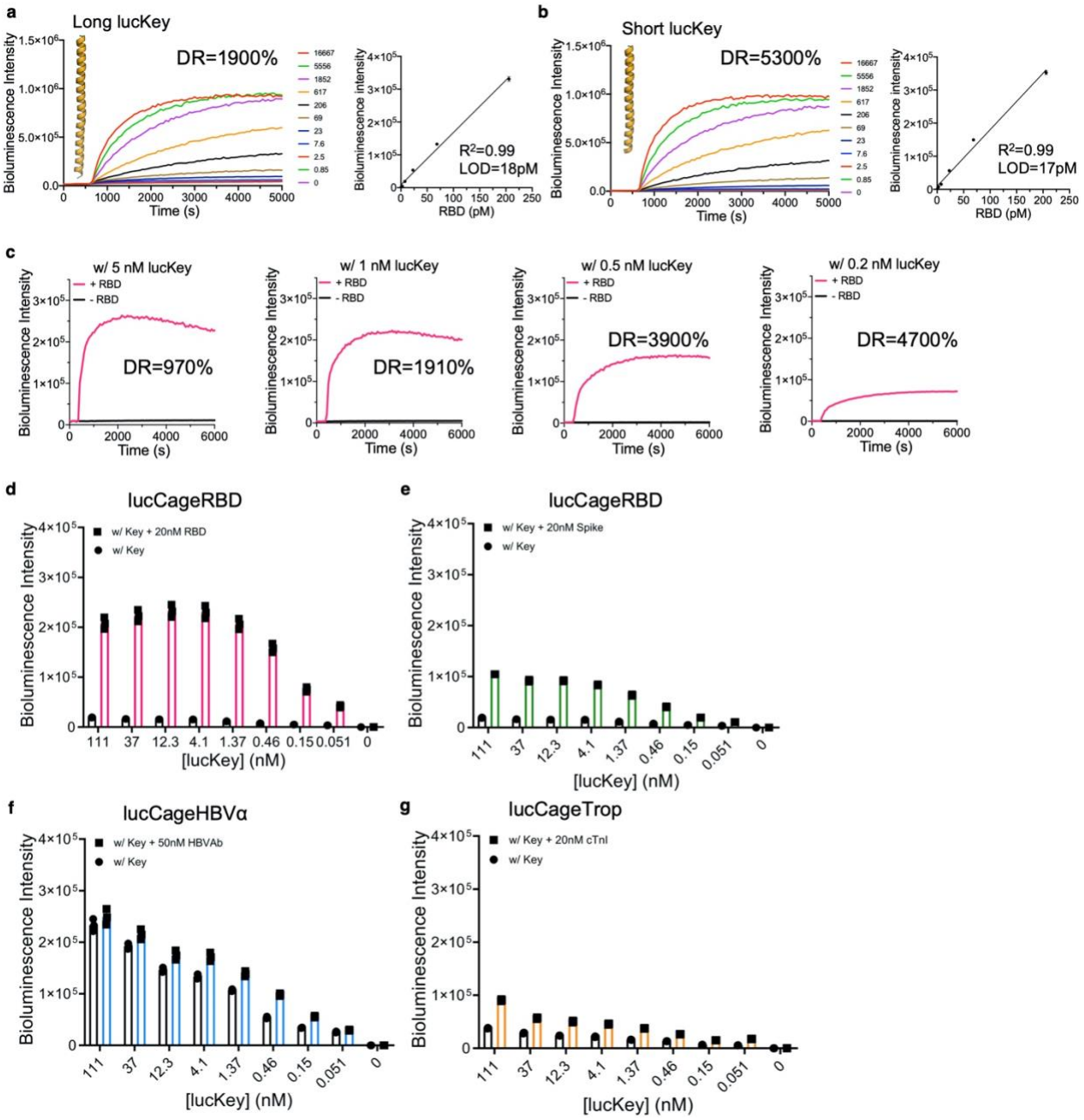
Extended Data Figure 8. Design and characterization of sensors for anti-SARS-CoV-2 antibodies. a-b, Experimental screening of *de novo* sensors for antibodies against the SARS-CoV-2 membrane protein (**a**), and the nucleocapsid protein (**b**). Selected epitopes of the membrane protein (M1, M3 and M4) and the nucleocapsid protein (N6 and N62) were computationally grafted into lucCage at different positions of the latch. Each design comprised two tandem copies of each epitope, separated by a flexible linker, to take advantage of the bivalent binding of antibodies. All designs were experimentally screened for increase in luminescence at 20nM of each lucCage design and 20nM of lucKey in the presence of anti-M rabbit polyclonal antibodies (ProSci, 3527) (**a**) or anti-N mouse monoclonal antibody at 100nM (clone 18F629.1) (**b**). These

experiments were performed in duplicate (**a**) or single replicate (**b**) in two independent instances, representative data are shown. The luminescence values were normalized to 100 in the absence of antibodies. Designs M3_1-17_334 and N62_369-382_340 were selected as the best candidates due to high sensitivity and stability, and were named lucCageSARS2-M and ucCageSARS2-N respectively. **c**, Left panel: structural model of lucCageSARS2-M, showing a close-up view of the predicted interface between the M3 epitope (red) and lucCage (blue). Middle panel: determination of lucCageSARS2-M sensitivity to anti-M pAb. Bioluminescence was measured over 4000 s in the presence of serially diluted anti-M pAb. From top to bottom - lucCageSARS2-M:lucKey concentration (nM) = 50:50, 5:5. Right panel: LOD calculations for the sensor at different concentrations. **d**, Left panel: structural model of lucCageSARS2-N, showing a close-up view of the predicted interface between the N62 epitope (purple) and lucCage (blue). Middle panel: determination of lucCageSARS2-N sensitivity to anti-N mAb. Bioluminescence was measured over 4000 s for lucCageSARS2-N + lucKey at 50 nM in the presence of serially diluted anti-N antibody. Right panel: LOD calculations for the sensor. All experiments were performed in triplicate unless specifically indicated, representative data are shown, and data are presented as mean values +/- s.d.



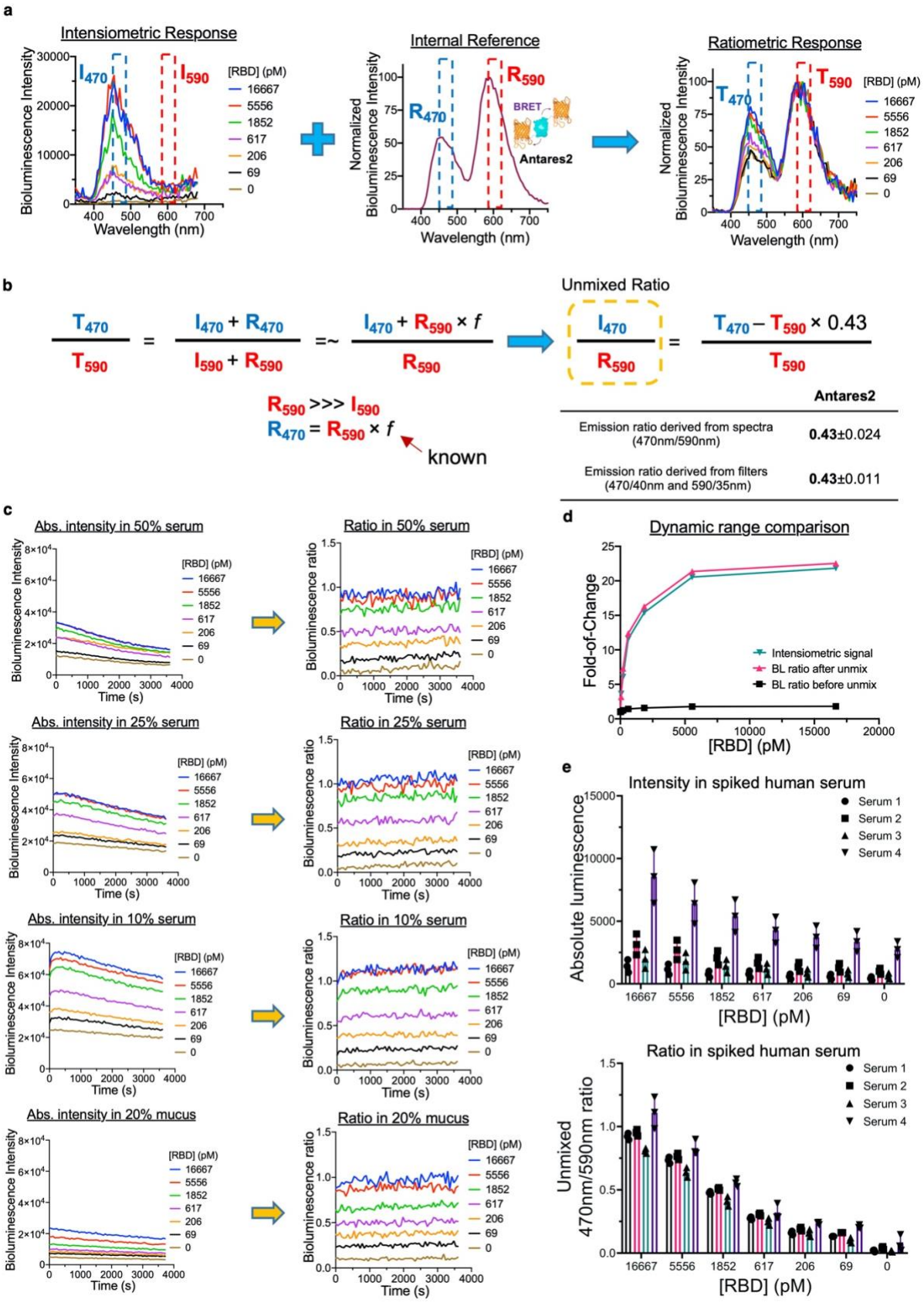
Extended Data Figure 9. Design and characterization of SARS-CoV-2 RBD sensors. a, Experimental screening of *de novo* sensors for the receptor-binding domain (RBD) of the SARS-CoV-2 Spike protein. All designs were experimentally screened for increase in luminescence at 20 nM of each lucCage design and 20 nM of lucKey in the presence of 200 nM RBD. The luminescence values were normalized to 100 in the absence of RBD. This experiment was performed in duplicate in two independent instances, representative data are shown. Design lucCageRBDdelta4_348 was selected as the best candidate due to high sensitivity and stability, and was named lucCageRBD. **b**, Structural model of lucCageRBD composed of the LCB1 binder (magenta) grafted into lucCage (blue) comprising a caged SmBiT fragment (gold). The black boxes show a close-up view of the interface of Cage (blue) and LCB1 binder (magenta) in the lucCageRBD design. **c**, Determination of lucCagerRBD's sensitivity. Bioluminescence was measured over 10000 s in the presence of serially diluted RBD protein. From top to bottom - lucCageRBD:lucKey concentration (nM) = 1:1, 1:10, 10:10. **d**, LOD calculations for the sensor at

different concentrations. From top to bottom - lucCageRBD:lucKey concentration (nM) = 1:1, 1:10, 10:10. **e**, Bioluminescence images acquired with a BioRad ChemiDoc imaging system. Changes in bioluminescence intensity levels were detected as a function of the concentration of RBD with lucCageRBD at 1 nM and lucKey at 10 nM. **f**, Detection of RBD in 10% simulated nasal matrix. Left: Bioluminescence was measured overtime in the presence of serially diluted RBD protein. Right: LOD was calculated to be 12 pM. **g**, Detection of spike protein in a 20% diluted pooled serum. Left: Bioluminescence was measured overtime in the presence of serially diluted HexaPro spike protein. Right: LOD was calculated to be 47 pM. All experiments were performed in triplicate unless otherwise indicated, representative data are shown, and data are presented as mean values +/- s.d.



Extended Data Figure 10. lucCageRBD tunability by varying the lucKey length (K_{CK}) and lucKey concentration (a-c) and the comparison of bioluminescent signals over a range of lucKey concentrations in the presence of target at saturating concentration (d-g). a-b, Experimental evaluation of the effect of K_{CK} on the dynamic range (DR) of lucCageRBD to detect monomeric SARS-CoV-2 RBD. A truncated lucKey (short lucKey), 14 residue shorter than the full-length key at its C-terminus (**b**), provides better dynamic range than the full-length lucKey (**a**) owing to reduced background signal, as predicted by the simulation in Extended Data Fig. 1f while the LOD remains the same. **c**, The effect of lucKey concentration on the dynamic range. Decreasing lucKey concentration increases the dynamic range of lucCageRBD due to reduced background signal, but with accompanying reduced maximum bioluminescence signal. **d-e,** lucCageRBD (1 nM) was incubated with RBD (20 nM, **d**) or spike protein (20 nM, **e**), which are expected to result in full reconstitution of the luciferase activity. In the presence of

spike protein, the same sensor was unable to yield the maximal bioluminescent signal, suggesting the effect of factors not captured by the simulations such as steric hindrance against complete luciferase reconstitution. **f**, lucCageHBVα (1 nM) incubated with 50 nM of the HBV antibody HzKR127-3.2 shows almost complete activation, but suffers from high background signal. **g**, lucCageTrop (1 nM) shows non-ideal background signal and moderate target-driven activation in the presence of 20 nM cTnI. All experiments were performed in triplicate, representative data are shown, and data are presented as mean values +/- s.d.


a
Intensiometric Response
Internal Reference
Ratiometric Response
Bioluminescence Intensity
Normalized Bioluminescence Intensity
Wavelength (nm)
I470
I590
R470
R590
T470
T590
BRET
Antares2
[RBD] (pM)
16667
5556
1852
617
206
69
0
b
Unmixed Ratio
R590 >>> I590
R470 = R590 × f
known
Antares2
Emission ratio derived from spectra (470nm/590nm)
0.43±0.024
Emission ratio derived from filters (470/40nm and 590/35nm)
0.43±0.011
c
Abs. intensity in 50% serum
Ratio in 50% serum
Abs. intensity in 25% serum
Ratio in 25% serum
Abs. intensity in 10% serum
Ratio in 10% serum
Abs. intensity in 20% mucus
Ratio in 20% mucus
Bioluminescence Intensity
Bioluminescence ratio
Time (s)
d
Dynamic range comparison
Fold-of-Change
Intensiometric signal
BL ratio after unmix
BL ratio before unmix
[RBD] (pM)
e
Intensity in spiked human serum
Absolute luminescence
Serum 1
Serum 2
Serum 3
Serum 4
Ratio in spiked human serum
Unmixed 470nm/590nm ratio

Extended Data Figure 11. Integration of Antares2 as the internal reference for calibration of lucCageRBD in different biological matrices. a, The bioluminescent emission spectra of lucCageRBD (Left) in response to varying concentrations of RBD. Antares2 is an efficient CyOFP1-teLuc-CyOFP1 BRET system[40] with a peak emission at 590 nm (Middle). The emission spectra were recorded from a mixture of lucCageRBD and lucKey (both at 1 nM), Antares2 (0.1 nM) and RBD at varying concentrations (Right). By acquiring the individual signal from 470/40 nm and 590/35 nm channels, the intensiometric responses from lucCageRBD were converted into ratiometric readouts. **b,** Equations to calculate the spectrally unmixed ratio. The total signal from the 470/40 nm channel (T_{470}) is the sum of the signals from the lucCageRBD sensor (I_{470}) and the Antares2 reference (R_{470}), while the total signal from the 590/35 nm channel (T_{590}) is equal to the sensor signal (I_{590}) plus reference signal (R_{590}). Since lucCageRBD gives negligible emission at 590/35 nm channel, T_{590} is approximately equal to R_{590} ($R_{590}>>I_{590}$). R_{470} is $R_{590} \times f$, a predetermined constant for Antares2, and therefore the unmixed ratio (I_{470}/R_{590}) could be calculated in real time during signal acquisition. The constant f for Antares2 was consistently determined to be 0.43 by either recording the full spectra or from the filter set. **c,** RBD at varying concentrations were spiked in 50%, 25%, 10% pooled serum or in 20% simulated nasal fluid. Absolute bioluminescence intensities and the emission kinetics were different across the matrices due to matrix inhibition effect and substrate turnover. In contrast, calibration with Antares2 resulted in stable ratiometric signals (I_{470}/R_{590}). **d,** The bioluminescence intensity of lucCageRBD at saturating RBD concentration (green curve) is ~20 folds higher than the background level. Reporting the raw ratio (T_{470}/T_{590}) as a function of the RBD concentration compromises the sensor dynamic range (black curve) due to a significant emission at 470/40 nm channel (R_{470}) from Antares2. After calculation and conversion of the unmixed ratio, the dynamic range becomes ~20 folds over the background with ratiometric readouts (magenta curve). **e,** Detection of spiked RBD in four different anonymized human sera (50%) shows that calibration using spectrally resolved Antares2 as an internal reference can minimize the variations of the intensiometric bioluminescence in these matrices. Bioluminescent signals and s.d. were measured in triplicate, and a representative one is shown for emission spectra and emission kinetics, respectively. Data are presented as mean values +/- s.d.

Supplementary Information

Supplementary discussion:

Sensor proteins have emerged as an active area of research. Traditional ELISA methods require multiple liquid-handling steps, preventing its use at the bedside. Lateral flow immunochromatographic assays are fast and cheap, but they have limited sensitivity, reproducibility, and poor quantitative performance [1]. ELISA and lateral flow also require two binding modules for the target being sensed, one for capture and the other for readout. Sensor proteins capture analytes with excellent molecular precision and generate signals in a homogeneous way that do not require time-consuming incubation and washing steps of traditional immunoassays. A typical sensor protein is composed of two components - an analyte-binding domain for target capturing, and a reporter component to convert and amplify the actuation into a measurable signal. The main hurdle of the protein sensor construction is finding analyte binding domains that undergo sufficient conformational changes. The most commonly used binding domains (e.g., antibodies) undergo only minor structural changes of the loops upon ligand binding. Coupling an appropriate reporter with optimal geometry to amplify the conformational change is also key to a successful biosensor. Protein engineers have devoted considerable work to build biosensors for a specific application, constructing tailored protein sensors remains a major challenge [2,3]. Computational [4], synthetic [5], and semisynthetic [6] strategies have also aimed to develop generalizable platforms for arbitrary biosensors. However, computationally designing small molecule binding sites into protein interfaces and generating semisynthetic protein sensors are both quite challenging problems currently. Therefore, generalized approaches for designing biosensors with a simple and robust computational protocol that requires little empirical optimization are needed.

Herein, we select and compare several state-of-the-art sensor configurations and design strategies that seek to detect protein targets (Supplementary Fig. 1 and Table S5). These strategies generally include a protein switch at an initial "closed" state by an intramolecular β-lactamase-to-inhibitor (Supplementary Fig. 1a), Src domain-to-proline rich peptide (Supplementary Fig. 1b), or benzenesulfonamide-to-carbonic anhydrase (Supplementary Fig. 1c) interaction. Upon binding to the target, the conformational change triggers the dissociation of these interactions eventually to an "open" state, resulting in the active β-lactamase for colorimetric assay (Supplementary Fig. 1a) or a decreased BRET efficiency (Supplementary Fig. 1b,c). Both assays exhibit ideal features for POC diagnostics such as wash-free, sensitive (pM-nM), fast (min), and simple readout (camera) to detect therapeutic antibodies, protein biomarkers, and virus particles in patients' serum or plasma [7–10]. In terms of design, their target specificity can be easily switched by swapping known binding domains (epitope, affimer, or inhibitor). The length and type of semiflexible linkers play the most critical role in the success rate, which has to be re-designed or re-optimized for each binding domain swap. In some cases, the interaction partners to hold the switch at a "close" state also need to be re-evaluated [11]. The abovementioned engineering work is unavoidable primarily due to the limited understanding of the dynamic nature of protein interaction, folding, and flexibility. Our generalized protein sensory system based on a *de novo* switch (Supplementary Fig. 1d) relies on the thermodynamic coupling (see Fig. 1) between a defined close state (K_{open}) and an open state (K_{LT} and K_{CK}). With our system (Table S5), the target specificity to arbitrary targets can be achieved not only by incorporating known binding domains but also de novo binders where we have full control over protein fold and geometry. Because there is no flexible or semi-flexible

linker in our system and we are capable of designing different types of interaction to cage binding domains, the conformational change is thus decoupled from the binder-target interaction, which makes this system more structurally predictable at open state. A newly developed GraftSwitchMover in Rosetta allows sensor design in one step, bypassing the need with the other formats to empirically re-engineer sensor configuration. The intermolecular association of the LucKey with the open form of the sensor generates the luminescent signal, providing an additional tunable parameter K_{CK} that can be optimized along with K_{open} to maximize sensor dynamic range, analytical range, and sensitivity.

It is important to note that the sensors we describe here have not been extensively optimized for POC settings. Fluorescence- or colorimetric-based assays suffer from autofluorescence, light scattering, and interference in whole-blood samples, limiting these assays for immediately quantifiable POC applications [12]. Background-free bioluminescence readout is ideal and ~100-fold more sensitive than fluorescence [13,14], but the absolute bioluminescence intensity is known to be influenced by factors in serum or whole blood not related to analyte [15,16]. Incorporating an internal reference as we showed here or transition to a ratiometric BRET response would facilitate the adaptation of the sensors for optimal POC settings. Furthermore, the sensitivity of protein sensors described here has not been fully tuned to cover the complete linearity of the clinical diagnostic range. In addition, there are pros and cons for two-component sensors where the performance and sensitivity depend on sensor component concentrations [17]. This tunability can be advantageous at the developmental stage but could become problematic for accurate and reproducible quantification in clinical settings. To avoid such problems, the lucKey could be attached to lucCage through a flexible linker.

Supplementary methods

1. Thermodynamic model

The thermodynamic model shown in Fig. 1b was adapted and modified from the previously published *de novo* switch model [18]. The equilibrium constants were defined as K_{open} for latch opening (Equation 1), K_{CK} for the dissociation constant of the lucCage and lucKey (Equation 2 and 3), and K_{LT} for the dissociation constant of the latch and target (Equation 4 and 5). K_R describes the equilibrium of the reconstituted luciferase, which is determined by the reported dissociation constant of the NanoBit system (190 μM [19]) and the effective local concentration (C_{eff}) of split counterparts (Equation 6 and 7). We set C_{eff} to 1 mM here as the literature suggested high micromolar to low millimolar range for intramolecular interaction partners [20], and our modular switch should span much shorter distance than flexible linkers. The total amount of each component is constant, so Equations 8, 9, and 10 were introduced. Given four equilibrium constants (K_{open}, K_{CK}, K_{LT}, and K_R) and three total concentrations ($[lucCage]_{total}$, $[lucKey]_{total}$, and $[target]_{total}$), python module sympy.nsolve was used to solve these ten equations numerically and find the concentration of each species at equilibrium. The total concentration of luminescent species **6** and **7** was extracted from the solution, divided by $[lucCage]_{total}$, and plotted for corresponding figures with various K_{open} for Extended Data Fig. 1a, K_{LT} for Extended Data Fig. 1b, $[lucCage]_{total}$, $[lucKey]_{total}$ for Extended Data Fig. 1c,d and K_{CK} for Extended Data Fig. 1e. The fraction of reconstituted luciferase is normalized between 0-1. R_{max} and R_{min} are the fraction of reconstituted luciferase at saturating and at low target concentration respectively. The sensor dynamic range is derived from the equation $(R_{max} - R_{min}) / R_{min} \times 100\%$**.** The heatmaps in Extended Data Fig. 1f and 1g were generated with a 2D array of indicated K_{open} and K_{CK} values, to plot the dynamic range and the fraction of reconstituted luciferase at saturated concentration, respectively. The code for the numerical simulations shown in this work are available at http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/model_simulation.py

Equation 1:

$$K_{open} = \frac{[2]}{[1]}$$

Equation 2:

$$K_{CK} = \frac{[2] \times [luckey]_{free}}{[5]}$$

Equation 3:

$$K_{CK} = \frac{[3] \times [luckey]_{free}}{[4]}$$

Equation 4:

$$K_{LT} = \frac{[6] \times [target]_{free}}{[7]}$$

Equation 5:

$$K_{LT} = \frac{[2] \times [target]_{free}}{[3]}$$

Equation 6:

$$K_R = \frac{190\ \mu M}{C_{eff}} = \frac{[5]}{[6]}$$

Equation 7:

$$K_R = \frac{[4]}{[7]}$$

Equation 8:

$$[lucCage]_{total} = [1] + [2] + [3] + [4] + [5] + [6] + [7]$$

Equation 9:

$$[lucKey]_{total} = [lucKey]_{free} + [4] + [5] + [6] + [7]$$

Equation 10:

$$[target]_{total} = [target]_{free} + [3] + [4] + [7]$$

2. Computational grafting of sensing domains into lucCage

The main feature sought for when incorporating new sensing domains into lucCage is whether a target-binding interface can be blocked by the cage domain. Finding the right orientation is usually simple when grafting small peptides. In the case of globular proteins being used as sensing domains, finding the right orientation is often facilitated by finding target-binding interfaces neighboring the N-terminus of the sensing domain. In our experience, sensing domains with an α-helical segment at the N-terminus are preferable to ensure a more predictable design process, though this is not a requirement. Binding domains with other secondary structure elements have been effectively incorporated into lucCage, as shown here with the SmBiT peptide (β-strand), and the PreS1 peptide and SARS-CoV-2 epitopes (loops). These features of the lucCage platform make it compatible with diverse sensing domains, however, it can be challenging to use analyte binding domains with complex and/or distant target-binding interfaces or that are bigger in size than the lucCage itself.

The structural models of the lucCage sensors were created by grafting each sensing domain onto the latch of the lucCage scaffold (See Table S3). The design was performed using a RosettaScripts[21] protocol, (*{name}_GraftSwitchMover_relax.xml*, See Code availability) to thread each sensing domain sequence into the model of lucCage (*lucCage.pdb*, See Code Availability). A bash script (*run_rosettascripts.sh*) was used to call RosettaScripts. This protocol uses two successive Rosetta movers: (i) GraftSwitchMover to thread the desired sensing domain sequence into a defined region of the lucCage latch (amino acids 325-359) and to select designs with the defined "important residues" that are buried in the cage/latch interface; (ii) and MultiplePoseMover to capture all models generated by GraftSwitchMover, minimize them (FastRelax[22] to find the lowest energy structure given the mutations from the previous mover.) and score each output model. The resulting designs were further evaluated by eye in PyMol 2.0. This was done by selecting designs showing favorable hydrophobic packing interactions between the newly threaded sequence and the cage and discarding designs with unfavorable buried hydrophilic residues that could destabilize the closed state of the sensor (unless these residues were identified as "important residues"). At the time of publication, GraftSwitchMover is limited to inserting a sequence within the range of the input structure (i.e., the inserted sequence cannot exceed the full length of the initial lucCage model being used). In order to design sensors that would extend the latch of lucCage (e.g., insertion of cTnT), we used Rosetta Remodel[23] to model the extended C-terminus of lucCage (*{name}_Remodel_relax.xml*), which uses a blueprint file (*{name}.bp*) called by a bash script containing all necessary flags (*run_rosetta_scripts_REMODEL.sh*).

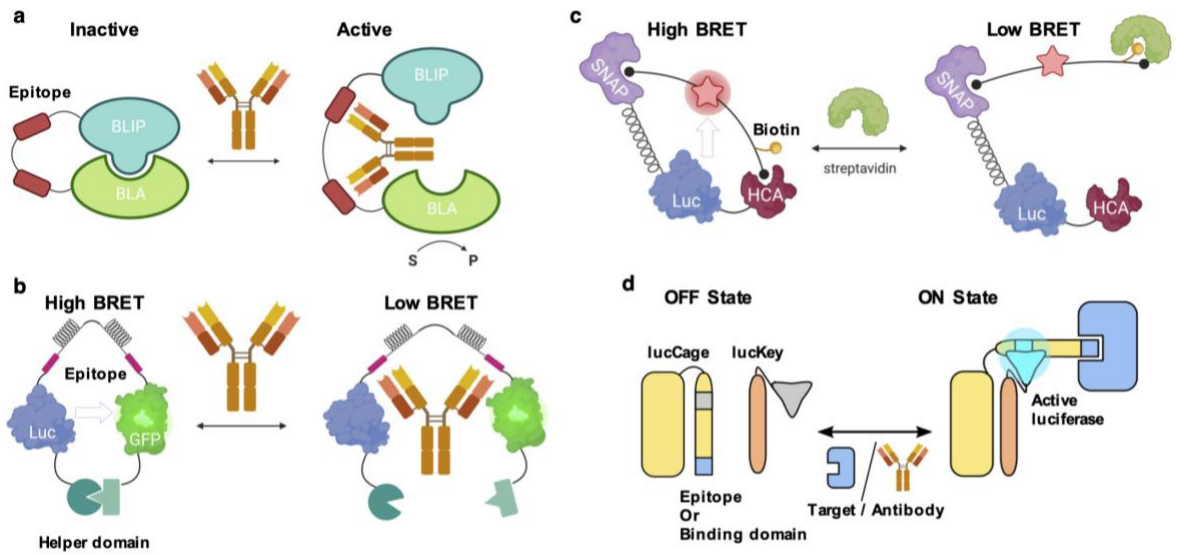
For grafting mini-protein binders with a pre-defined tertiary structure (i.e., Bot.671.2, SpaC, the Her2 affibody, and LCB1) we first identified the primary interaction surface of the binding protein to its target and identified the main secondary structure elements involved in it. We identified the main amino acid sequence of the N-terminal secondary structure element and threaded it into lucCage ({name}*_GraftSwitchMover_relax.xml)*, to obtain output protein models including the partial sequence of a sensing domain onto the latch. At this point, any latch residues remaining after the sensing domain were manually removed from the pdb file to prepare the protein model for the next step. Then, we used an additional script (*{name}_MergePDB_relax.xml*) to align the resulting lucCage designs with the model (or crystal structure) of the full-length sensing domain, create a complete model including the full sensing domain in the context of the switch, and energy-minimize the designs. The input sensing domains used are named as: Her2_3mzw.pdb,

Bot_5vid.pdb, ProA_4WWI.pdb, LCB1_7jzu.pdb, and delta4LCB1_7jzu.pdb. Finally, the best designs were selected by eye using PyMol 2.0. The design models and RosettaScripts code used in this work have been deposited to http://files.ipd.uw.edu/pub/de_novo_design_of_tunable_biosensors_2021/designcode_and_models.zip

Supplementary Information references

1. Bahadır, E. B. & Sezgintürk, M. K. Lateral flow assays: Principles, designs and labels. *Trends Analyt. Chem.* **82**, 286–306 (2016).
2. Yeh, H.-W. & Ai, H.-W. Development and Applications of Bioluminescent and Chemiluminescent Reporters and Biosensors. *Annu. Rev. Anal. Chem.* **12**, 129–150 (2019).
3. Greenwald, E. C., Mehta, S. & Zhang, J. Genetically Encoded Fluorescent Biosensors Illuminate the Spatiotemporal Regulation of Signaling Networks. *Chem. Rev.* **118**, 11707–11794 (2018).
4. Glasgow, A. A. *et al.* Computational design of a modular protein sense-response system. *Science* **366**, 1024–1028 (2019).
5. Guo, Z. *et al.* Generalizable Protein Biosensors Based on Synthetic Switch Modules. *J. Am. Chem. Soc.* **141**, 8128–8135 (2019).
6. Yu, Q. *et al.* Semisynthetic sensor proteins enable metabolic assays at the point of care. *Science* **361**, 1122–1126 (2018).
7. van Rosmalen, M. *et al.* Dual-Color Bioluminescent Sensor Proteins for Therapeutic Drug Monitoring of Antitumor Antibodies. *Anal. Chem.* **90**, 3592–3599 (2018).
8. Adamson, H. *et al.* Affimer–Enzyme–Inhibitor Switch Sensor for Rapid Wash-free Assays of Multimeric Proteins. *ACS Sens.* **4**, 3014–3022 (2019).
9. Tenda, K. *et al.* Paper-Based Antibody Detection Devices Using Bioluminescent BRET-Switching Sensor Proteins. *Angewandte Chemie* **130**, 15595–15599 (2018).
10. Griss, R. *et al.* Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring. *Nat. Chem. Biol.* **10**, 598–603 (2014).
11. Arts, R. *et al.* Detection of Antibodies in Blood Plasma Using Bioluminescent Sensor Proteins and a Smartphone. *Anal. Chem.* **88**, 4525–4532 (2016).
12. Lopez-Ruiz, N. *et al.* Smartphone-based simultaneous pH and nitrite colorimetric determination for paper microfluidic devices. *Anal. Chem.* **86**, 9554–9562 (2014).
13. Leippe, D. M. *et al.* A bioluminescent assay for the sensitive detection of proteases. *Biotechniques* **51**, 105–110 (2011).
14. Troy, T., Jekic-McMullen, D., Sambucetti, L. & Rice, B. Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. *Mol. Imaging* **3**, 9–23 (2004).
15. Yeh, H.-W., Wu, T., Chen, M. & Ai, H.-W. Identification of Factors Complicating Bioluminescence Imaging. *Biochemistry* **58**, 1689–1697 (2019).
16. Yeh, H.-W. *et al.* ATP-Independent Bioluminescent Reporter Variants To Improve in Vivo Imaging. *ACS Chem. Biol.* **14**, 959–965 (2019).
17. Edwardraja, S. *et al.* Caged activators of artificial allosteric protein biosensors. *ACS Synth. Biol.* (2020) doi:10.1021/acssynbio.9b00500.
18. Langan, R. A. *et al.* De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
19. Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
20. Krishnamurthy, V. M., Semetey, V., Bracher, P. J., Shen, N. & Whitesides, G. M. Dependence of Effective Molarity on Linker Length for an Intramolecular Protein−Ligand System. *Journal of the American Chemical Society* **129**, 1312–1320 (2007).
21. Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6**, e20161 (2011).
22. Khatib, F. *et al.* Algorithm discovery by protein folding game players. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18949–18953 (2011).

23. Huang, P.-S. *et al.* RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS One* **6**, e24109 (2011).
24. Banala, S., Aper, S. J. A., Schalk, W. & Merkx, M. Switchable reporter enzymes based on mutually exclusive domain interactions allow antibody detection directly in solution. *ACS Chem. Biol.* **8**, 2127–2132 (2013).
25. Schena, A., Griss, R. & Johnsson, K. Modulating protein activity using tethered ligands with mutually exclusive binding sites. *Nat. Commun.* **6**, 7830 (2015).



Supplementary Fig. 1. Working mechanisms of recent biosensor platforms to detect protein targets. a, An antibody-sensing enzyme-inhibitor switch. The initially inactive TEM1-β-lactamase (BLA) is tethered to its inhibitor protein (BLIP) via a semi-flexible linker. Antibody binding to two epitopes in the semi-flexible linker triggers the dissociation of BLA-BLIP, leading to active BLA as the reporter[24]. **b**, LUMinescent AntiBody Sensor platform, LUMABS. High BRET efficiency is achieved when the luciferase (Luc) donor and the GFP acceptor are in close proximity mediated by an interaction between a Src Homology 3 domain and a proline-rich peptide (helper domain). Binding of an antibody to epitope sequences in the linker disrupts the interaction between these helper domains, resulting in loss of BRET efficiency[7]. **c**, Chemical ligand–associated steric hindrance platform, CLASH. The intramolecular tether contains one benzenesulfonamide ligand for human carbonic anhydrase (HCA) binding and one benzylguanine moiety for SNAP-tag. One additional synthetic biotin on the tether can bind to streptavidin as a specific effector protein, which increases the overall steric hindrance and prevents the interaction of benzenesulfonamide to the HCA domain, leading to a low BRET efficiency open state[25]. **d**, Modular biosensor platform based on a *de novo* switch. A structurally preorganized lucCage containing a caged sensing domain (or epitope) stabilizes the complex in a closed state. The sensing domain binding to the target of interest (or the epitope binding to the antibody of interest) facilitates the association of lucCage and lucKey into an open state, which is reported by bioluminescence emission after split-luciferase reconstitution.

Table S1. X-ray data collection and structure refinement statistics.

Data Collection	SelMET-sCageHA_267-1S(E99Y/T144Y/I30M)
Space group	*C*2
Unit cell dimensions	
a, b, c (Å)	178.993 60.127 71.799
a, β, γ (°)	90, 112.463, 90
Wavelength (Å)	0.9794
Resolution (Å)	50-2.03 (2.03-2.00)[a]
R_{sym}	6.6 (16.5)[a]
$I/\sigma(I)$	24.0 (3.5)[a]
Completeness (%), >1σ	70.6 (33.8)[a]
Redundancy	2.5 (1.6)[a]
Refinement	
Resolution (Å)	46.09-1.99 (2.06-1.99)[a]
No. of reflections	37603
R_{work} / R_{free}	0.2078/0.2515
R.m.s deviations	
bond (Å) / angle (º)	0.007/0.910
Average B-values (Å^2)	38.19
Ramachandran plot (%)	
Favored / Additional allowed	97.7/2.3
Generously allowed	0.0

[a]The numbers in parentheses are the statistics from the highest resolution shell.

Table S2. Summary of biosensors in this work

Biosensor	Analytical target	Dynamic range[a]	LOD (nM)	Detection range (nM)
lucCageBim	Bcl-2	360%	0.2	0.2-12.5
lucCageBoT	Botulinum neurotoxin B	130%	0.4	0.4-50
lucCageProA	Fc domain	350%	0.39	0.39-12.5
lucCagHer2	Her2 receptor	380%	0.23	0.23-25
lucCageTrop	Troponin I	250%	0.009	0.009-0.3
lucCageHBV	Anti-HBV antibody (HzKR127-3.2)	98%	2	2-100
lucCageHBVα	Anti-HBV antibody (HzKR127-3.2)	225%	0.26	0.26-100
lucCageHBV+ HzKR127-3.2	PreS1	80%	0.6	0.6-100
lucCageSARS2-M	anti-SARS-Cov-M	50%	2.9	2.9-250
lucCageSARS2-N	anti-SARS-Cov-NP	70%	3.0	3.0-100
lucCageRBD	SARS-CoV-2 RBD	1700%	0.015	0.015-6

[a]Defined as intensiometric change ($\Delta E/E_{min}$) of total bioluminescence intensity. $\Delta E = E_{max} - E_{min}$ where E_{max} is the maximal bioluminescence intensity at saturating target concentration and E_{min} is the emission in the absence of the analytical target.

Table S3. List of sensing domains used in this work

Biosensor	Sensing domain	Sensing domain sequence
lucCageBim	Bim	EIWIAQELRRIGDEFNAYYA
lucCageBoT	Bot.0671.2	MFAELKAKFFLEIGDRDAARNALRKAGYSDEEAERIIRKYELE
lucCageProA	Protein A domain C (SpaC)	EQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLNDAQAPK
lucCagHer2	Her2 affibody	EMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLNDAQAPK
lucCageTrop	cTnI + cTnC	EDQLREKAKELWQTIYNLEAEKFDLQEKFKQQKYEINVLRNRINDNQKVSKTKDDSKGKSEEELSDLFRMFDKNADGYIDLEELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVE
lucCageHBV	preS1 (a.a. 35–46)	GANSNNPDWDFN
lucCageHBVα	preS1 (a.a. 35–46) 2x	GANSNNPDWDFNGGSGGGSSGFGANSNNPDWDFNPN
lucCageSARS2-M	SARS-CoV-2 nucleocapsid protein (a.a. 369-382) 2x	MADSNGTITVEELKKLLEGGSGGMADSNGTITVEELKKLLE
lucCageSARS2-N	SARS-CoV-2 membrane protein (a.a. 1-17) 2x	KKDKKKKADETQALGGSGGKKDKKKKADETQAL
lucCageRBD	LCB1	ILQKIYEIMRLLDELGHAEASMRVSDLIYEFMKKGDERLLEEAERLLEEVER

Table S4. Comparison of background intensity and dynamic range between the lucCage sensors under the same experimental conditions

Biosensor	No target (% ± SD)	Saturating target (% ± SD)	Fold Activation
lucCageBim	6.685 ± 0.056	21.689 ± 0.223	3.24
lucCageBoT	1.261 ± 0.030	2.419 ± 0.022	1.92
lucCageProA	11.403 ± 0.030	42.503 ± 0.786	3.73
lucCagHer2	2.989 ± 0.010	9.562 ± 0.300	3.20
lucCageTrop	8.773 ± 0.057	30.717 ± 0.420	3.50
lucCageHBVα	62.945 ± 0.896	87.414 ± 0.385	1.39
lucCageSARS2-M	22.286 ± 0.432	28.357 ± 0.836	1.27
lucCageRBD	7.340 ± 0.126	100 ± 1.731	13.62

Percentages calculated assuming activation of lucCageRBD at saturating target concentration as 100% luciferase intensity.
Experimental conditions: (1nM lucCage + 1nM lucKey in 1:1 of buffer 0.01 M HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20 Nano-Glo assay buffer). These conditions may differ from optimal experimental conditions for some lucCage sensors, hence, the reduced dynamic range for some of them was observed.

Table S5. Summary of recent biosensor platforms to detect protein targets

Platform	Enzyme–inhibitor switch	LUMABS	CLASH	lucCage-lucKey (this work)
Switch module	BLA-BLIP	SH3-PRP	HCA-Nluc-p30-SNAP	Two components (lucCage and lucKey)
Switch tunability	Introducing point mutations	Replacing helper domain	Changing the affinity of tethered ligands	Computational interface design
Binder geometry	Small linear epitope or affimer	Small linear epitope or mimotope	Small molecule-target interaction	*De novo* binders, affibody, linear epitope, natural binders
No. of Binding sites required	2	2	1	1
Linker module	Semi-flexible or flexible linker	Semi-flexible linker	30-proline linker	No linker
Readout	Colorimetric by nitrocefin	BRET (450/517nm)	BRET (450/570nm)	Intensiometric in luminescence
Target tunability	Swapping binding motif	Swapping binding motif	Changing ligand	Swapping binding motif
Sensitivity	pM	pM	nM	pM
Sensitivity tunability	Monovalent affinity & avidity	Monovalent affinity & avidity	Mutations in the receptor protein	Monovalent affinity & avidity (K_{LT})
Dynamic range tunability	Linker optimization	Replacing helper domain	Geometry optimization	Cage-latch & Cage-key interface design (K_{open} & K_{CK}). Key length.
Number of sensors built	3: Herceptin, hCRP, plant virus	7: therapeutic antibodies	2: streptavidin, DHFR	10: Bcl-2, BoNT/B, Fc, Her2, cTnI, anti-HBV, HBV PreS1, SARS-CoV-2 RBD, anti-SARS-CoV-N, anti-SARS-CoV-M