

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Antibody control of protein conformation in pathology and for therapeutic pharmacology

Permalink

<https://escholarship.org/uc/item/3d56h5pt>

Author

Faust, Bryan Thomas

Publication Date

2022

Supplemental Material

<https://escholarship.org/uc/item/3d56h5pt#supplemental>

Peer reviewed|Thesis/dissertation

Antibody control of protein conformation in pathology and for therapeutic pharmacology

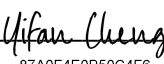
by
Bryan Faust

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in
Biophysics

in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

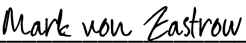
DocuSigned by:

87A0F4E0B50C4F6... Yifan Cheng
Chair

DocuSigned by:

Aashish Manglik

DocuSigned by:

Adam Frost

DocuSigned by:

A4DF1018F0A2414... Mark von Zastrow

Committee Members

Copyright 2022

by

Bryan Faust

Dedication

To the memory of my mother, Anne Faust. Her perseverance fueled mine.

This is your dissertation, too, Mom.

Acknowledgements

I did not get here alone. I want to first recognize my two exceptional advisors, Yifan Cheng and Aashish Manglik. At the beginning of graduate school, I did not know what being co-advised would entail, but now have come to realize that it meant having two champions in your corner – scientifically, professionally, and personally. Their openness to my desired co-advised research direction, encouragement in the “optimization trenches”, and empathy through some tough sections of graduate training are only some of the reasons I made it through. I will continue to carry their limitless enthusiasm for every facet of science throughout my career.

To the remarkable scientists I have learned from throughout my time at UCSF: Nick Hoppe, Julian Harris, Chase Webb, Ben Barsi-Rhyne, Simone Harrison, Jiahao Liang, Christian Billesbølle, Ishan Deshpande, Allie Born, Daniel Asarnow, Evan Green, Eugene Palovcak, Adamo Mancino, Marcell Zimanyi, Henriette Autzen, Cristina Puchades, Caleigh Azumaya, Junrui Li, Jianhua Zhao, and Migliang Jin – each of you has played a significant role in my scientific development and I am so thankful to have experienced this time together and for the opportunity to call all of you friends.

My pre-UCSF scientific mentors: Sandro Vivona, Jeremy Wilbur, and Bob Stull – thank you all for believing I could actually do this, for establishing an environment where I could grow professionally at Stemcentrx, and for keeping me on my toes, scientifically, year after year.

To my closest friends who have stood by along the highs and lows of graduate school and the external challenges I faced along the way: Richard Erb, Adam Zuckerman, Matthew Menninger,

Gino Grilli, Sara Okholm, Ronan Hayes, Shannon Burke, and Michael Smith – you all provided a welcome refuge from the throes of graduate school, kept me grounded, and are amongst the most important people in my life.

To my family: my father, Curt Faust, my brother Troy Faust, my sister Brittany Owens, my brother-in-law Brandon Owens, and my grandfather Thomas Faust. I could not have done this without your unwavering support from afar. Dad, you are always willing to listen to how the day went in the lab, and constantly push me to excel in life. Thank you for instilling your sense of curiosity in me. Troy, every conversation with you is a good one. I look forward to many more backpacking trips in the future and am so proud to be your brother. Britt and Brandon, thank you for being a sounding board on so many aspects of life. I look up to you both. Grandpa, thank you for always asking if I was done with graduate school when I called – it was great motivation to move things along.

I would be remiss to not also mention my “little dude” George (the Welsh terrier). Thank you for keeping me sane and being an attentive listener during my countless practice talks.

Lastly, to my wife, Ellen: There a few words to express the gratitude I feel towards you and for the life we’re building. From the initial decision to start this program through every crossroad, you have been my biggest ally. Thank you for being so patient with my late nights in lab, for always being willing to proofread my esoteric scientific work, and for keeping me loved and well-fed throughout this journey. Because of you, I am a better scientist and person.

Contributions

The chapters in this dissertation were performed under the guidance of Dr. Aashish Manglik and Dr. Yifan Cheng. Dr. Adam Frost and Dr. Mark Von Zastrow also provided key scientific guidance on facets of each chapter. Specifically, chapter 1 includes results formally accepted for publication in *Nature* from a manuscript titled “Autoantibody mimicry of hormone action at the thyrotropin receptor” written by **Faust B**, Billesbølle CB, Suomivuori CM-, Singh I, Zhang K, Hoppe N, Pinto AFM, Diedrich J, Muftuoglu Y, Szkudlinski MW, Saghatelian A, Dror RO, Cheng Y, Manglik A. Faust B and Billesbølle CB contributed equally to this work. At the time of dissertation submission, these results are pending publication and citation information though represent the final accepted form of the manuscript.

Chapter 2 includes a reprint of “An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike” as it appears in Schoof M, **Faust B**, Saunders RA, Sangwan S, Rezelj V, Hoppe N, Boone M, Billesbølle CB, Puchades C, Azumaya CM, Kratochvil HT, Zimanyi M, Deshpande I, Liang J, Dickinson S, Nguyen HC, Chio CM, Merz GE, Thompson MC, Diwanji D, Schaefer K, Anand AA, Dobzinski N, Zha BS, Simoneau CR, Leon K, White KM, Chio US, Gupta M, Jin M, Li F, Liu Y, Zhang K, Bulkley D, Sun M, Smith AM, Rizo AN, Moss F, Brilot AF, Pourmal S, Trenker R, Pospiech T, Gupta S, Barsi-Rhyne B, Belyy V, Barile-Hill AW, Nock S, Liu Y, Krogan NJ, Ralston CY, Swaney DL, García-Sastre A, Ott M, Vignuzzi M; QCRG Structural Biology Consortium, Walter P, Manglik A. *Science*. 2020 Dec 18; **370**(6523):1473-1479. Schoof M, Faust B, Saunders RA, Sangwan S, Rezelj V contributed equally to this work.

The co-authors, Yifan Cheng, and Aashish Manglik directed and supervised the research described in this dissertation. Bryan Faust provided the experimentation and manuscript preparation for much of the following dissertation, which is a substantive contribution comparable to other dissertations in biochemistry and biophysics.

Each chapter in this dissertation reflects a substantial collective team effort without which none of this work would be possible.

Antibody control of protein conformation in pathology and for therapeutic pharmacology

Bryan Faust

Abstract

Proteins are the primary mediators of cellular activity and play a critical role in disease etiology and treatment. The conformation of proteins, either intrinsic or induced by other interacting partners, can inform their localization, their capacity to transmit signaling cascades, and, ultimately, their function. Immunoglobulins, or antibodies, are glycoproteins generated by the immune system with the primary function of protecting the host by blocking or neutralizing the spread of pathogens. The blocking activity of antibodies can proceed via simple mechanisms such as occlusion of interacting interfaces or through more complex allosteric pathways for blocking pathogenic protein function. Paradoxically, immune system dysregulation can also produce self-reactive, pathogenic antibodies causative for disease. Amongst these disparate examples is a shared feature: the capacity for antibodies to effect protein function through conformational control of their interacting partner. This dissertation uses a combination of cryo-electron microscopy, pharmacological signaling studies, and biochemical reconstitutions to understand how natural, patient-derived, self-reactive antibodies can select for active state conformations of the thyrotropin G protein-coupled receptor, and how synthetic single-domain antibodies can restrict SARS-CoV-2 S protein conformation and subsequent viral entry.

In the first chapter, I and others structurally describe the activation mechanism of the thyrotropin receptor both in the context of normal, hormone-mediated signaling and in pathogenic, autoantibody-induced activation. These results describe the hormone mimicry and ensuing conformational selection employed by autoantibodies in the pathogenesis of Graves' disease via their action at the thyrotropin receptor. This work is the first to structurally describe how

autoantibodies can activate G protein-coupled receptors and further refines the model for glycoprotein hormone receptor activation by the eponymous glycoprotein hormones.

In the second chapter, I and others discover and biophysically characterize single-domain antibodies directed against the SARS-CoV-2 S protein. The single-domain antibodies we initially describe only inhibit viral entry through steric occlusion of the angiotensin-converting enzyme II binding interface on the viral S protein. In a structure-guided manner, we further engineer one single-domain antibody into a multivalent construct that restricts the viral S protein into a closed conformation and show enhanced potency in viral entry inhibition after additional affinity maturation. We demonstrate that this construct is stable to common routes of aerosolized delivery and propose its use as a viral prophylactic. This work also describes single-domain antibodies with allosteric, yet incompletely described mechanisms of S protein entry inhibition.

Table of Contents

AUTOANTIBODY MIMICRY OF HORMONE ACTION AT THE THYROTROPIN RECEPTOR	1
1.1 ABSTRACT.....	3
1.2 INTRODUCTION	3
1.3 RESULTS	5
1.4 DISCUSSION	18
1.5 FIGURES	21
1.6 TABLES	43
1.7 SUPPLEMENTAL FILES.....	45
1.8 MATERIALS AND METHODS	46
1.9 ACKNOWLEDGEMENTS.....	70
1.10 AUTHOR CONTRIBUTIONS.....	70
1.11 REFERENCES	72
AN ULTRAPOTENT SYNTHETIC NANOBODY NEUTRALIZES SARS-COV-2 BY STABILIZING INACTIVE SPIKE	82
2.1 ABSTRACT.....	85
2.2 INTRODUCTION	85
2.3 RESULTS	86
2.4 DISCUSSION	93
2.5 FIGURES	95
2.6 TABLES	109
2.7 MATERIALS AND METHODS	114
2.8 ACKNOWLEDGEMENTS.....	128
2.9 AUTHOR CONTRIBUTIONS.....	128
2.10 REFERENCES	133

List of Figures

Figure 1.1 Cryo-EM structures of native human TSH and TR1402 bound to active TSHR complexed with heterotrimeric G _s	21
Figure 1.2 Activation mechanism of the TSHR revealed by inactive structure bound to inverse agonist CS-17.....	22
Figure 1.3 Activation of the TSHR by a Graves' disease autoantibody.....	23
Figure 1.4 Membrane bilayer interactions are critical for TSHR activation.	24
Figure 1.5 Model for TSHR action.	26
Figure 1.6 TSHR construct expression matching and supplemental cAMP production assays.	27
Figure 1.7 Cryo-EM data processing for TSH-bound TSHR-Gs complex.....	29
Figure 1.8 Model map correlation coefficients for cryo-EM structures.	30
Figure 1.9 Cryo-EM data processing for TR1402-bound TSHR-Gs complex.	31
Figure 1.10 Glycosylation status of native human TSH.	32
Figure 1.11 Receptor:hormone interaction comparisons across the glycoprotein hormone receptor family.	33
Figure 1.12 Cryo-EM data processing for Org 274179-0 bound TSHR.....	35
Figure 1.13 Cryo EM data processing for the CS-17 bound TSHR:Org 274179-0 complex.	36
Figure 1.14 Comparison of Org 274179-0 bound- and CS-17 bound TSHR EM maps.....	37
Figure 1.15 Comparison of TSHR activation with other GPCRs.	38
Figure 1.16 Cryo-EM data processing for M22-bound TSHR-Gs complex.....	39
Figure 1.17 Extent of ECD activation for all individual simulations across four simulation conditions.	40
Figure 1.18 Glycosylation of engineered K1-70 ^{glyco} construct.	42
Figure 2.1 Discovery of two distinct classes of anti-Spike nanobodies	95
Figure 2.2 Cryo-EM structures of Nb6 and Nb11 bound to Spike.	96
Figure 2.3 Multivalency improves nanobody affinity and inhibitory efficacy.....	97
Figure 2.4 Affinity maturation of Nb6 yields a picomolar SARS-CoV-2 neutralizing molecule.	98
Figure 2.5 mNb6 and mNb6-tri retain activity after aerosolization, lyophilization, and heat treatment.....	99
Figure 2.6 Cryo-EM workflow for Nb6.....	100
Figure 2.7 Cryo-EM workflow for Nb11.....	101
Figure 2.8 Resolution of cryo-EM maps and models.	102

Figure 2.9 Modeling of distances for multimeric nanobody design	103
Figure 2.10 Radiolytic hydroxyl radical footprinting of Spike ^{S2P}	104
Figure 2.11 Multivalent Nb3 construct inhibits Spike ^{S2P} :ACE2 interaction.	105
Figure 2.12 CryoEM workflow for mNb6.	106
Figure 2.13 mNb6 and Nb3-tri are additive for viral neutralization.	107
Figure 2.14 Stability of Nb6 and its derivatives.	108

List of Tables

Table 1.1 Cryo-electron microscopy data collection, refinement, and validation statistics	43
Table 1.2 Summary of TSHR signaling studies.....	44
Table 2.1 Nanobody affinities and efficacies in neutralization assays.	109
Table 2.2 Cryo-electron microscopy data collection, refinement, and validation statistics	111
Table 2.3. X-ray crystallography data collection and refinement statistics.....	112
Table 2.4. Nanobody expression plasmids	113

Chapter 1

Autoantibody mimicry of hormone action at the thyrotropin receptor

Contributing Authors

Bryan Faust^{1,2,3,15}, Christian B. Billesbølle^{1,15}, Carl-Mikael Suomivuori^{4,5,6,7}, Isha Singh¹, Kaihua Zhang², Nicholas Hoppe^{1,3}, Antonio F. M. Pinto⁸, Jolene K. Diedrich⁸, Yagmur Muftuoglu⁹, Mariusz W. Szkudlinski¹⁰, Alan Saghatelian¹¹, Ron O. Dror^{4,5,6,7}, Yifan Cheng^{2,3,12}, Aashish Manglik^{1,3,13,14}

1. Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA
2. Department of Biochemistry & Biophysics, University of California, San Francisco, CA, USA
3. Biophysics Graduate Program, University of California, San Francisco, CA, USA
4. Department of Computer Science, Stanford University, Stanford, CA, USA
5. Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA, USA
6. Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, USA
7. Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA
8. Mass Spectrometry Core for Proteomics and Metabolomics, Salk Institute for Biological Studies, La Jolla, CA, USA
9. Stanford University School of Medicine, Stanford, CA, USA
10. Trophogen, Inc., Rockville, MD, USA
11. Clayton Foundation Laboratory for Peptide Biology Lab, Salk Institute for Biological Studies,

La Jolla, CA, USA

12. Howard Hughes Medical Institute, University of California, San Francisco, CA, USA

13. Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA, USA

14. Chan Zuckerberg Biohub, San Francisco, CA, USA

15. These authors contributed equally: Bryan Faust, Christian B. Billesbølle

1.1 Abstract

Thyroid hormones are vital to metabolism, growth, and development¹. Thyroid hormone synthesis is controlled by thyrotropin (TSH), which acts at the thyrotropin receptor (TSHR)². Autoantibodies that activate the TSHR pathologically increase thyroid hormones in Graves' disease³. How autoantibodies mimic TSH function remains unclear. We determined cryogenic-electron microscopy structures of active and inactive TSHR. In inactive TSHR, the extracellular domain lies close to the membrane bilayer. TSH selects an upright orientation of the extracellular domain due to steric clashes between a conserved hormone glycan and the membrane bilayer. An activating autoantibody from a Graves' disease patient selects a similar upright orientation of the extracellular domain. Reorientation of the extracellular domain transduces a conformational change in the seven transmembrane domain via a conserved hinge domain, a tethered peptide agonist, and a phospholipid that binds within the seven transmembrane domain. Rotation of the TSHR extracellular domain relative to the membrane bilayer is sufficient for receptor activation, revealing a shared mechanism for other glycoprotein hormone receptors that may also extend to other G protein-coupled receptors with large extracellular domains.

1.2 Introduction

The thyroid gland regulates organ development and metabolism in all vertebrates via the thyroid hormones triiodothyronine (T_3) and thyroxine (T_4)¹. Synthesis and secretion of the thyroid hormones is controlled by a homeostatic hypothalamic-pituitary-thyroid signaling axis². Hypothalamic and pituitary sensing of low thyroid hormone levels induces secretion of the pituitary thyrotropin hormone (also called thyroid stimulating hormone, TSH), which acts at the thyrotropin receptor (TSHR), a G protein-coupled receptor (GPCR) located on thyroid follicles⁴.

Activation of heterotrimeric G_s and G_q signaling pathways downstream of the TSHR leads to thyroid hormone production⁵, closing the negative feedback loop to set physiological thyroid hormone levels.

Dysregulation of the central hypothalamic-pituitary-thyroid signaling axis leads to inappropriately increased or decreased thyroid hormone levels, causing a disease burden that affects approximately 5% of the world population⁶. Hypothyroidism predominantly stems from iodine deficiency or autoimmune inflammation of the thyroid. The primary cause of hyperthyroidism in countries without iodine deficiency is Graves' disease, an autoimmune disorder leading to inappropriate activation of the TSHR by autoantibody thyroid stimulating immunoglobulins (TSI)^{7,8}. Persistent direct activation of the TSHR by TSI overcomes the physiological hypothalamic-pituitary-thyroid negative feedback loop, leading to elevated thyroid hormones despite low levels of serum TSH³. Notwithstanding the central role of the TSHR in regulating thyroid hormone physiology, no currently approved medications directly target this receptor to treat thyroid diseases⁹. Current interventions either target thyroid hormone synthesis, or in more severe cases, destroy the thyroid gland leading to a lifelong need for thyroid hormone replacement therapy³.

A deeper understanding of how the TSHR is activated, physiologically by TSH or pathologically by TSI, would enable approaches to precisely tune the hypothalamic-pituitary-thyroid signaling axis to correct thyroid disease. A central challenge, however, has been the design of TSHR selective molecules, due in part to a limited understanding of the structure and dynamics of TSHR function. The TSHR shares significant similarity to two homologous receptors that are critically important in reproductive physiology: the follicle-stimulating hormone receptor (FSHR) and the luteinizing hormone–choriogonadotropin receptor (LH/CGR). While the FSHR is specific for the

follicle stimulating hormone (FSH), the LH/CGR is activated by both luteinizing hormone (LH) and chorionic gonadotropin (CG). A hallmark of TSH, FSH, LH and CG, together termed glycoprotein hormones, is their complex N-linked glycosylation required for biological activity¹⁰⁻¹². Structures of the FSHR and LH/CGR bound to FSH and CG, respectively, have illuminated key aspects of glycoprotein hormone recognition and receptor activation^{13,14}. However, these studies leave unanswered how glycosylation drives hormone activity, and for the TSHR, how pathogenic autoantibodies mimic TSH action. Here, we use a combination of cryogenic electron-microscopy (cryo-EM), protein engineering, signaling studies, and molecular dynamics (MD) simulations to illuminate the molecular basis of action of TSH and autoantibodies at the TSHR. Our studies provide a model for physiological and pathological activation of the TSHR and a general activation mechanism for the glycoprotein hormone receptor family.

1.3 Results

Structures of hormone-bound activate TSHR

We first obtained a structure of activated TSHR bound to native human TSH (**Figure 1.1**). The TSHR is naturally proteolytically cleaved within the extracellular domain leading to removal of residues 317 to 366, a region termed the “C-peptide”¹⁵. While the physiological relevance of C-peptide excision remains unclear, previous biochemical studies have demonstrated that cleavage in this domain leads to lower expression levels of intact receptor¹⁶. To simplify purification of intact receptors for structural studies, we generated a TSHR construct devoid of the C-peptide; removal of the C-peptide did not affect TSH or TSI activation of G_s signaling (**Figure 1.6**). To further improve expression and purification of active TSHR, we generated a construct with a C-terminal fusion of an engineered “mini” G_s protein. This miniG_s protein contains only the Ras-

like GTPase domain of G_{α_s} and is thermostabilized to interact with a GPCR in a nucleotide-independent manner¹⁷. We purified this construct, termed TSHR-mini G_s , in complex with native TSH purified from human pituitary, recombinantly expressed $G\beta_1\gamma_2$, and the G_s stabilizing nanobody 35 (Nb35)¹⁸. The resulting preparation of TSH-activated, G_s -bound TSHR was stable enough to enable single particle cryo-EM. Because of conformational heterogeneity caused by flexibility between the TSHR extracellular domain (ECD) and the seven transmembrane domain (7TM) bound to G_s , we separately classified and refined these two regions of the complex to yield a 3.4 Å resolution map of the TSHR ECD bound to TSH and a 2.9 Å map of the TSHR 7TM bound to the G_s heterotrimer (**Figure 1.7, Table 1.1**). While these reconstructions enabled us to build a model for the TSHR 7TM domain and the G_s heterotrimer, key interacting regions between TSH and TSHR were of poor resolution (>4 Å) (**Figure 1.7, Figure 1.8**).

To gain higher resolution insights into how TSH interacts with the TSHR ECD, we determined a cryo-EM structure of active TSHR bound to TR1402, a high affinity TSH “superagonist”^{19,20}. Glycoprotein hormones are composed of a common glycoprotein α chain (GPH α) and hormone-specific β chains. While TSH and TR1402 share an identical TSH β chain, four arginine substitutions in the TR1402 GPH α lead to ~55-fold improved potency for G_s activation compared to native TSH (**Figure 1.6**). Likely due to this improved affinity, our reconstruction of TR1402-bound TSHR-mini G_s was resolved to 2.7 Å for the TSHR ECD bound to TR1402 and 2.4 Å for the 7TM bound to G_s (**Figure 1.1, Figure 1.9**). 3D variability analysis of the TR1402 reconstruction further supported flexibility between the TSHR ECD and 7TM domain (**Movie 1.1**). Notably, the root mean squared deviation (RMSD) for the TSHR between TR1402 and TSH-bound states is 1.26 Å, suggesting that these ligands stabilize highly similar receptor conformations.

Our reconstructions of both TSH and TR1402 resolved key glycosylation sites on the hormone, including at Asn52 of the GP α chain previously demonstrated to be required for glycoprotein hormone signaling^{10-12,21} (**Figure 1.1, Figure 1.10**). TSH and TR1402 bind to the concave surface of the TSHR ECD leucine rich repeat (LRR) in a similar orientation as follicle stimulating hormone (FSH) bound to FSHR and chorionic gonadotropin (CG) bound to LH/CGR^{13,14,22} (**Figure 1.11**). Our TR1402 reconstruction revealed that the hinge region of TSHR, which connects the TSHR ECD and the 7TM domain, is ordered as two disulfide-linked α -helices that make extensive contact with both hormone chains (**Figure 1.1**). One of these TSHR α -helices is coordinated by ionic interactions in both the TR1402 α and TSH β chains while the other interacts with the α -L1 and α -L3 loops of the TR1402 α -chain, likely via hydrophobic interactions (**Figure 1.1, Figure 1.11**). The TSHR hinge region positions Y385, a residue previously identified as important in TSH-mediated signaling²³, into a hydrophobic pocket at the hormone α/β chain interface (**Figure 1.1**). Sulfation of Y385 in TSHR, and a homologous tyrosine in FSHR and LH/CGR, has been proposed to be critical for glycoprotein hormone action¹³. However, our structure suggests that a sulfate group attached to Y385 would not be compatible with TSHR binding to the hormone. Consistent with this structural prediction and contrary to previous reports^{23,24}, we found that Y385 mutations in the TSHR that preclude sulfation (Y385F and Y385A) retained similar TSH potencies and efficacies as wild-type TSHR (**Figure 1.6**).

Selective binding of glycoprotein hormones to their cognate receptors is determined by the hormone-specific β chain^{25,26}, and more specifically, has been localized to a 16-residue segment between the 10th and 12th cysteine in the β chain termed the “seatbelt” loop^{27,28,29,30}. A key concept

from prior studies examining glycoprotein hormone selectivity is the presence of negative determinants that preclude binding of incorrectly paired hormones and receptors³¹. Structural comparison of the seatbelt loop between the three different glycoprotein hormones bound to their respective receptors provides insight into such negative determinants (**Figure 1.1**). We define two distinct regions of the seatbelt loop based on a conserved aspartate residue in the hormone β -chains known to be critical for receptor interaction (D94 in TSH, D99 in CG, D93 in FSH): an N-terminal region (region I) and a C-terminal region (region II). The net charge for region I is negative for TSH β /FSH β and positive for CG β (**Figure 1.1**)³². In TSH, D91 in region I contacts TSHR residue K209. For CG bound to LH/CGR, this ionic interaction is swapped, with CG β R95 contacting LH/CGR E206 (**Figure 1.1**). These swapped ionic interactions are important in determining faithful hormone-receptor pairings as underscored by previous studies that identified enhanced hCG signaling at the K209E TSHR mutant³³. Comparison of region II between TSH and FSH reveals another set of opposing ionic interactions likely important for hormone selectivity (**Figure 1.1**). In this region, TSH β presents a negatively charged E98 to contact K58 in the TSHR. By contrast, FSH β presents a positively charged R97 to contact E76 in FSHR (**Figure 1.1**). While it is likely that additional interactions are responsible for encoding hormone selectivity, these opposing ionic interactions provide a structural rationale for negative determinants that limit incorrect hormone-receptor pairing.

Our reconstruction of TR1402-bound TSHR revealed density for an endogenous phospholipid buried within the 7TM domain (**Figure 1.1**). A combination of mass spectrometry experiments to identify potential lipids enriched in our TSHR-G_s samples and potential candidates compatible for modeling within the observed cryo-EM density led us to tentatively assign this density as

dipalmitoylphosphatidylcholine (DPPC, **Figure 1.11**). To probe whether lipid binding at this site is important for TSHR activation, we generated cell-surface expression-matched TSHR cell lines with two mutations predicted to inhibit DPPC occupancy in the 7TM domain (A644K and A647K). Both mutations significantly decreased the efficacy and potency of TSH in cAMP production (**Figure 1.1**), suggesting that the presence of a lipid within the TSHR transmembrane domain is important for propagating ECD binding signals to the 7TM domain. Our hormone-bound structures of the TSHR thus offer insight into the molecular determinants of receptor-hormone interactions, hormone selectivity, and the presence of a lipidic conduit important for receptor activation.

Conformational changes on TSHR activation

To understand how TSH activates the TSHR, we next determined an inactive-state structure of the receptor (**Figure 1.2**). Our attempts to capture the inactive state of the TSHR stabilized by the small molecule negative allosteric modulator Org 274179-0³⁴ resulted in a low resolution reconstruction resolving the TSHR ECD and a detergent micelle (**Figure 1.12**). We speculated that this result may be due to orientational flexibility of the TSHR ECD compared to the 7TM domain associated with constitutive activity of the TSHR³⁵. To further stabilize an inactive conformation, we used a Fab fragment of the inverse agonist CS-17 antibody, which has been previously reported to suppress constitutive activity of the TSHR³⁶. We obtained a reconstruction of the TSHR–Org 274179-0–CS-17 complex to a global resolution of 3.1 Å, with higher resolution features for the TSHR ECD and CS-17 Fab in comparison to the lower resolution reconstruction of the TSHR 7TM domain (**Figure 1.2**, **Figure 1.13**). While this map enabled modeling into the TSHR ECD and CS-17 Fab densities, the resolution in the TSHR transmembrane region limited our ability to model side chains accurately or to define the binding site for Org 274179-0.

In the inactive state, the TSHR ECD rotates 55° towards the 7TM domain through an axis bisecting these two domains (**Figure 1.2**). We term this the “down” state. By contrast, we label the ECD in the “up” state as observed for active TSHR bound to TSH. Importantly, the internal leucine-rich-repeat structure of the ECD does not change between the inactive and active states of the TSHR (RMSD of 0.64 Å). The ECD therefore moves as a rigid body between the inactive and active states. The ECD orientation is similar between the low-resolution reconstruction of unbound, inactive TSHR and CS-17-bound, inactive TSHR (**Figure 1.14**), suggesting that CS-17 likely stabilizes an inactive, ECD down orientation to achieve its inverse agonist efficacy. To test whether rigid body motion of the TSHR ECD is sufficient to activate the receptor in the absence of other stimuli, we designed two mutations in TSHR (K262C and N483C) that would form a disulfide bond to trap the ECD in the up orientation (**Figure 1.2**). The K262C-N483C double mutant, but not the single cysteine substitutions, exhibited a 16-fold increase in constitutive activity compared to expression-matched wild-type TSHR (**Figure 1.2, Figure 1.6**). Incubating cells with the disulfide reductant tris(2-carboxyethyl)phosphine (TCEP) markedly decreased constitutive activity of the K262C-N483C mutant and restored TSH-stimulated cAMP production. Rotation of the TSHR ECD to the up state is therefore sufficient to activate the receptor, even in the absence of extensive interactions between TSH and TSHR.

Comparison of these new TSHR structures with previously determined structures of LH/CGR showed a similar relative orientation of the ECD in the inactive and active orientations (**Figure 1.15**). Given these similarities, we surmise the activation-associated rigid-body motion of the ECD is a common feature of active-state glycoprotein hormone receptors. Additionally, the TSHR 7TM

domain displays several classic hallmarks of GPCR activation. Compared to the inactive state, TM6 of active TSHR is displaced outward by 14 Å (measured from C α carbon of D617) to accommodate the α 5 helix of miniG α_s and TM7 moves \sim 4 Å inward relative to the transmembrane core of the receptor (measured from C α carbon of Y678) (**Figure 1.2**). These movements are similar to both the prototypical G $_s$ -coupled family A receptor, the β_2 -adrenoceptor¹⁸, and to the LH/CGR¹⁴ (**Figure 1.15**).

A central question is how changes in the TSHR ECD orientation connect to activation of the TSHR 7TM domain. The C-terminal end of the TSHR ECD comprises two key regions of interaction with the 7TM domain - the hinge helix and the p10 peptide (**Figure 1.2**). The hinge helix at the base of the TSHR ECD contains two disulfide bonds that connect the last LRR β -sheet with the p10 peptide, which is a conserved 10-amino acid region connecting the ECD to transmembrane helix 1 (TM1). The p10 peptide has been proposed to be a tethered intramolecular agonist for glycoprotein hormone receptors^{37,38,39}. Indeed, previous studies have found that mutation of residues within the p10 peptide are poorly tolerated in all glycoprotein hormone receptors⁴⁰. We therefore aimed to understand how rotation of the ECD is coupled to conformational changes in the p10 peptide, and how this translates to activation of the TSHR 7TM. Comparison of active and inactive TSHR structures revealed that ECD transition into the up state results in rotation of the hinge helix, which makes extensive contacts with ECL1 in both active and inactive TSHR. Rotation of the hinge helix lifts the N-terminal portion of the p10 peptide \sim 5 Å away from the 7TM core (measured at F409 C α) (**Figure 1.2**). Upward movement of the p10 peptide produces a void into which the extracellular tip of TM7 shifts inward by \sim 4 Å (measured at N658 C α) (**Figure 1.2**). In active state structures, we observed a well resolved interaction between TSHR TM7

residue K660 and E409 in the p10 peptide (**Figure 1.2**). Although, we do not resolve the E409 side chain in inactive TSHR, the peptide backbone for this residue is shifted 3 Å away from the active state, suggesting that the E409-K660 interaction stabilizes the fully active state of the TSHR. Our structures therefore reveal that TSHR activating changes in p10 peptide couple to the 7TM domain via TM7.

We tested the importance of these interactions with structure-guided mutations. To test the relevance of the hinge helix-ECL1 interaction, we designed a substitution in ECL1 (I486F) that has been annotated as a driving mutation in thyroid follicular adenoma⁴¹. Our structures of the TSHR show that receptor activation is associated with a 6 Å movement of ECD residue Y279 relative to I486; replacing isoleucine with the bulkier phenylalanine would therefore be predicted to perturb TSHR activation (**Figure 1.2**). In signaling studies, I486F is constitutively active but has diminished sensitivity to TSH (**Figure 1.2**). A previous study demonstrated that the I486F mutant is not constitutively active if the TSHR ECD is truncated⁴², further highlighting that the contact between the ECD and ECL1 is a critical conduit for propagation of the ECD orientation to the 7TM domain. Although mutagenesis of the majority of the p10 peptide has been previously shown to severely compromise glycoprotein hormone receptor expression and sensitivity to glycoprotein hormones⁴⁰, we tested the importance of p10 peptide-TM7 coupling by disrupting the E409-K660 interaction. We find that the E409A TSHR mutant shows diminished constitutive activity and reduced TSH potency and efficacy, supporting the importance of p10 peptide-TM7 contacts in TSHR activation (**Figure 1.2**). Our structural observations and mutagenesis studies, combined with extensive prior mutagenesis studies for the TSHR, therefore demonstrate that

rotation of the ECD is critical for receptor activation and highlight a conduit for this conformational change to impinge on the 7TM domain.

Agonist M22 autoantibody mimics TSH to activate the TSHR

We next aimed to understand how agonistic autoantibodies activate the TSHR. Previous efforts to characterize thyroid stimulating immunoglobulins (TSI) identified M22, a monoclonal antibody isolated from a patient with Graves' disease that potently activates the TSHR⁴³. A prior X-ray crystal structure of M22 bound to the extracellular domain of the TSHR provided evidence of an epitope overlapping with the predicted binding site for TSH⁴⁴. However, the mechanism for M22 mimicry of TSH action remains poorly understood^{45,46}.

We prepared a complex of active TSHR-miniG_s bound to the Fab fragment of the M22 autoantibody and analyzed the structure and dynamics of this complex with cryo-EM. In 2D class averages, we identified two distinct orientations of the TSHR ECD bound to the M22 Fab (**Figure 1.3**). One class appeared consistent with a TSHR ECD orientation in the up state similar to that observed for TSH. In the other class, the TSH ECD–M22 Fab complex is distinct from either the up or down orientations of the TSHR ECD. Instead, the TSH ECD appears to embed within the side of the detergent micelle in a region that would normally occupy the lipid bilayer. Although this “side” orientation of the TSHR ECD is not physiological and likely results from detergent-solubilization of the TSHR, it suggests that the M22-bound TSHR ECD is more orientationally heterogeneous than when bound to glycosylated TSH. We were unable to determine a high-resolution reconstruction of this “side” complex. However, we successfully reconstructed two maps for the up state: the TSHR ECD bound to M22 Fab at 3.0 Å and the TSHR 7TM domain

bound to G_s at 2.8 Å (**Figure 1.3, Figure 1.16**). In agreement with our TR1402-bound TSHR-G_s reconstruction, we also resolved a well-defined two-tailed lipid density in the transmembrane core of M22-bound TSHR, suggesting that this lipid is a common feature of activated TSHR (**Figure 1.3**).

We next compared TSH- and M22-activated TSHR (**Figure 1.3**). With M22 bound, the ECD up state is in an orientation that is highly similar to the TSH-activated state, with a small 5° rotation when the 7TM domains are aligned. The conformation of the 7TM domain and orientation of G_s heterotrimer are also highly similar between M22- and TSH-activated TSHR, with a RMSD of 0.87 Å. A key distinction between TSH-, TR1402- and M22-activated TSHR is that the hinge region is unresolved in M22-activated TSHR, which is consistent with prior studies showing that the hinge region is dispensable for activation of the TSHR²⁴. The lack of a direct interaction with the hinge region is potentially responsible for the increased orientational flexibility of the TSHR ECD when bound to M22 as compared to TSH or TR1402. The structure of M22- and TSH-activated TSHR therefore demonstrate remarkable similarity in ECD orientation.

Membrane bilayer is critical for TSHR activation

Our structures of the TSHR revealed that rotation of the ECD is coupled to 7TM domain activation by direct interactions between the p10 peptide and TM7. However, this does not address how binding of TSH and activating TSI lead to the up ECD orientation. From structures of CG-bound LH/CGR, it has been proposed that a clash between the CG-β chain and the membrane bilayer drives receptor activation by “pushing” the LH/CGR ECD away from the membrane¹⁴. Simultaneously, an additional interaction between the hinge region and the common GPH-α chain

“pulls” the ECD to the up state. While our structures of inactive and active TSHR revealed overall receptor conformation and ECD orientation changes similar to LH/CGR, we arrive at a distinct model for glycoprotein hormone action based on our structures of TSH, the M22 antibody agonist, and the CS-17 antibody inverse agonist.

To establish our model, we first used a computational approach to orient the active and inactive state TSHR structures within a membrane bilayer and to define hypothetical membrane bilayer boundaries⁴⁷ (**Figure 1.4**). While this static modeling approach is limited in completely describing the significant receptor and membrane dynamics important in TSHR function, it provided a first order approximation to understand how TSH and antibodies exert their efficacy. Using this approach, we first modeled the binding poses of the agonist M22 and inverse-agonist CS-17 Fab fragments onto either the active (up) or inactive (down) ECD states. Modeling the M22 Fab bound to the inactive, down orientation of TSHR ECD revealed significant expected clashes between the antibody light chain and both the receptor 7TM domain and the outer membrane plane (**Figure 1.4**). These steric clashes suggest that M22 binding to TSHR is incompatible with the inactive TSHR. Conversely, modeling the CS-17 Fab bound to the active, up orientation of the TSHR ECD revealed clashes between the antibody constant domains and the outer membrane plane (**Figure 1.4**), suggesting that CS-17 binding is incompatible with the active state. Even with the limitations inherent in this modeling approach, we inferred that M22 and CS-17 exert their efficacy at TSHR by preventing the transition of ECD orientations between active and inactive states.

To determine whether M22 and CS-17 constrain the TSHR ECD orientation in a more realistic membrane environment, we performed all-atom molecular dynamics (MD) simulations of TSHR

using both the M22-bound active structure and the CS-17-bound inactive structure as starting points. For each starting structure, we performed simulations with and without the respective antibody present. To quantify ECD orientation, we used a projection metric to define a reaction coordinate between the inactive and active ECD states from our cryo-EM models (see Methods). In simulations with the antibodies removed, the TSHR ECD fluctuates between active and inactive orientations, irrespective of the starting structure (**Figure 1.4, Figure 1.17**). In contrast, in simulations of TSHR bound to M22 or CS-17, the antibodies largely constrain the TSHR ECD to its initial state. With M22, the TSHR ECD is constrained to an active, up orientation. Conversely, CS-17 constrains the TSHR ECD to an inactive, down orientation. These simulations therefore support our static modeling, and indicate that M22 and CS-17 constrain an otherwise dynamic TSHR ECD to exert their efficacy.

We next turned to understanding how TSH activates the TSHR. We used our static modeling approach to examine how TSH would interact with the membrane bilayer in inactive TSHR with the ECD in the inactive orientation. Unlike a prior model proposed for CG activation at LH/CGR, the TSH β chain is not predicted to clash with the membrane bilayer in inactive TSHR. Instead, glycosylation at Asn52 in the GPH α chain directly points orthogonal to the membrane bilayer (Fig. 4f). We propose that in the inactive state, the down orientation of the TSHR ECD would be unable to bind TSH due to clashes between the Asn52 glycosylation and the membrane bilayer. Two lines of evidence support this proposal. 1) Our cryo-EM reconstruction of native human TSH resolves two monosaccharide moieties (**Figure 1.10**). At this site, a single additional glycan monosaccharide would be expected to induce a membrane clash analogous to those observed for M22 and CS-17. Typical biantennary N-linked glycans are composed of 5-12 monosaccharide

units with an average size of ~ 9 Å per monosaccharide⁴⁸. Mass spectrometry of native human TSH used to solve the cryo-EM structure identified Asn52 glycan chains with up to 16 monosaccharide units (**Figure 1.10**). The size of the glycosylation at Asn52 is therefore sufficiently large to be incompatible with TSH binding to inactive TSHR. 2) Prior studies have demonstrated that glycosylation of TSH and CG is necessary for TSHR and LH/CGR activation, respectively. Deglycosylated TSH or CG bind to their respective receptors but function as competitive antagonists for the native hormone^{11,49,50,51}. Taken together, these prior data and our modeling demonstrate that the conserved GPH α glycosylation at Asn52 drives TSHR activation by inhibiting the transition of the TSHR ECD to the inactive, down state.

To directly test whether glycosylation of an ECD-bound ligand is sufficient to promote full agonism at TSHR (in a manner analogous to the proposed model for TSH-mediated activation), we designed a gain-of-function experiment. Here we used an Fab fragment from the K1-70 antibody, which has previously been characterized as a TSHR neutral antagonist⁴⁵. Contrary to previous reports⁴⁵, we find that K1-70 is a weak partial agonist in a cAMP signaling assay (**Figure 1.4**). We aligned the previously solved X-ray crystal structure of K1-70 bound to the TSHR ECD with the inactive state of the TSHR from our cryo-EM structure (**Figure 1.4**). Notably, the TSHR ECD is highly similar in the K1-70 crystal structure and our inactive-state cryoEM structure (RMSD 0.62 Å). This modeling suggests that, unlike M22, K1-70 is not predicted to clash with the membrane bilayer (**Figure 1.4**). The low efficacy agonism of K1-70 in the absence of a clear membrane interaction highlights limitations of our static modeling approach: it likely does not capture the full complement of TSHR ECD orientations or may not reflect subtle conformational changes associated with K1-70 efficacy. Using the K1-70 Fab as a starting point, we engineered a

new Fab construct, K1-70^{glyco}, with a predicted glycosylation motif in a loop that is the closest contact between K1-70 and the membrane bilayer (**Figure 1.4**). Our prediction was that the engineered glycan would mimic the effect of TSH glycosylation at Asn52 and clash with the membrane bilayer in the inactive down state of TSHR, thereby converting K1-70 from a weak partial agonist into a full agonist. We isolated a glycosylated preparation of K1-70^{glyco} by Concanavalin A affinity chromatography and confirmed its glycosylation status by native mass spectrometry (**Figure 1.18**). Consistent with our model, K1-70^{glyco} is significantly more efficacious at activating cAMP signaling at TSHR compared to either WT K1-70, a non-glycosylated control K1-70^{glyco}(N16Q), (**Figure 1.4**) or a K1-70 construct with glycosylation introduced at another position distant from the membrane bilayer (**Figure 1.18**). With these modeling exercises, MD simulations, and gain-of-function experiments, we conclude that steric interactions between ECD-interacting ligands and the membrane bilayer is a critical feature of TSHR activation.

1.4 Discussion

We propose the following model for TSHR activation (**Figure 1.5**). In the unliganded state, the TSHR ECD is structurally dynamic and can transition between the up and down orientations. Transient excursions to the ECD up state lead to basal signaling. TSH selects the ECD up state due to clashes between the GPH α Asn52 glycan and the membrane bilayer if the TSH ECD converts to the down state. Activation of the 7TM domain proceeds, in part, via conformational rearrangement of the p10 peptide, which is coupled to inward motion of TM7. An endogenous phospholipid also participates as a conduit for propagating activating signals from the ECD into the 7TM domain. While interactions between the hinge region and TSH are required for potent

hormone binding and signaling, it is not required for activation *per se*, as TSHR can be activated by the M22 autoantibody without similar contacts. Finally, orientational selection of the ECD down state by CS-17 leads to inverse agonist activity. The orientation of the ECD relative to the membrane bilayer is therefore a critical determinant of TSHR activity.

Our studies revise current models of glycoprotein hormone receptor activation. We demonstrate that sulfation of Tyr385 is not necessary for hormone-mediated receptor activation. Based on our structures of TSH- and TR1402-bound TSHR and associated modeling, we propose a structural rationale for the importance of GPH α Asn52 glycosylation. Indeed, glycosylation at the right position is sufficient to fully activate the TSHR as illustrated by our experiments with K1-70, underscoring the importance of this post-translational modification in glycoprotein hormone action. Our structural studies with agonistic and inverse agonistic antibodies demonstrate that selection of ECD orientations is sufficient to activate or deactivate the receptor. Indeed the lower constitutive activity of both LH/CGR and FSHR^{52,53} may be explained by decreased ECD dynamics in these receptors relative to TSHR. Finally, structures with TR1402 and M22 revealed a phospholipid embedded within the transmembrane domain of TSHR in a region that is overlapping with an inhibitory small molecule for LH/CGR. We therefore predict that a similar phospholipid may be a common feature that underlies glycoprotein hormone receptor activation.

The structural basis of how autoantibodies mimic the TSH hormone to activate the TSHR provides insight into the molecular pathophysiology of Graves' disease, and more broadly highlights the unique way in which breaking of immune tolerance to self-antigens leads to aberrant GPCR signaling. For Graves' disease in particular, our structure of the TSHR bound to M22 suggests that the progression of thyroid autoimmune diseases is largely dependent on the geometric binding

behavior of the TSHR ECD interacting antibodies relative to the membrane bilayer. This provides some context for the presence of both TSH blocking and agonistic autoantibodies in patients with autoimmune thyroid disease - although these antibodies likely share similar epitopes on the TSHR ECD, it is their orientation relative to the membrane bilayer that determines their efficacy. More broadly, our studies provide a first structural perspective on how autoantibodies pathologically activate GPCRs, of which there are now dozens of examples in the literature⁵⁴. While the specific mechanisms for such pathological antibodies acting at various GPCRs are likely unique, a shared feature is likely recognition of GPCR extracellular domains and selection of distinct ECD orientations to mimic hormone function.

Our observations for the TSHR may also have broader implications for GPCR activation by other receptors containing large leucine-rich repeat extracellular domains such as LGR4/5/6 and the relaxin-family receptors. As the LRR domains in these receptors present static interfaces for ligand binding, we suspect that the rigid-body motions and membrane-dependent interactions described for the agonist-bound TSHR ECD are shared amongst these receptors in the transition to activated orientations. Together, our work establishes a key structural role for hormone glycosylation in receptor activation, introduces additional lipid-dependent factors for consideration in the development of small-molecules targeting the TSHR, and illuminates the mystery of how both stimulatory and TSH-blocking autoantibodies can exist within a single patient with autoimmune thyroid diseases. Our structural insights therefore set the foundation for discovery of new therapies targeting the TSHR for diseases of thyroid homeostasis.

1.5 Figures

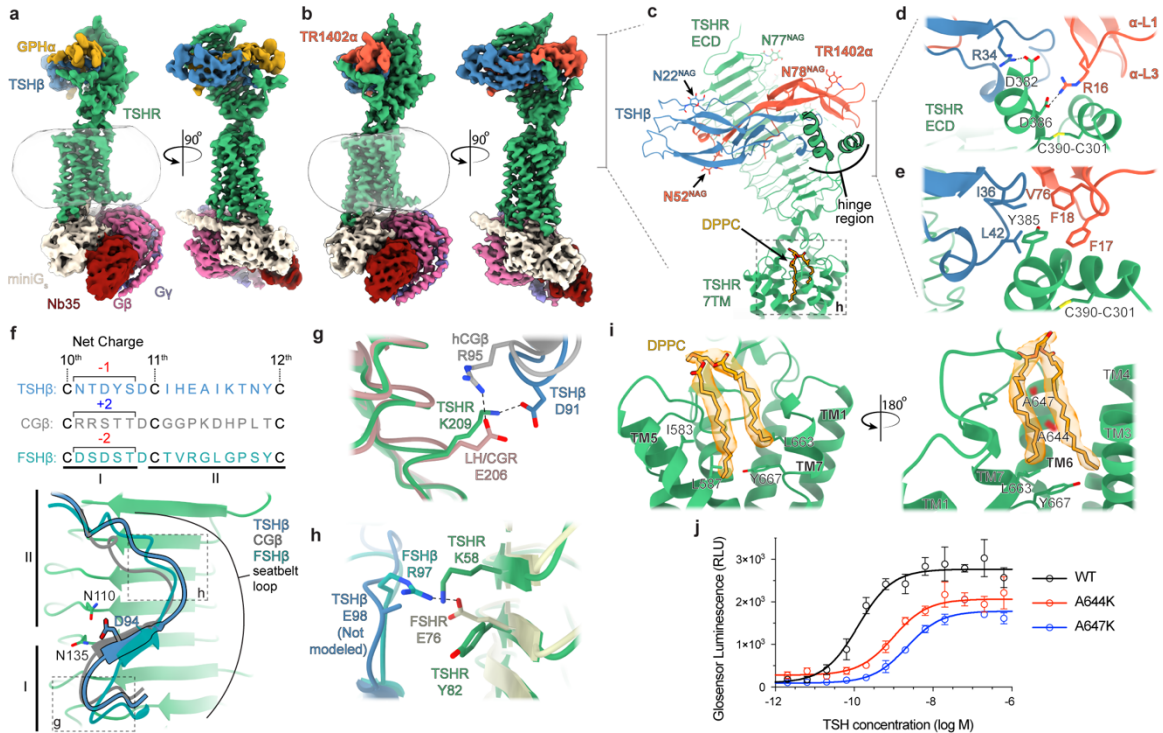


Figure 1.1 Cryo-EM structures of native human TSH and TR1402 bound to active TSHR complexed with heterotrimeric G_s

Cryo-EM maps of TSHR–G_s–Nb35 complex bound to TSH (a) and TR1402 superagonist (b). TR1402 and TSH bind to the TSHR extracellular domain (ECD). c) Model of TR1402 bound to TSHR. Resolved N-linked glycans are highlighted for both TR1402 and TSHR (N-acetylglucosamine, NAG). The disulfide-linked α -helices of the TSHR hinge are outlined in black d) TSH β R34 and TR1402 α R16 coordinate D382 and D386, respectively, in the TSHR hinge region, which positions Y385 into a hydrophobic pocket at the interface of hormone α and β chains (e). f) Sequences of TSH, FSH and hCG seatbelt loop regions between the 10th and 12th cysteine residues in the hormone-specific β chains. The seatbelt loop is further divided into regions I and II by a conserved aspartate (D94 in TSHR). Net charges in region I differ between the glycoprotein hormones. Region II of the seatbelt loop is conformationally divergent among the glycoprotein hormones. Close-up views of selectivity determinants in region I (g) and region II (h). i) Cryo-EM density for a lipid (dipalmitoylphosphatidylcholine, DPPC) in the TSHR transmembrane pocket. A644 and A647 side chains (highlighted in red) line the lipid binding pocket. j) cAMP production assay for mutations in the lipid binding site. Plotted data points are means of triplicate measurements \pm SD from a representative experiment of n=3 biological replicates.

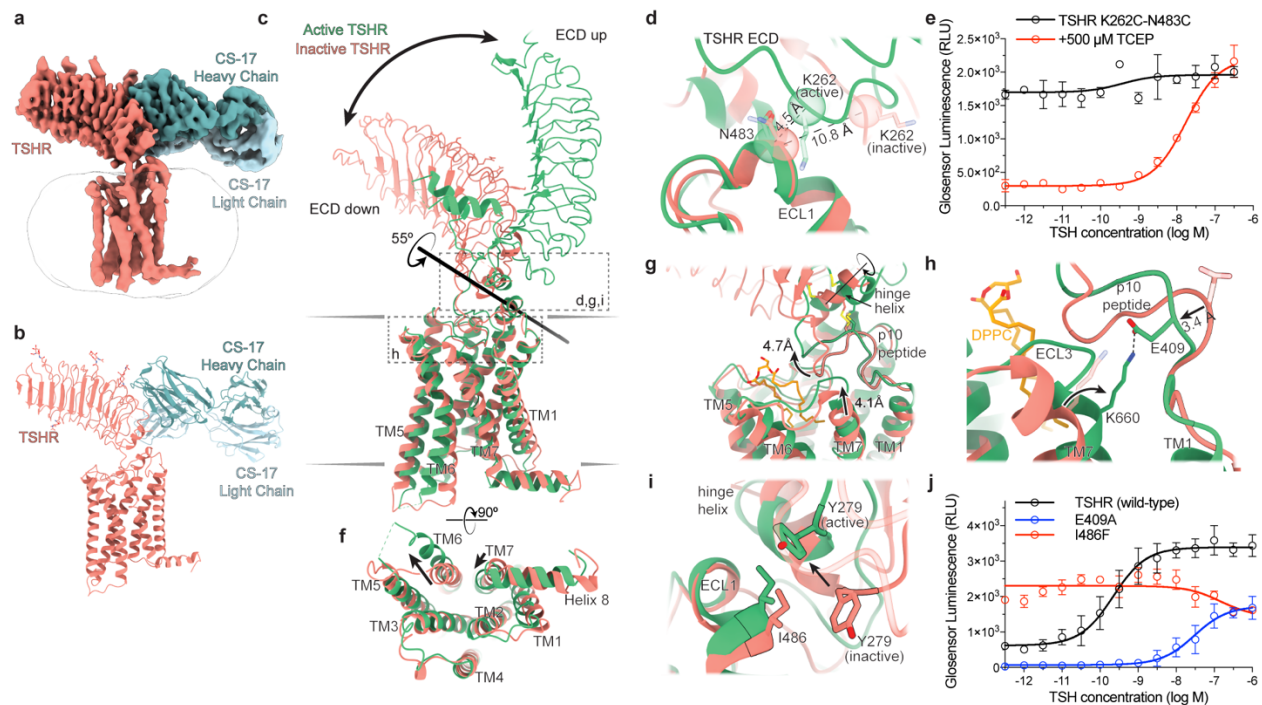


Figure 1.2 Activation mechanism of the TSHR revealed by inactive structure bound to inverse agonist CS-17.

Cryo-EM map (a) and model (b) of inactive TSHR bound to CS-17 Fab. c) Structural comparison of inactive and active TSHR with the 7TM domain aligned. In inactive TSHR, the extracellular domain (ECD) is in a down orientation close to the membrane bilayer. In active TSHR, the ECD is in an up orientation. The ECD rotates 55° along an axis as calculated by Dyndom3D⁵⁵. d) Disulfide trapping of the active TSHR ECD conformation using a K262C-N483C mutant TSHR. The C α distance between positions 262 and 483 (indicated in parentheses) would only enable a disulfide bond when the ECD is in the active, up conformation but not in the inactive, down conformation. e) The K262C-N483C TSHR mutant is more constitutively active than wild-type TSHR (see Figure 1.6). Addition of 500 μ M TCEP reduces basal activity of the K262C-N483C disulfide-locked construct. f) In the active state, TM6 of TSHR moves outward by 14 \AA , while TM7 moves inward by 4 \AA . g) Upon activation, rotation of the ECD leads to a rotation of the hinge helix, an extracellular displacement of the p10 peptide, and an inward movement of the extracellular tip of TM7. h) E409 in the p10 peptide interacts with K660 in TM7 in active TSHR. Inactive state side chains (transparent) are not resolved but peptide backbone suggests the E409-K660 interaction is not maintained. i) Y279 traverses ~ 6 \AA across the ECL1-hinge helix interface directly over I486. j) Disruption of p10-TM7 interactions (E409A) or perturbation of ECL1-hinge helix interface (I486F) affects TSH-mediated receptor activation and basal activity, respectively. Plotted data points are means of triplicate measurements \pm SD from a representative experiment of n=3 biological replicates.

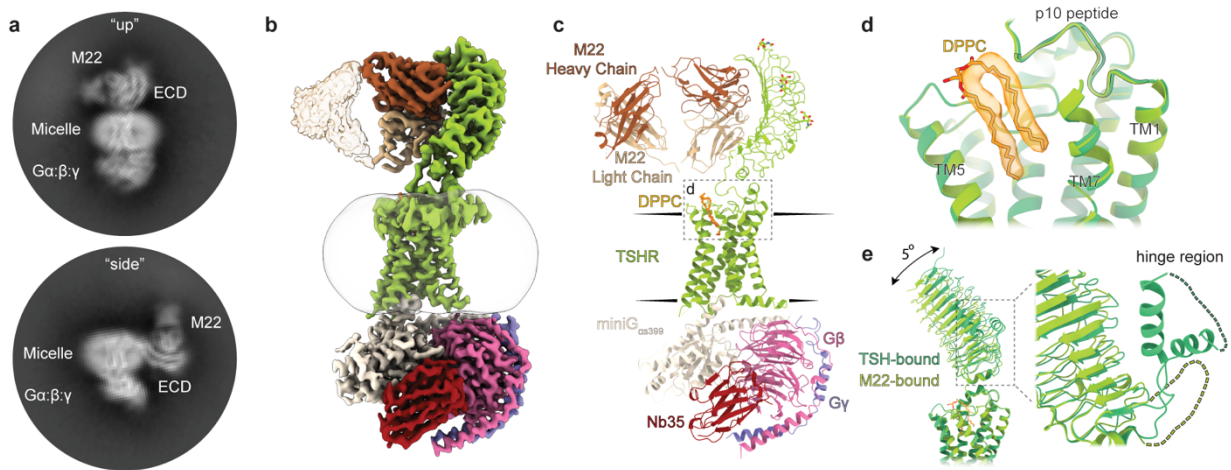


Figure 1.3 Activation of the TSHR by a Graves' disease autoantibody.

a) Selected 2D class averages of the “up” and “side” TSHR ECD orientations for TSHR bound to M22. Cryo-EM map (**b**) and model (**c**) of the M22 Fab-TSHR-G_s-Nb35 complex. **d)** Dipalmitoylphosphatidylcholine (DPPC, orange) is modeled into the density present in the TSHR transmembrane pocket. TM6 and ECL3 are hidden for clarity. **e)** Alignment of TSHR 7TM domain between TSH and M22-bound models reveals minimal (~5°) change in the orientation of the ECD. The TSHR α -helical hinge region is not resolved in the M22-bound receptor complex.

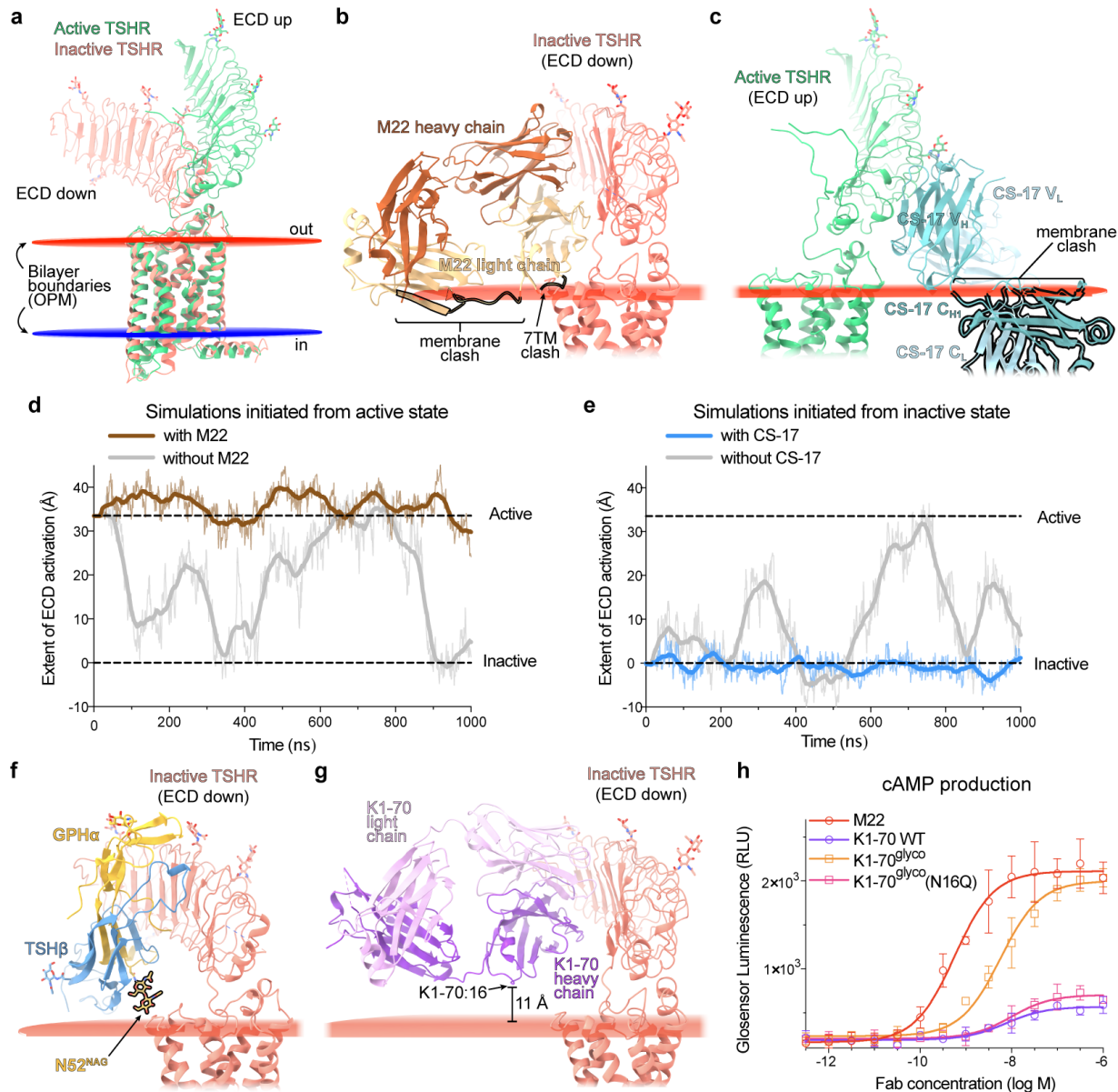


Figure 1.4 Membrane bilayer interactions are critical for TSHR activation.

a) Orientation of active and inactive TSHR in a mammalian plasma membrane bilayer as defined by the Orientations of Proteins in Membranes server. **b)** Modeling of autoantibody agonist M22 binding to inactive TSHR with the ECD in the down orientation shows expected clashes between the M22 light chain, the TSHR 7TM domain, and the outer membrane bilayer **c)** Modeling of inverse agonist CS-17 binding onto active TSHR shows expected clashes between the CS-17 constant domains and the membrane bilayer. **d, e)** In simulations with no antibodies present (grey traces), the TSHR ECD fluctuates between ECD up (active) and ECD down (inactive) orientations, regardless of the starting structure. M22 constrains the ECD to an ECD up orientation (brown trace), and CS-17 constrains the ECD to a down (inactive) orientation (blue trace). ECD orientation is quantified using a projection metric as outlined in Methods. Thick traces represent smoothed values with an averaging window of 25 ns; thin traces represent unsmoothed values. Black dashed horizontal lines represent ECD conformations in cryo-EM structures. **f)** Modeling of TSH binding

to inactive TSHR with the ECD in the down orientation. The membrane-proximal Asn52 glycan is highlighted. **g**) Modeling of K1-70 to inactive state TSHR. K1-70 is compatible with binding to inactive state TSHR. The location of an engineered membrane-proximal glycosylation site (K1-70: Q16N), and distance to the OPM-determined static membrane plane is indicated. **h**) An engineered version of K1-70 with N-linked glycosylation at residue 16 (K1-70^{glyco}) is more potent and efficacious at cAMP production than K1-70. Plotted data points are means of triplicate measurements \pm SD from a representative experiment of n=3 biological replicates.

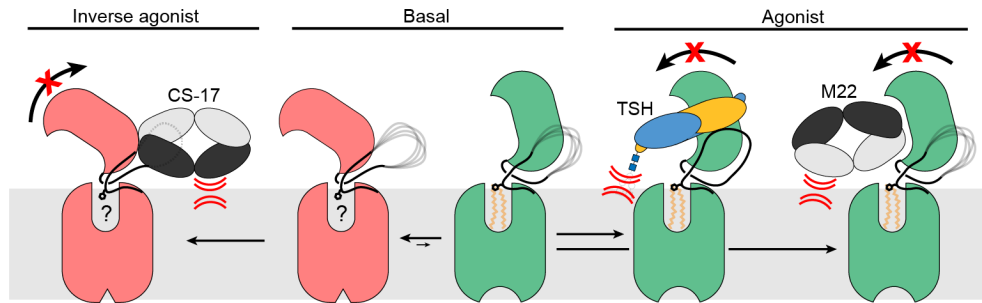


Figure 1.5 Model for TSHR action.

In the basal state, the TSHR ECD can spontaneously transition to the up state, leading to constitutive activity. TSH stabilizes an upright ECD because steric clashes between the $GPH\alpha$ N52 glycan and the membrane bilayer prevent conversion of the ECD to the down state. Agonistic autoantibodies like M22 activate TSHR in a similar manner by preventing the ECD down state. Conversely, inverse agonistic antibodies like CS-17 prevent the ECD from assuming the up state, thereby locking TSHR in an inactive orientation.

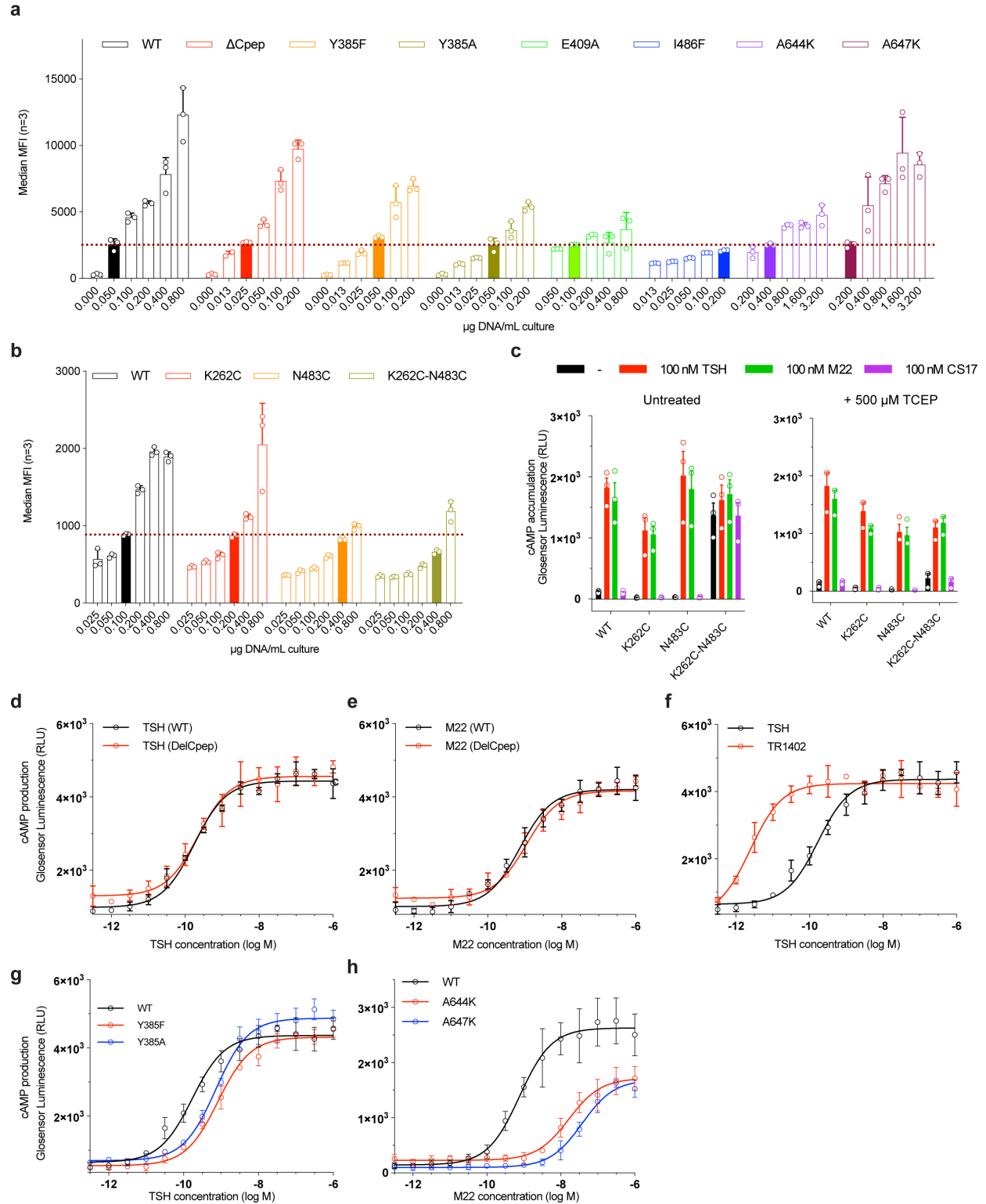


Figure 1.6 TSHR construct expression matching and supplemental cAMP production assays.

a, b) DNA titrations of TSHR mutants tested in cAMP production assays. Median fluorescence intensity \pm SD ($n=3$) of anti-FLAG-A647 staining is plotted. Filled in bars represent the DNA

concentration used for co-transfection experiments with pGlo cAMP biosensor plasmid. Horizontal red line indicates the mean wild-type (WT) TSHR fluorescence intensity. Flow cytometry experiments were performed in triplicate, with identical gating across all cell lines tested. **c**) cAMP production of TSHR cysteine mutants comparing the basal level (-) to 100 nM of TSH, M22, or CS-17 for untreated cells (left panel) and cells treated with 500 μ M TCEP (right panel). Data points are means \pm SD from three biological replicates. **d, e**) cAMP production curves comparing TSH and M22-mediated activation of WT and C-peptide deleted TSHR constructs. **f**) cAMP production curves for TSH and TR1402-mediated WT TSHR activation. **g**) cAMP production curves for TSH-mediated WT and TSHR mutant cell line (Y385F, Y385A) activation. **h**) cAMP production curves for M22-mediated WT and lipid-displacing TSHR mutant cell line (A644K, A647K) activation. Plotted data points in panels c-g are means of triplicate measurements \pm SD from a representative experiment of n=3 biological replicates.

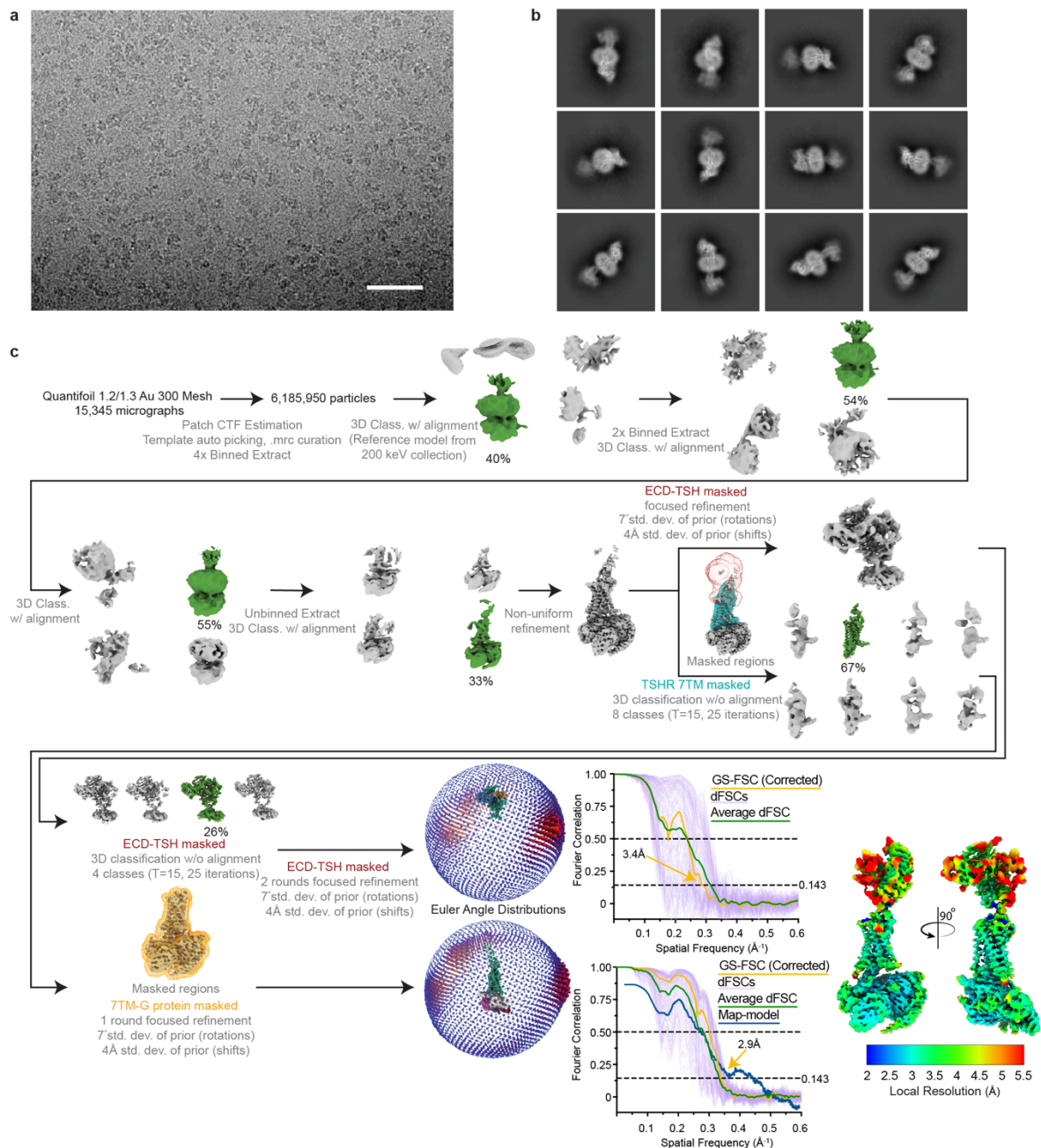


Figure 1.7 Cryo-EM data processing for TSH-bound TSHR-G_s complex.

a) Representative image from 15,345 micrographs. Scale bar, 50 nm. **b)** Selected 2D class averages. **c)** Processing approach used for reconstruction of TSH-bound TSHR-G_s complex. A local resolution map was calculated from cryoSPARC using masks from indicated local refinement, then visualized with the composite map in the same scale. A viewing distribution plot was generated using scripts from the pyEM software suite⁸³ and visualized in ChimeraX. GS-FSC and Directional FSC (dFSC, shown as purple lines) curves were generated in cryoSPARC and as previously described in Dang, S. *et al. Nature* 552, 426-429 (2017).

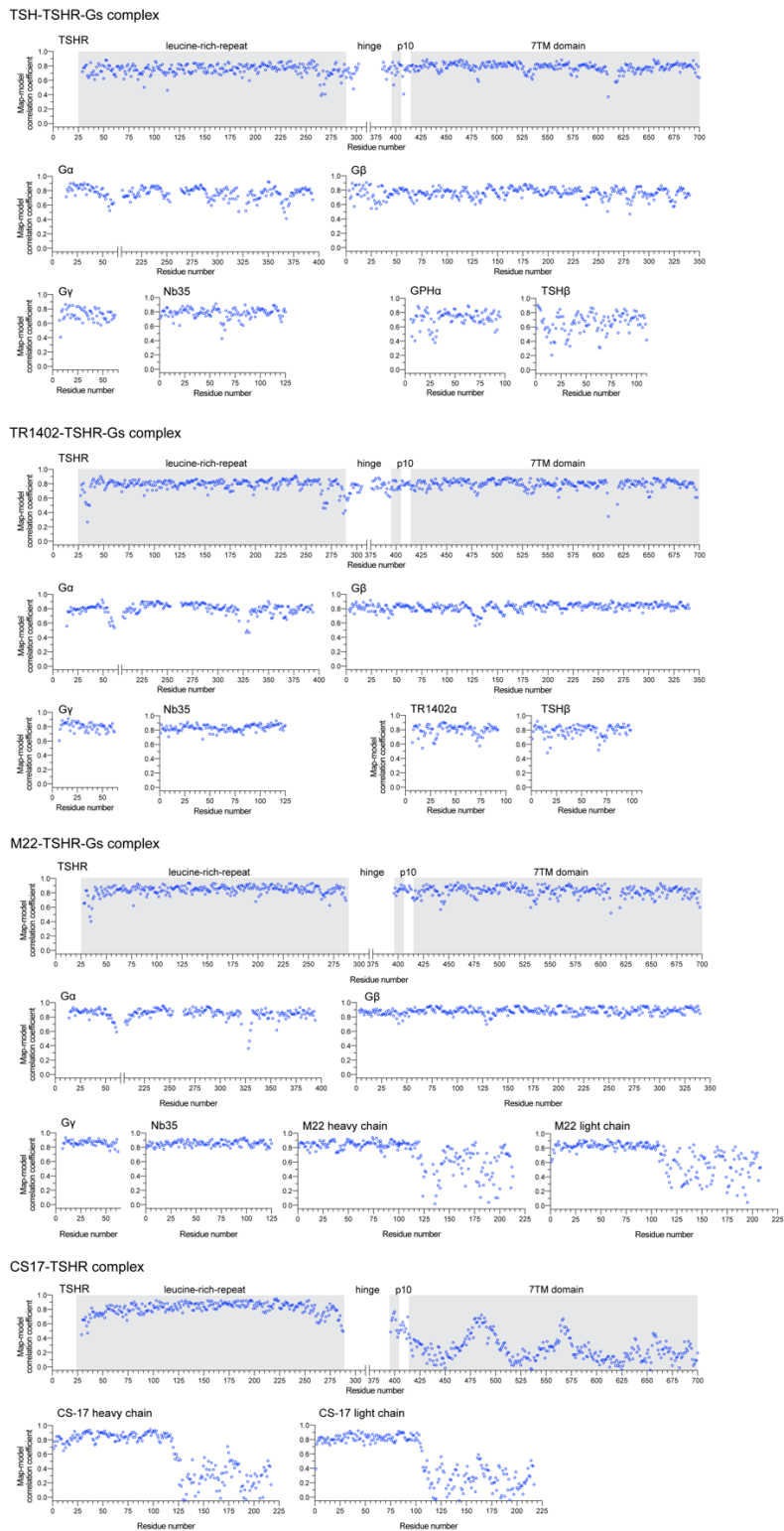


Figure 1.8 Model map correlation coefficients for cryo-EM structures.

Correlation values for resolved residues in each modeled chain are shown. Low values indicate regions that are poorly resolved, e.g. the 7TM domain of TSHR in CS-17-TSHR complex or the constant domains of the Fab fragments for CS-17 and M22.

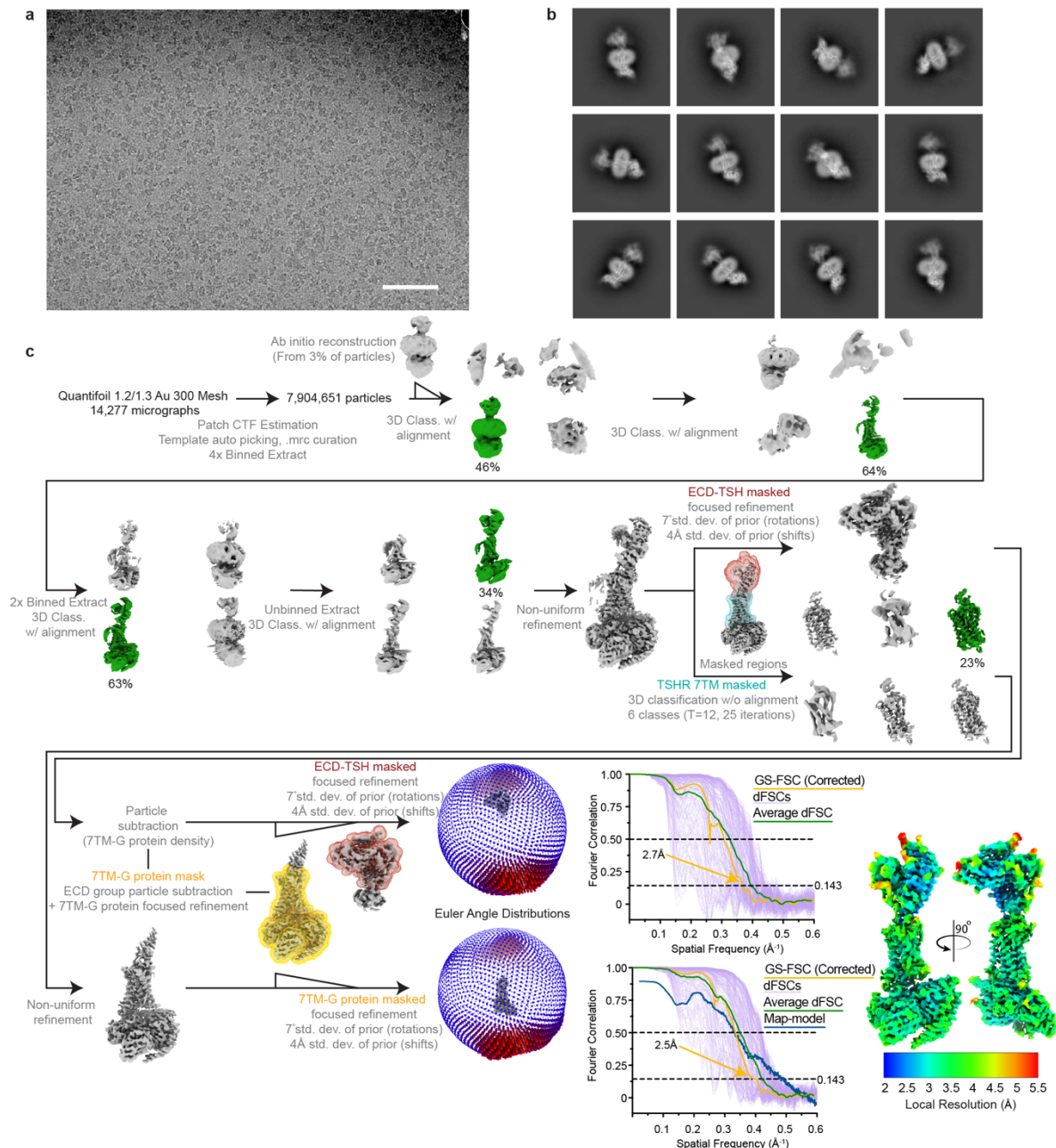


Figure 1.9 Cryo-EM data processing for TR1402-bound TSHR-G_s complex.

a) Representative image from 14,277 micrographs. Scale bar, 50 nm. **b)** Selected 2D class averages from final reconstruction. **c)** Processing approach used for reconstruction of TR1402-bound TSHR-G_s complex. A local resolution map was calculated from cryoSPARC using masks from indicated local refinement, then visualized with the composite map in the same scale. A viewing distribution plot was generated using scripts from the pyEM software suite and visualized in ChimeraX. GS-FSC and dFSC curves were generated in cryoSPARC and as previously described in Dang, S. *et al. Nature* 552, 426-429 (2017).

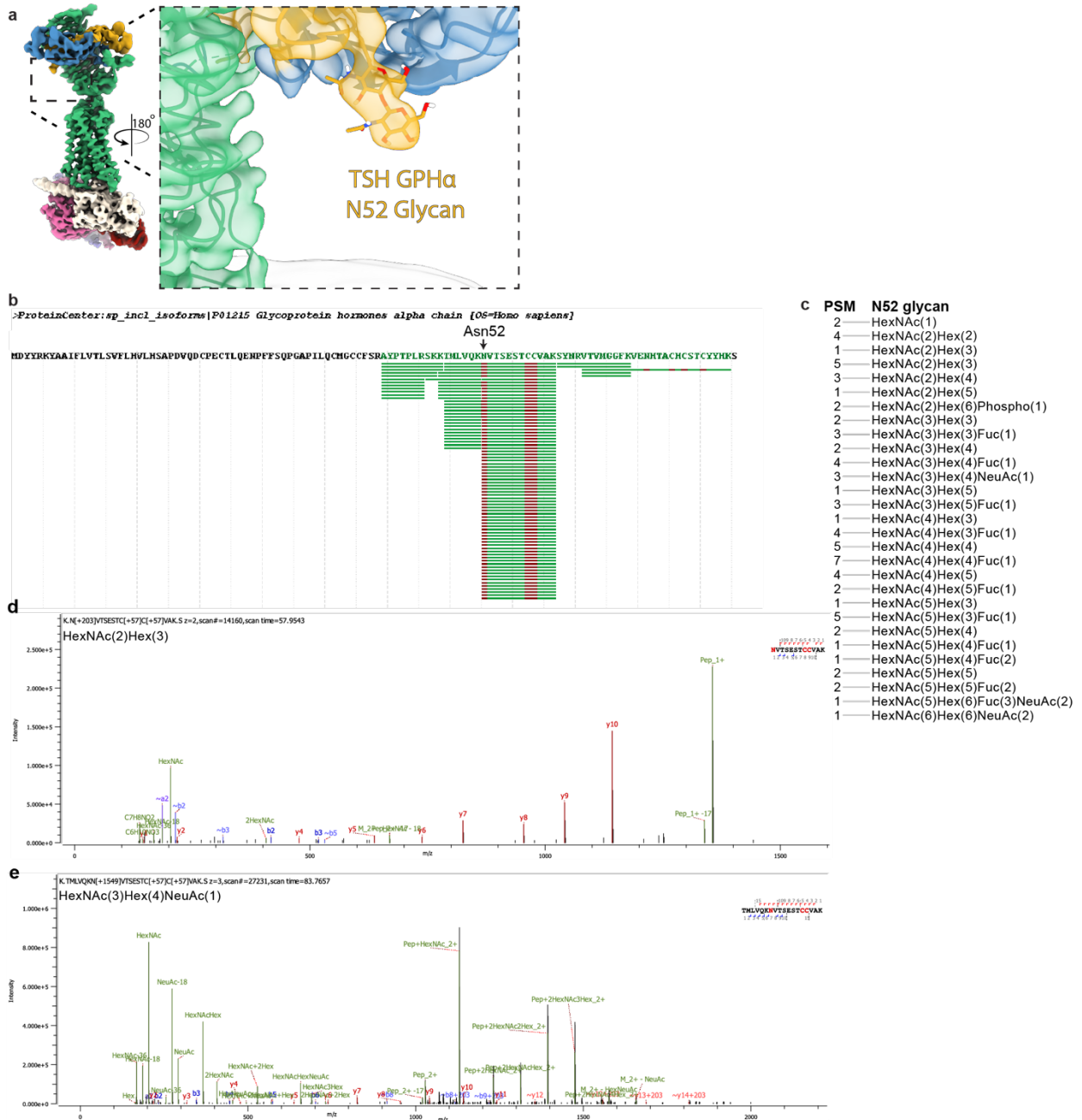


Figure 1.10 Glycosylation status of native human TSH.

a) Cryo-EM density map of TSH-bound TSHR showing resolved glycan density for Asn52. **b)** Peptide coverage of Asn52 from native human TSH GPH α chain in mass spectrometry experiments. **c)** Peptide spectra match (PSM) of glycans detected on Asn52 of GPH α chain. **d,e)** Representative MS/MS spectra showing Asn52 with various length glycans.

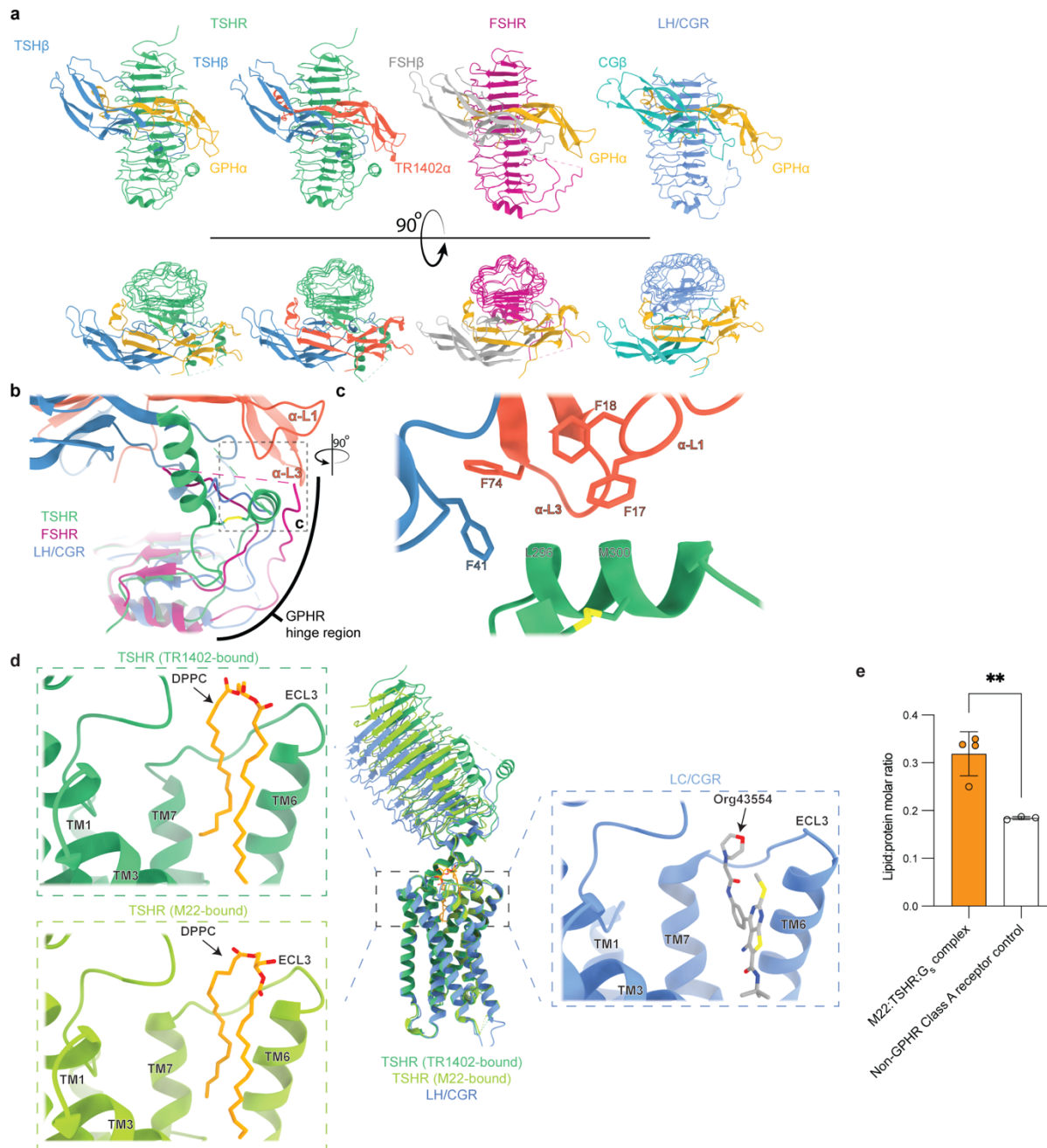


Figure 1.11 Receptor:hormone interaction comparisons across the glycoprotein hormone receptor family.

a) Comparison of TSHR, FSHR (PDB: 4AY9), and LH/CGR ECD:hormone (PDB: 7FIH) interactions shows the comparable concave binding-interface and lateral-ECD contacts made by GPH α and receptor-specific β subunits. **b)** Alignment of the GPHR ECDs reveals the variability in modeled poses of hinge-hormone contacts observed in prior GPHR:hormone structures. **c)** The N-terminal hinge-hormone disulfide-linked α helix presents a hydrophobic face to the α -L1 and α -L3 loops of the TR1402 α -chain and is stabilized by multiple phenylalanine contacts in the GPH α and TSH β chains. C-terminal hinge-region α helix hidden for visualization. **d)** Comparison of the TSHR and LHCGR (PDB: 7FIH) transmembrane pockets highlights the similarity in the DPPC and the allosteric agonist Org43554 binding sites. TM4/5 and ECL2 not shown.

e) Lipid:protein molar ratio comparison between the M22:TSHR:G_s complex and a non-GPHR Class A receptor control. Data points represent individual measurements from technical replicates (n=3 for non-GPHR control and n=4 for M22:TSHR:G_s complex) of the ratio of pmol of lipid DPPC per pmol of protein. ***P* = 0.0088; Unpaired two-tailed t test was used to calculate statistical differences in lipid:protein molar ratios.

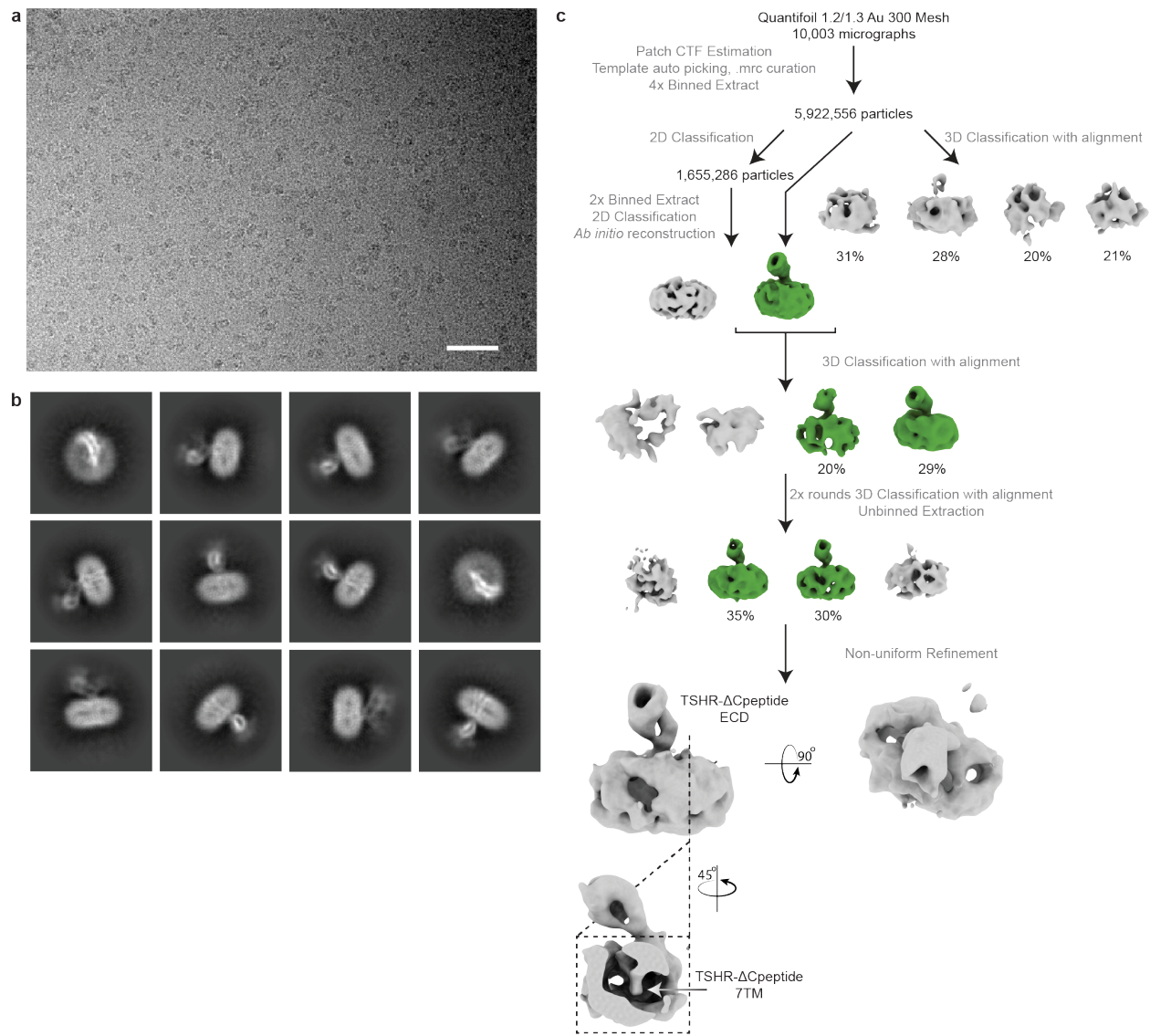


Figure 1.12 Cryo-EM data processing for Org 274179-0 bound TSHR.

a) Representative image from 10,003 micrographs. Scale bar, 50 nm. **b)** Selected 2D class averages generated from curated particles. **c)** Processing approach used for low resolution reconstruction. Despite starting with a similar or larger sized dataset as for other TSHR samples, the TSHR–Org 274179-0 complex did not yield high resolution reconstruction of the 7TM domain. A low-resolution reconstruction of the TSHR ECD was observed. This suggests potential flexibility between the TSHR 7TM domain and the ECD.

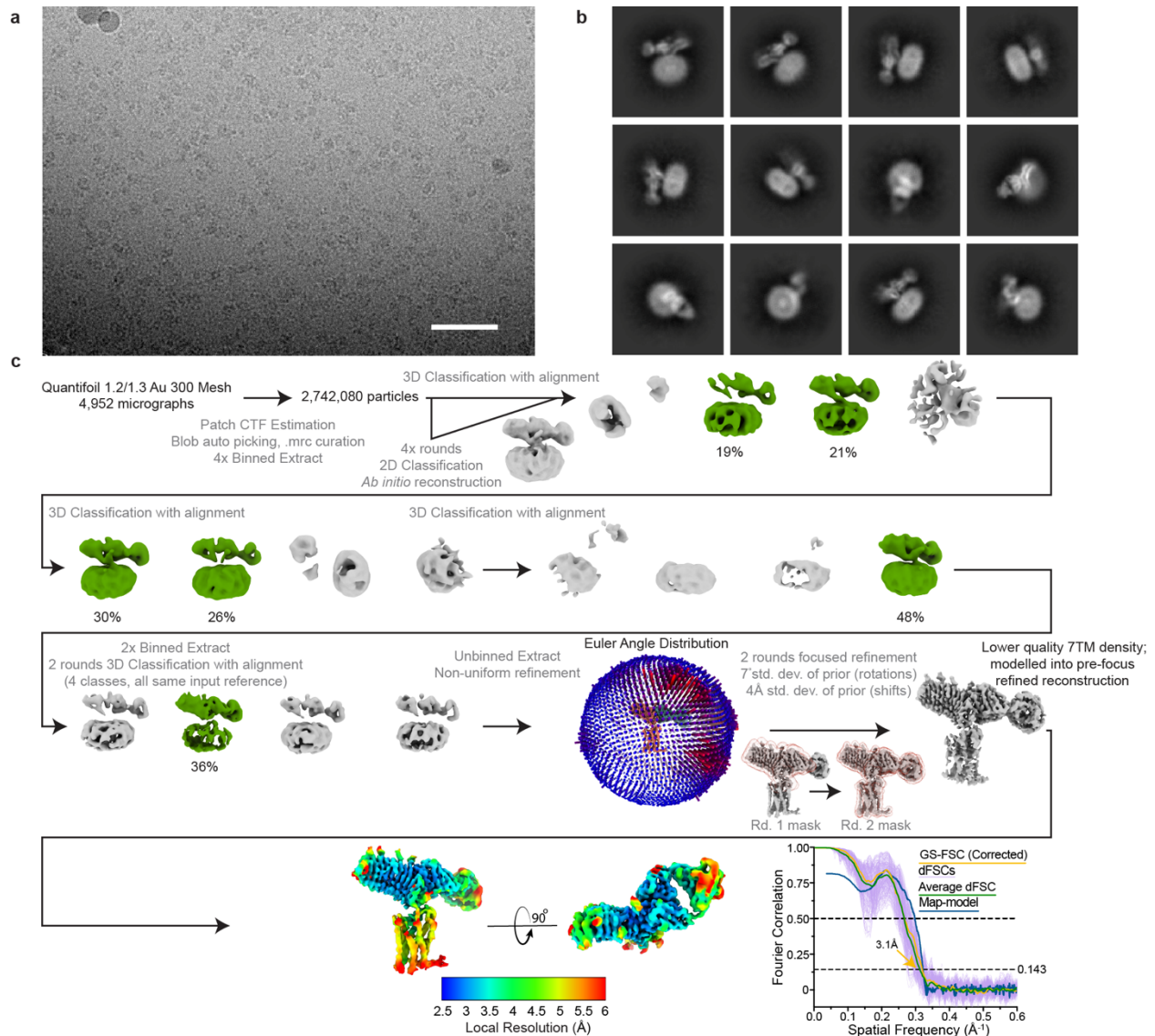


Figure 1.13 Cryo EM data processing for the CS-17 bound TSHR:Org 274179-0 complex. **a)** Representative image from 4,952 micrographs. Scale bar, 50 nm. **b)** Selected 2D class averages generated from the final reconstruction. **c)** Processing approach used for reconstruction of the complex. A viewing distribution plot was generated using scripts from the pyEM software suite and visualized in ChimeraX. Local resolution map generated from non-uniform refinement mask in cryoSPARC. GS-FSC and dFSC curves were generated in cryoSPARC and as previously described in Dang, S. *et al. Nature* 552, 426-429 (2017).

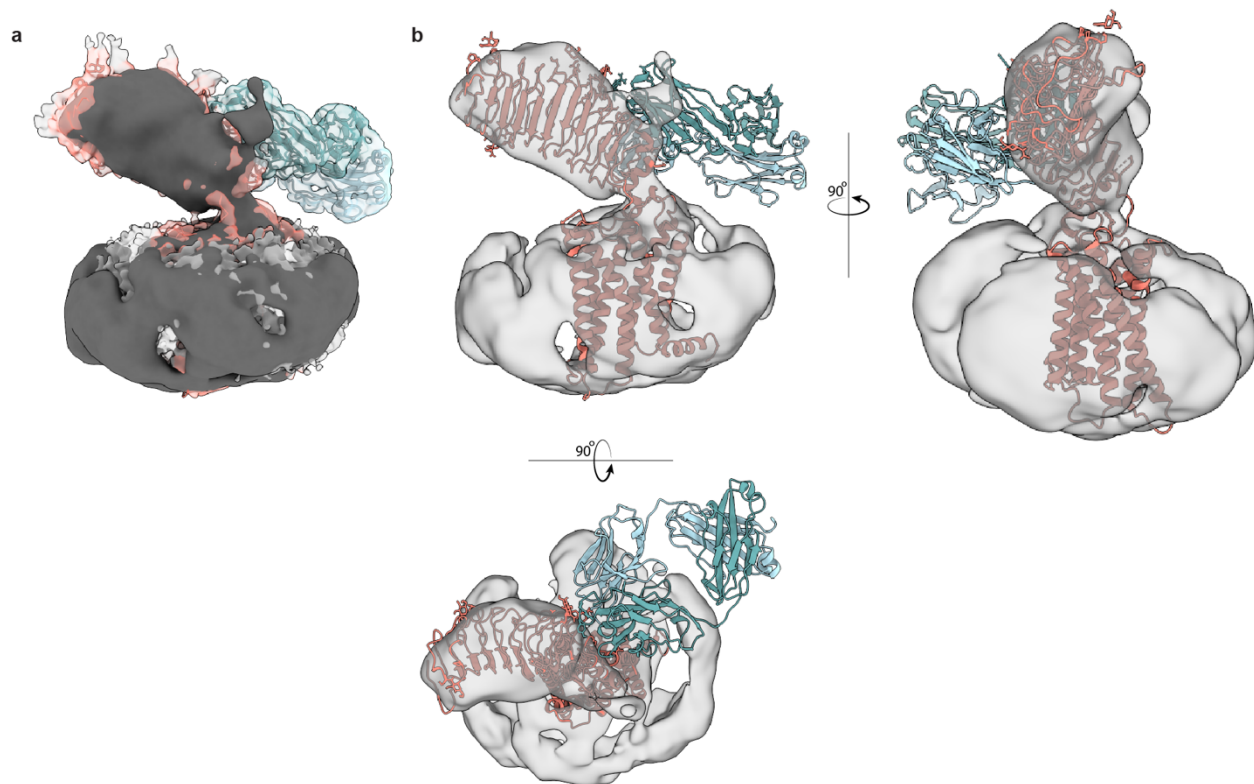


Figure 1.14 Comparison of Org 274179-0 bound- and CS-17 bound TSHR EM maps.

a) Org 274179-0 bound-TSHR EM density map (dark) fit into the CS-17-bound TSHR map (colored by TSHR-CS-17 model). **b)** CS-17-bound TSHR model fit into Org 274179-0 bound-TSHR map, suggesting that inactive state TSHR ECD orientations are similar between CS-17 and Org 274179-0-bound states.

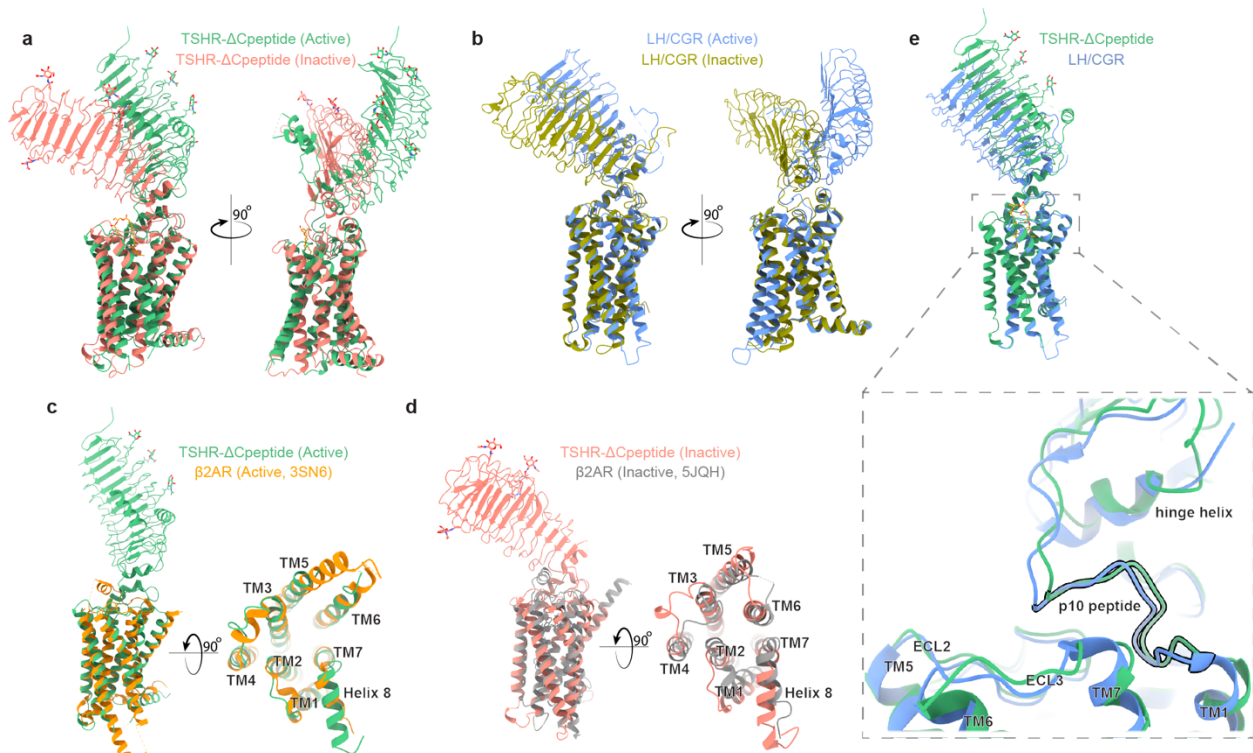


Figure 1.15 Comparison of TSHR activation with other GPCRs.

a) Ribbon diagram of TSHR in active and inactive conformations. **b)** Ribbon diagram of LH/CGR in active (PDB:7FIH) and inactive (PDB:7FIJ) conformations. A similar reorientation of the ECD is shared between TSHR and LH/CGR upon activation. **c)** Comparison of active TSHR to active β 2-adrenoceptor (β 2AR, PDB:3SN6). **d)** Comparison of inactive TSHR to inactive β 2AR, PDB:5JQH). **e)** Comparison of active conformations of TSHR and LH/CGR reveals similar overall structures of the 7TM domain and the p10 peptide.

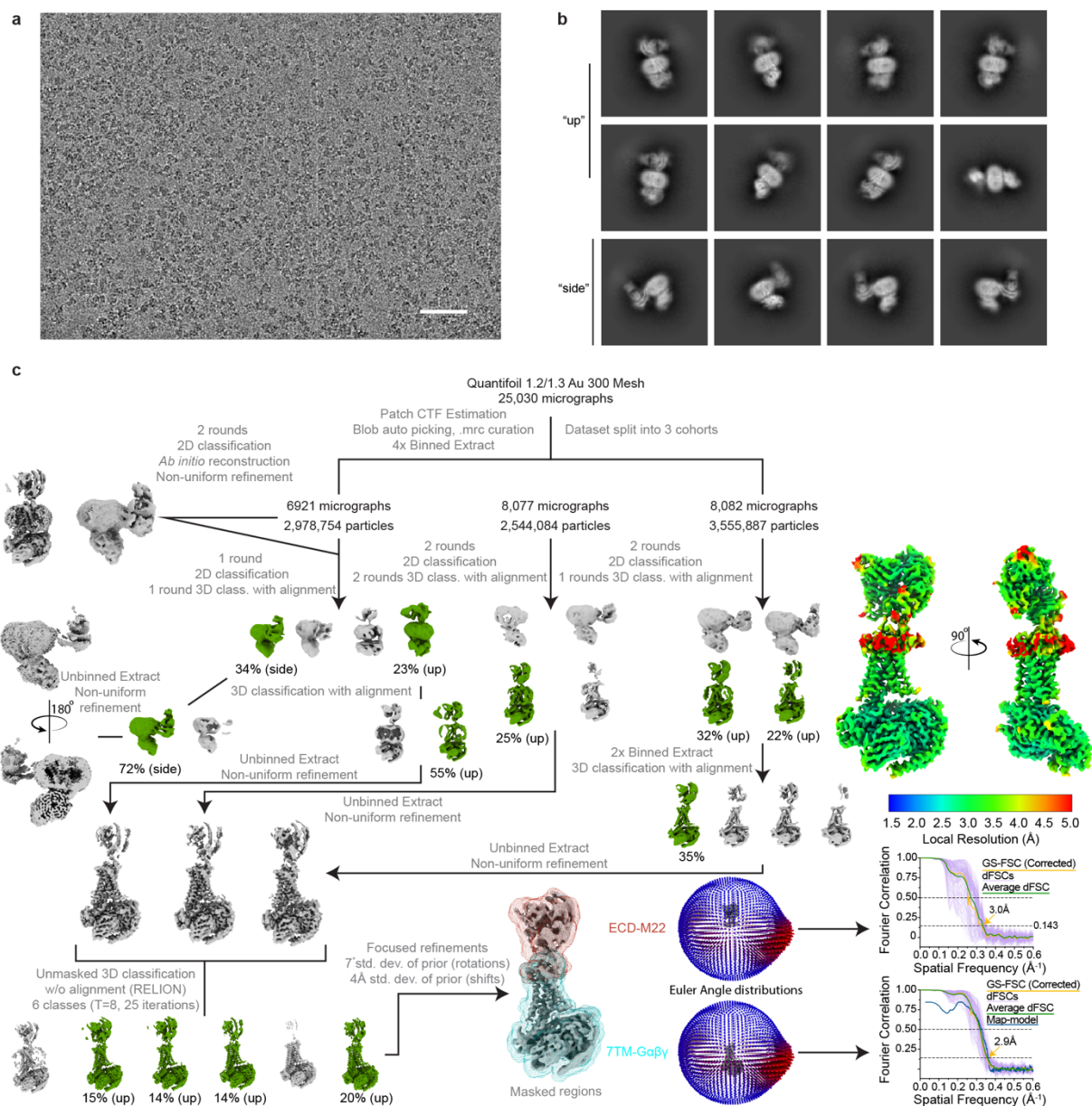


Figure 1.16 Cryo-EM data processing for M22-bound TSHR-Gs complex.

a) Representative image from 25,030 micrographs. Scale bar, 50 nm. **b)** Selected 2D class averages from final reconstruction. **c)** Processing approach used for reconstruction of M22-bound TSHR-G_s complex. A local resolution map was calculated from cryoSPARC using masks from indicated local refinement, then visualized with the composite map in the same scale. A viewing distribution plot was generated using scripts from the pyEM software suite and visualized in ChimeraX. GS-FSC and dFSC curves were generated in cryoSPARC and as previously described in Dang, S. *et al. Nature* 552, 426-429 (2017).

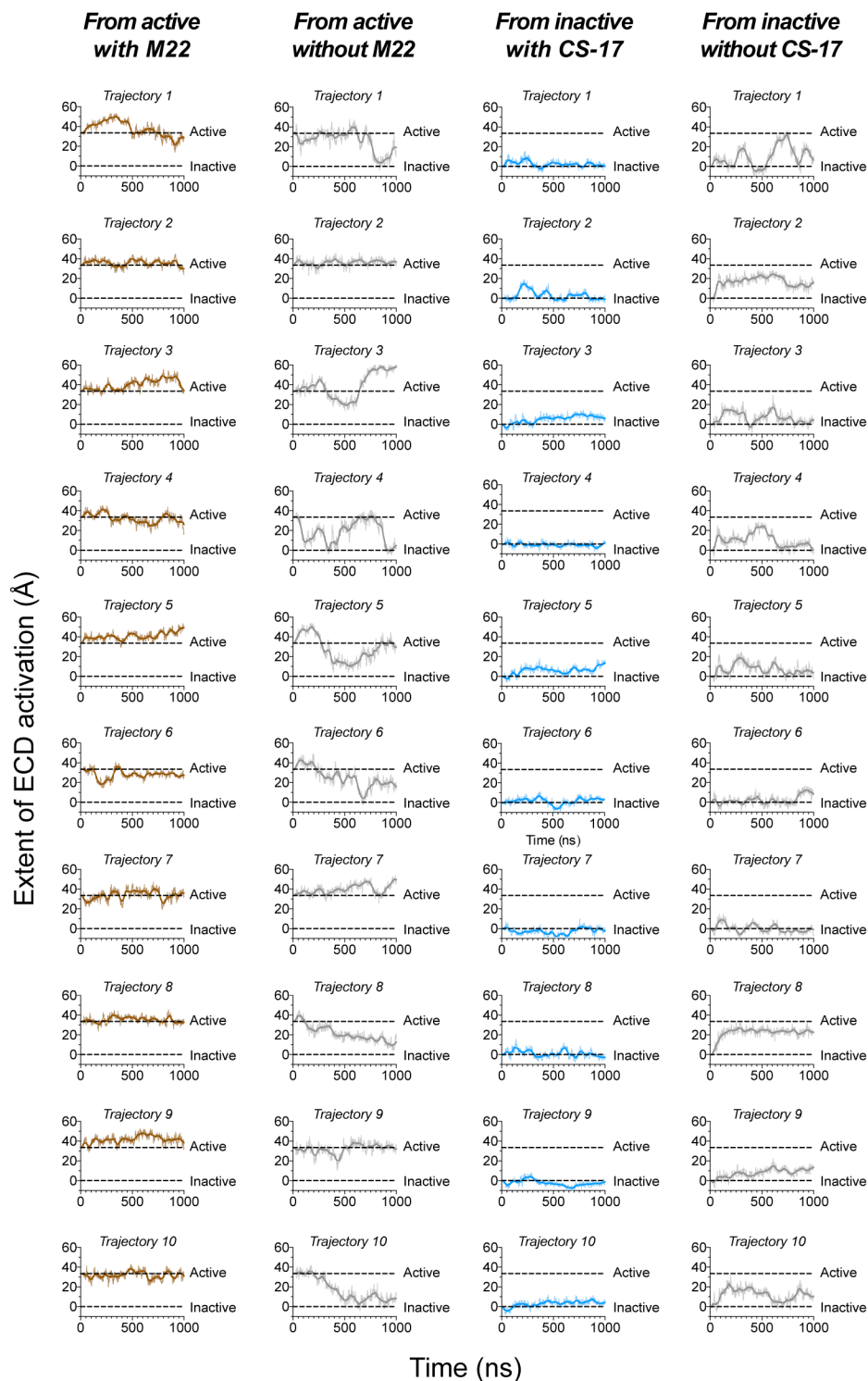


Figure 1.17 Extent of ECD activation for all individual simulations across four simulation conditions.

Dashed lines indicate the projection metric values (see Methods) in the inactive (0.0 Å) and active (33.5 Å) state cryo-EM structures. Thick traces indicate the moving average smoothed over a 25-ns window, and thin traces represent unsmoothed data. Each column represents a distinct simulation condition: started from the active structure, with M22 bound (brown traces, first

column); started from the active structure, with M22 removed (grey traces, second column); started from the inactive structure, with CS-17 bound (blue traces, third column); started from the inactive structure, with CS-17 removed (grey traces, fourth column).

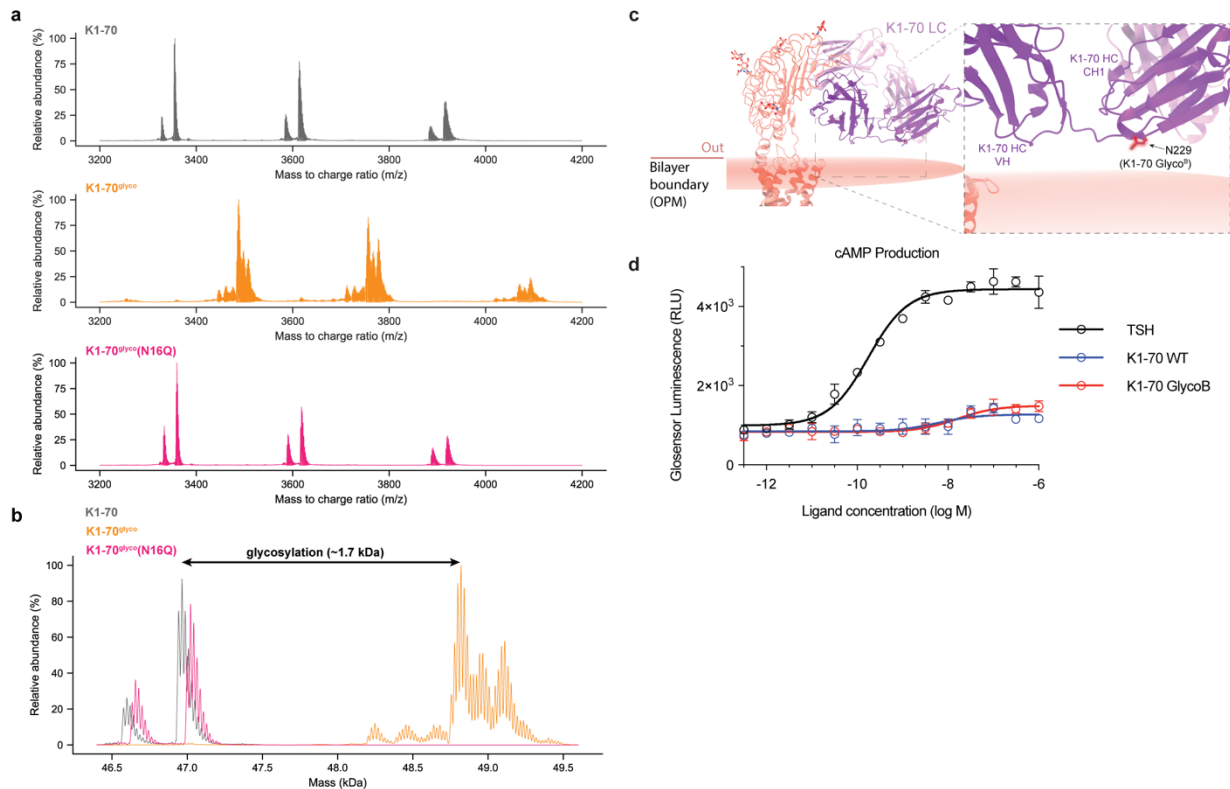


Figure 1.18 Glycosylation of engineered K1-70glyco construct.

a) Native mass spectrum (nMS) of K1-70, K1-70^{glyco}, and K1-70^{glyco}(N16Q) Fab fragments. **b)** Accurate mass assignment of Fab fragments. nMS demonstrates ~1.7 kDa increased mass for K1-70^{glyco} and more heterogeneity consistent with N-linked glycosylation. The smaller mass and increased homogeneity of the K1-70^{glyco}(N16Q) construct further supports glycosylation at the engineered N16 position. **c)** Crystal structure of K1-70 TSHR-ECD complex (PDB: 2XWT) aligned to CS-17 bound TSHR. The membrane-proximal N229 residue (of the “Glyco^B” glycosylation motif) is highlighted in red. **d)** cAMP production comparing K1-70 WT, K1-70 Glyco^B and M22 Fab fragment-mediated activation of TSHR. Plotted data points are means of triplicate measurements ± SD from one of three biological replicates.

1.6 Tables

Table 1.1 Cryo-electron microscopy data collection, refinement, and validation statistics

	TSH-bound TSHR-G _s (EMDB-25758) (EMDB-27649) (EMDB-27650) (PDB: 7T9I)	TR1402-bound TSHR-G _s (EMDB-26795) (EMDB-27647) (EMDB-27648) (PDB: 7UTZ)	Org 2274179-0- bound TSHR (EMDB-27640)	CS-17 Fab/ Org 2274179-0- bound TSHR (EMDB-25762) (PDB: 7T9M)	M22-bound TSHR-G _s (EMDB-25763) (EMDB-27651) (EMDB-25762) (PDB: 7T9N)
Data collection and processing					
Magnification	130,000	105,000	105,000	130,000	81,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e-/Å ²)	77	60	50	77	60
Defocus range (µm)	-0.8 to -2.2	-0.8 to -2.2	-0.8 to -2.4	-0.8 to -2.2	-0.8 to -2.0
Pixel size (Å) (physical)	0.662	0.831	0.85	0.644	0.844
Symmetry imposed	<i>C1</i>	<i>C1</i>	<i>C1</i>	<i>C1</i>	<i>C1</i>
Initial particle images (no.)	6,185,950	7,940,651	5,922,556	2,742,080	9,078,725
Final particle images (no.)	60,446 (7TM-G protein) 80,483 (ECD-TSH)	259,702 (7TM-G protein) 375,930 (ECD-TSH)	57,018	41,054	244,973
Map resolution (Å) (masked)	2.9 (7TM-G protein) 3.4 (ECD-TSH)	2.4 (7TM-G protein) 2.7 (ECD-TSH)	6.7	3.1	2.9 (7TM-G protein) 3.0 (ECD-M22)
FSC threshold	0.143	0.143	0.143	0.143	0.143
Refinement					
Initial model used (PDB code)	AlphaFold (TSH) TR1402-bound TSHR- G _s structure	AlphaFold (TR1402) M22-bound TSHR-G _s structure		AlphaFold (TSHR) AlphaFold (CS-17)	AlphaFold (TSHR), 7LJC (G protein) 3SN6 (Nb35) 3G04 (M22)
Model resolution (Å) (unmasked/masked)	3.9/3.8	3.2/3.1		3.5/3.3	3.2/3.1
FSC threshold	0.5	0.5		0.5	0.5
Model resolution range (Å)	3.8-50			3.5-50	3.2-50
Map sharpening <i>B</i> factor (Å ²)	-102 (ECD), -83 (7TM/MiniGα _s βγ)	-100 (ECD), -73 (7TM/MiniGα _s βγ)		-91	-90 (ECD) -80 (7TM/MiniGα _s βγ)
Model composition					
Non-hydrogen atoms	11991	12057		7827	13405
Protein residues	1532	1536		995	1723
Ligands	NAG: 7	NAG:7, DPC:1		NAG: 5	NAG: 3, DPC:1
<i>B</i> factors (Å²)					
Protein	104.8	79.2		93.5	89.9
Ligand	96.2	88.5		49.4	45.6
R.m.s. deviations					
Bond lengths (Å)	0.005	0.004		0.006	0.006
Bond angles (°)	0.916	0.678		1.221	1.013
Validation					
MolProbity score	1.36	1.36		1.53	1.58
Clashscore	3.3	3.91		5	5.62
Poor rotamers (%)	0.62	0		0.91	0.14
CaBLAM outliers (%)	1.34	1.34		1.63	1.32
Ramachandran plot					
Favored (%)	96.42	96.96		96.05	95.99
Allowed (%)	3.58	3.04		3.95	4.01
Disallowed (%)	0	0		0	0

Table 1.2 Summary of TSHR signaling studies.

Values are expressed as mean pEC₅₀ or mean E_{max} ± s.e.m. from (n) biological replicates. ND = Not determined.

		cAMP Production Assays			
		TSHR Construct	Ligand	cAMP pEC ₅₀	cAMP E _{max}
WT/ΔC Peptide Comparison	WT	TSH		-9.8 ± 0.006 (9)	100 ± 5 (9)
		M22		-9.2 ± 0.1 (9)	95 ± 3.6 (9)
		TR1402		-11.5 ± 0.09 (9)	105 ± 4 (9)
	ΔC Peptide	TSH		-9.8 ± 0.06 (9)	109 ± 7 (9)
		M22		-9.1 ± 0.03 (9)	102 ± 8 (9)
Y385 Mutants	Y385F	TSH		-9.1 ± 0.02 (9)	97 ± 12 (9)
		M22		-9.0 ± 0.1 (9)	87 ± 17 (9)
	Y385A	TSH		-9.1 ± 0.07 (9)	96 ± 14 (9)
		M22		-9.1 ± 0.03 (9)	93 ± 14 (9)
	TSHR Construct	Ligand	cAMP pEC ₅₀	cAMP E _{max}	
Lipid site displacement mutants	WT	TSH		-9.9 ± 0.004 (6)	100 ± 2 (6)
		M22		-9.1 ± 0.1 (6)	91.4 ± 0.6 (6)
	A644K	TSH		-9.0 ± 0.05 (6)	60 ± 7 (6)
		M22		-7.4 ± 0.05 (6)	54 ± 10 (6)
	A647K	TSH		-8.6 ± 0.03 (6)	59 ± 3.1 (6)
		M22		-7.5 ± 0.4 (6)	58 ± 2 (6)
	TSHR Construct	Ligand	cAMP pEC ₅₀	cAMP E _{max}	
Disulfide Mutant Assays	WT	TSH		-10 ± 0.5 (9)	100 ± 18 (9)
	WT (+500 μM TCEP)			-10.2 ± 0.5 (9)	112 ± 22 (9)
	K262C			ND	ND
	N483C			ND	ND
	K262C-N483C			ND	ND
	K262C-N483C (+500 μM TCEP)			-7.7 ± 0.3 (9)	79 ± 15 (9)
	TSHR Construct	Ligand	cAMP pEC ₅₀	cAMP E _{max}	
Activation Mutant Assays	WT	TSH		-9.5 ± 0.4 (8)	100 ± 3 (8)
	E409A			-7.5 ± 0.3 (8)	60 ± 0.6 (6)
	I486F			ND	ND
	TSHR Construct	Ligand	cAMP pEC ₅₀	cAMP E _{max}	
Glycosylation Mutant Assays	WT	M22		-9.2 ± 0.1 (9)	100 ± 8 (9)
		K1-70		-8.1 ± 0.2 (9)	17 ± 5 (9)
		K1-70 ^{glyco}		-8.2 ± 0.1 (9)	87 ± 9 (9)
		K1-70 ^{glyco} B		-8.3 ± 0.5 (9)	21 ± 4 (9)
		K1-70 (N16Q)		-8.1 ± 0.2 (9)	22 ± 6 (9)

1.7 Supplemental Files

Movie 1.1 TR1402-bound TSHR 3D Variability analysis

A movie displaying the conformational flexibility between the TSHR ECD and transmembrane domains.

1.8 Materials and Methods

Expression and purification of TSHR for active state structures

The human *TSHR* gene with an N-terminal influenza hemagglutinin signal sequence and FLAG (DYKDDDK) epitope tag was cloned into a custom pCDNA3.1 vector containing a tetracycline inducible cassette. To improve expression of intact receptor for structural studies, the sequence for the 50-residue hinge-region “C-peptide” (Ala317-Phe366) was removed via site directed mutagenesis (New England Biolabs). The construct further included the miniG_{s399} protein¹⁷, which was fused to the C-terminus of TSHR with a human rhinovirus 3C protease cleavage sequence flanked by three-residue Gly/Ser linkers. This construct (TSHR-ΔCpep-miniG_{s399}) was transfected into inducible Expi293F TetR cells (Thermo Fisher) using the ExpiFectamine transfection reagent per manufacturer instructions. After 24 hours, protein expression was induced with 2 μg/mL doxycycline hyclate, and the culture was placed in a 30°C incubator for 36 hours prior to harvesting by centrifugation. Pelleted cells were washed with 50 mL of phosphate buffered saline, pH 7.5 prior to storage at -80°C. For receptor purification, frozen cells were hypotonically lysed in 50 mM HEPES, pH 7.5, 1 mM EDTA, 160 μg/mL Benzamidine, 2 μg/mL Leupeptin for 10 minutes at 25°C. The membrane fraction was collected by centrifugation, and the fusion protein was extracted with 50 mM HEPES, pH 7.5, 300 mM NaCl, 1% (w/v) glyco-diosgenin (GDN, Anatrace), 0.1% (w/v) cholesteryl hemisuccinate (CHS, Steraloids), 2 mM MgCl₂, 2 mM CaCl₂, 160 μg/mL Benzamidine, 2 μg/mL Leupeptin with dounce homogenization and incubation with stirring for one hour at 4°C. The soluble fraction was separated from the insoluble fraction by centrifugation and was incubated in batch for 1 hour at 4°C with homemade M1-FLAG-antibody conjugated sepharose beads. Sepharose resin was then washed extensively with 50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1% (w/v) GDN, 0.01% (w/v) CHS, 2 mM MgCl₂, 2 mM CaCl₂ and then with 50 mM HEPES,

pH 7.5, 150 mM NaCl, 0.0075% (w/v) GDN, 0.00075% (w/v) CHS, 2 mM MgCl₂, 2 mM CaCl₂ prior to elution with 50 mM HEPES, pH 7.5, 150 mM NaCl, 0.0075% (w/v) GDN, 0.00075% (w/v) CHS, 5 mM EDTA, 0.2 mg/mL FLAG peptide. Post-elution, TSHR-ΔCpep-miniG_{s399} fusion protein was concentrated in a 100 kDa MWCO Amicon spin concentrator, and injected onto a Superdex200 Increase 10/300GL (Cytiva) gel filtration column equilibrated in 50 mM HEPES, pH 7.5, 150 mM NaCl, 0.005% (w/v) GDN, and 0.0005% CHS to isolate monodisperse material for further complexing with agonist, Gβ₁γ₂ heterodimer, and Nb35¹⁸.

Expression and purification of TSHR for inactive state structures

A pCDNA3.1-containing TSHR-ΔCpep construct was generated via site directed mutagenesis. Expi293 cell transfection, doxycycline induction, and expression were performed as described for TSHR-ΔCpep-miniG_{s399} with the additional supplement of dimethylsulfoxide solubilized 1 μM Org 274179-0³⁴ post-induction. Purification of TSHR-ΔCpep was performed as described for TSHR-ΔCpep-miniG_{s399} with all buffers supplemented with 1 μM Org 274179-0.

Expression and purification of Gβ₁γ₂

Human Gβ₁γ₂ heterodimer was expressed in *Trichoplusia ni* (Hi5) insect cells using a single baculovirus generated in *Spodoptera frugiperda* (Sf9) insect cells. A bicistronic pVLDual construct contained the Gβ₁ subunit with a N-terminal 6x His-tag, and an untagged human Gγ₂ subunit. For expression, Hi5 insect cells were transduced with baculovirus at a density of ~3.0 x 10⁶ cells/mL, grown at 27 °C shaking at 130 rpm. 48 hours post-transduction, cells were harvested and washed in a hypotonic buffer containing 20 mM HEPES, pH 8.0, 5 mM β-mercaptoethanol (β-ME), and protease inhibitors (20 μg/mL leupeptin, 160 μg/mL benzamidine). The membrane fraction was

then separated by centrifugation and solubilized with 20 mM HEPES pH 8.0, 100 mM sodium chloride, 1.0% sodium cholate, 0.05% dodecylmaltoside (Anatrace), and 5 mM β -mercaptoethanol (β -ME). Solubilized G β γ ₂ heterodimer was then incubated with HisPur™ Ni-NTA resin (Thermo Scientific) in batch. Bound G β γ ₂ heterodimer was washed extensively and detergent was slowly exchanged to 0.1% (w/v) lauryl maltose neopentyl glycol (L-MNG, Anatrace) and 0.01% CHS before elution with 20 mM HEPES pH 7.5, 100 mM NaCl, 0.1% L-MNG, 0.01% CHS, 270 mM imidazole, 1 mM dithiothreitol (DTT), and protease inhibitors. Eluted G β γ ₂ heterodimer was pooled and rhinovirus 3C protease was added to cleave the N-terminal 6x His-tag during overnight dialysis in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.02% L-MNG, 0.002% CHS, 1 mM DTT, and 10 mM imidazole. To remove uncleaved G β γ ₂, dialyzed material was incubated with HisPur™ Ni-NTA resin in batch. The unbound fraction was then incubated for 1 hour at 4 °C with lambda phosphatase (New England Biolabs), calf intestinal phosphatase (New England Biolabs), and antarctic phosphatase (New England Biolabs) for dephosphorylation. Final anion exchange chromatography was performed using a MonoQ 4.6/100 PE (Cytiva) column to purify only geranylgeranylated heterodimer. The resulting protein was pooled and dialyzed overnight in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.02% L-MNG, and 100 μ M TCEP, and concentrated with a 3 kDa centrifugal concentrator to a final concentration of 162 μ M. Glycerol was added to a final concentration of 20%, and the protein was flash frozen in liquid nitrogen and stored at -80 °C until further use.

Expression and purification of Nb35

A pET-26b vector containing the Nb35 sequence with a carboxy-terminal Protein C affinity tag (EDQVDPRLIDGK) was transformed into BL21 Rosetta *Escherichia coli* cells and inoculated into

8 L of Terrific Broth supplemented with 0.1% glucose, 2 mM MgCl₂, and 50 µg/mL kanamycin. Cells were induced with 400 µM IPTG at OD₆₀₀ = 0.6 and allowed to express at 20 °C for 21 hours. Harvested cells were incubated SET Buffer (200 mM Tris pH 8.0, 500 mM sucrose, 0.5 mM EDTA) in the presence of protease inhibitors (20 µg/mL leupeptin, 160 µg/mL benzamidine) and benzonase. To initiate hypotonic lysis, 2 volumes of deionized water were added to the cell mixture after 30 minutes of SET buffer mixing. Following lysis, NaCl was added to 150 mM, CaCl₂ was added to 2 mM, and MgCl₂ was added to 2mM and lysate was spun down to remove the insoluble fraction. Supernatant was incubated with homemade anti-Protein C antibody-coupled sepharose. Nb35 was eluted with 20 mM HEPES pH 7.5, 100 mM NaCl, and 2 mM CaCl₂, 0.2 mg/mL Protein C peptide, and 5 mM EDTA pH 8.0, concentrated in a 10 kDa MWCO Amicon filter and injected over a Superdex S75 Increase 10/300 GL column (Cytiva) size-exclusion chromatography column equilibrated in 20 mM HEPES pH 7.5, 100 mM NaCl. Monodisperse Nb35 fractions were pooled, concentrated, and supplemented with 20% glycerol prior to flash freezing in liquid nitrogen for storage at -80°C until further use.

Expression and purification of M22 agonist Fab

The M22 heavy and light chains⁵⁷ were cloned into pFastBac with a GP67 signal peptide and N-terminal FLAG epitope and a P2A self-cleaving peptide sequence between the heavy and light chains. The construct also included a C-terminal 8x Histidine epitope tag on the heavy chain. Baculovirus was generated using the Bac-to-bac method (Thermo Fisher), and used to transduce *Trichoplusia ni* (Hi5) insect cells (Expression Systems) at a density of 2.0 x 10⁶ cells/mL. Cells were cultured with shaking at 120 RPM at 25°C for 60 hours. M22 Fab was purified from the cell supernatant by first adjusting pH to 8.0 with 1 M Tris prior to a 1 hour incubation with 5 mM CaCl₂,

5 mM MgCl₂, and 1 mM NiCl₂ to precipitate chelators. After separating the insoluble fraction, the 8x-His tagged M22 Fab was captured on HisPur™ Ni-NTA resin, washed in a buffer comprised of 50 mM HEPES, pH 7.5, 500 mM NaCl, 20 mM imidazole, pH 7.5 and eluted with the same buffer supplemented with 500 mM imidazole. The eluate was further purified by size exclusion chromatography over a Superdex200 Increase 10/300GL gel filtration column equilibrated in 50 mM HEPES, 100 mM NaCl, pH 7.5. Monodisperse M22 Fab fractions were pooled, concentrated, and supplemented with 20% glycerol prior to flash freezing in liquid nitrogen for storage at -80°C until further use.

Expression and purification of WT K1-70, K1-70^{glyco} and K1-70^{glyco}(N16Q) Fabs

Heavy and light chain sequences for K1-70 IgG⁴⁷ were cloned into a pCDNA3.1 vector containing human IgG1 constant regions. Resulting constructs were transfected into Expi293F cells using a 2:1 mass ratio of heavy to light chain using the Expifectamine transfection kit, per the manufacturer's instructions. After 5 days, cultures were harvested and the supernatant purified over a 1 mL MabSelect SuRe HiTrap column (Cytiva) equilibrated in 100 mM sodium phosphate, 150 mM sodium chloride, pH 7.2. Monoclonal antibodies were eluted with 100 mM Glycine, pH 3.0, neutralized with 1.5 M Tris pH 8.0, and dialyzed into 20mM sodium phosphate, 10 mM EDTA, pH 7.0 for generation of Fab fragments. Immobilized papain agarose was equilibrated in the IgG dialysis buffer with freshly added 20 mM Cysteine-HCl, pH adjusted to 7.0. IgG was concentrated to ~40 mg/mL with a 100 kDa MWCO Amicon filter, diluted 1:1 with the IgG dialysis buffer plus Cysteine-HCl, and added to the equilibrated papain agarose. The IgG-papain suspension was placed on a shaker at 37C for overnight digestion. Post-digest, Fc-fragments were removed via 1 hour, room temperature batch incubation with phosphate buffered saline-equilibrated Protein A agarose

(Pierce). Fab fragments were concentrated in a 10 kDa MWCO Amicon filter and injected onto a Superdex200 Increase 10/300GL gel filtration column equilibrated in 50 mM HEPES, 100 mM NaCl, pH 7.5. Monodisperse Fab fractions were pooled, and digestion/-Fc removal was confirmed by SDS-PAGE.

To produce a glycosylated version of K1-70 (K1-70^{glyco}), a glycosylation motif was introduced at position 16 in the heavy chain in a loop region that connects the A and B strands of the VH IgG domain. The sequence “KKPGQS” was replaced with “KKPGNGS” to generate an N-linked N-x-S/T glycosylation motif. As a control for glycosylation, we also generated a version of K1-70^{glyco} with a N16Q mutation K1-70^{glyco}(N16Q). The resulting constructs were then purified as full length IgG as described above for K1-70. Despite introduction of the glycosylation motif in K1-70^{glyco}, the extent of glycosylation was not complete. To enrich for the glycosylated fraction of K1-70^{glyco}, monodisperse K1-70^{glyco} Fab was brought to 1 mM Mn²⁺ and Ca²⁺ and loaded over a Concanavalin A-conjugated Sepharose 4B packed HiTrap column (Cytiva) pre-equilibrated in 20 mM HEPES, 500 mM NaCl, 1 mM Mn²⁺, 1 mM Ca²⁺, pH 7.4. The glycosylated fraction K1-70^{glyco} was eluted with 20 mM HEPES, 500 mM NaCl, 500 mM methyl alpha-D-mannopyranoside, pH 7.4, fractions were concentrated to 100 μM and flash frozen in liquid nitrogen for storage at -80°C until further use.

Expression and purification of CS-17 Fab

Heavy and light chain sequences of CS-17 were determined from sequencing of CS-17 murine hybridoma cell line PTA-8174 (ATCC) by Genscript. The resulting sequences were cloned into a pCDNA3.4-containing murine IgG2a construct. Heavy and light chain constructs were transfected,

expressed and purified as previously described for K1-70 IgG. CS-17 Fab generation also followed identical steps to K1-70 Fab generation.

Expression and purification of TR1402

The human TSH analog TR1402 was produced at Trophogen, Inc. (Rockville, MD), following the procedures for site-directed mutagenesis and expression of recombinant hormones as described previously¹⁹. Briefly, TR1402 was produced in stably transfected Chinese hamster ovary (CHO) cells and purified by a combination of dye, ion exchange, and gel filtration HPLC.

Preparation of inactive and- active-state TSHR complexes

To prepare the TSH, TR1402, and M22-activated TSHR complexes, purified TSHR- Δ Cpep-miniG_{s399} was incubated with a 2-fold molar excess of purified G β ₁ γ ₂, Nb35, and either native human TSH (National Hormone and Pituitary Program, resuspended from lyophilized powder in 50 mM HEPES, pH 7.4 150 mM NaCl), human TSH analog TR1402 (resuspended from lyophilized powder in 50 mM HEPES, pH 7.4 150 mM NaCl), or M22 Fab and incubated overnight at 4 °C. After incubation, the complexed material was purified with anti-Protein C antibody Sepharose resin to purify Nb35-bound complex. Protein C Sepharose was washed with 20 column volumes of 50 mM HEPES, 150 mM NaCl, 0.0075% GDN (w/v), 2 mM CaCl₂ prior to elution with 50 mM HEPES, 150 mM NaCl, 0.005% GDN (w/v), 5 mM EDTA, 0.2 mg/mL Protein C peptide. The eluted fractions were concentrated with a 100 kDa MWCO Amicon filter, and injected onto a Superdex200 Increase 10/300GL gel filtration column equilibrated in 50 mM HEPES, 150 mM NaCl, 0.005% GDN (w/v), pH 7.5. Monodisperse fractions were concentrated with a 100 kDa MWCO Amicon filter immediately prior to cryo-EM grid preparation.

For formation of the inactive state CS-17-bound TSHR- Δ Cpep complex, the CS-17 Fab was incubated at a 2x molar excess with SEC-purified TSHR- Δ Cpep overnight at 4°C, then concentrated in a 100 kDa MWCO Amicon filter prior to cryo-EM grid preparation prior to size exclusion chromatography over a Superdex200 Increase 10/300GL gel filtration column equilibrated in 50 mM HEPES, 150 mM NaCl, 0.005% GDN (w/v), pH 7.5 and 1 μ M Org 274179-0 . Monodisperse fractions of the resulting complex were concentrated with a 50 kDa MWCO concentrator immediately prior to cryo-EM grid preparation.

Cryo-EM vitrification, data collection, and processing

TSH-bound TSHR-G_s complex

The TSH-bound TSHR-G_s complex was concentrated to 23 μ M and 3 μ L was applied onto a glow discharged 300 mesh 1.2/1.3 gold grid covered in a holey gold film (UltraAufoil). After a 30 second hold at 4 °C, excess sample was removed with a blotting time of 3 seconds and a blotting force of 0 prior to plunge freezing into liquid ethane using a Vitrobot Mark IV (Thermo Fisher). 15,346 super-resolution videos were recorded with a K3 detector (Gatan) on a Titan Krios (Thermo Fisher) microscope operated at 300 keV with a BioQuantum post-column energy filter set to a zero-loss energy selection slit width set of 10 eV. 66-frame videos were recorded for 2 s at a nominal magnification of 130,000x (physical pixel size of 0.664 Å/pixel) and a defocus range of -0.8 to - 2.2 μ m for a total dose of 77 e-/Å².

Super-resolution videos of the TSH-TSHR-G_s complex were motion corrected, binned to physical pixel size, and dose fractionated on-the-fly during data collection using UCSF MotionCor2⁵⁸.

Corrected micrographs were imported into cryoSPARC v3.1⁵⁹ for CTF estimation via the “Patch CTF Estimation” job type. Templates for particle picking were generated from projections of the same TSHR complex reconstructed from a previous 200 keV imaging session. Particle picking templates were low-pass filtered to 20 Å and used to pick 9,151,778 particles. After picking, estimated CTF fit resolution >5 Å and relative ice thickness outlier measurements were used to remove low-quality micrographs before further processing. 6,185,950 curated particles were extracted in a 512-pixel box and Fourier cropped to 128 pixels before undergoing a round of 3D classification with alignment utilizing one 20Å low-pass filtered reconstruction and three “random” reconstructions generated from a prematurely truncated *Ab initio* reconstruction job. 2,474,380 particles classified into the “TSHR” class were extracted in a 512-pixel box and Fourier cropped to 256 pixels for two additional rounds of 3D classification with alignment, utilizing the same class distributions as previously described. From these 2 rounds of 3D classification, and because the TSHR ECD remained poorly resolved, 734,891 particles belonging to the “TSHR” class were extracted into a 512-pixel box for a new 3D classification workflow. First, these particles were subject to one round of “non-uniform” refinement⁶⁰ in cryoSPARC and then classified utilizing two different masking schemes (TSHR 7TM domain or TSHR ECD:TSH complex) using alignment-free 3D classification in RELION v3.1.2⁶¹. Particles in qualitatively “good” looking classes were re-imported into cryoSPARC for focused refinements. Particles from the TSHR 7TM domain-masked 3D classification were subject to a round of focused refinement with a mask encompassing the TSHR 7TM domain and G protein/Nb35 complex. Particles from the TSHR ECD:TSH complex masked-3D classification were subject to two rounds of focused refinement using the same mask as in RELION, decreasing mask pixel dilation between rounds. For each focused refinement, pose/shift gaussian priors (7° standard deviation of prior (rotation),

4Å standard deviation of prior (shifts)) were used to limit large deviations from the initially determined poses. Directional FSC curves for each final reconstruction were calculated, and each map was loaded into ChimeraX⁶² for generation of a composite map via the *vop maximum* command.

TR1402-bound TSHR-G_s complex

The human TSH analog TR1402-bound TSHR-G_s complex was concentrated to 11.6 μM and 3 μL was applied onto a glow discharged 300 mesh 1.2/1.3 gold grid covered in a holey gold film and plunge frozen into liquid ethane with identical blotting conditions as with the TSH complex. 14,277 super-resolution videos were recorded with a K3 detector on a 300 kV FEI Titan Krios microscope located at the HHMI Janelia Research Campus equipped with a spherical aberration corrector and a post-column BioQuantum energy filter set to a zero-loss energy selection slit width set of 20 eV. 60-frame videos were recorded across a defocus range of -0.8 to - 2.2 μm at a nominal magnification of 105,000x (physical pixel size of 0.831 Å/pixel) for 4.4s, resulting in a total exposure dose of 60 e-/Å².

14,276 super-resolution videos were motion corrected, binned to physical pixel size and dose weighted, post acquisition, using UCSF MotionCor2. Contrast transfer function estimation was performed with the cryoSPARC “Patch CTF Estimation” job prior to template-assisted, automated particle picking. Templates were generated from projections of the final native human TSH-bound TSHR-G_s complex. Micrographs for additional data processing were curated as previously described, and the subsequent 7,904,651 particles were extracted in a 416-pixel box and Fourier cropped to 104 pixels. To generate initial models for 3D classification, 200,000 particles were used

to generate 3 *Ab initio* reconstructions. The class containing a reconstruction resembling a TR1402-bound complex was used along with three additional reconstructions generated from a truncated *Ab initio* reconstruction job, as previously described. The picked, curated particles were then subject to 2 rounds of 3D classification without alignment. Particles classified into the TR1402-bound class at the end of these 2 rounds were extracted in a 416 pixel box and Fourier cropped to 208 pixels. Particles then underwent one round of 3D classification with alignment, using four TR1402-bound TSHR-G_s complex reconstructions as initial models. The 375,930 particles in the highest resolution reconstruction class were then extracted in a 416-pixel box without additional Fourier cropping. One round of Non-Uniform refinement was performed to assign accurate initial angles for subsequent alignment-free classification jobs. At this point, based on the native human TSH processing particles results, these particles were split into separate ECD and transmembrane domain:G protein refinement pathways. For the ECD refinement approach, using a mask encompassing only the ECD, hinge region and TR1402 density, particles were locally refined with the same pose/shift gaussian priors as in the native human TSH processing. Then, particle subtraction using a mask of the transmembrane domain:G protein complex was performed. These subtracted particles were subject to a final round of local refinement using the same parameters as previously described.

For the transmembrane domain:G protein classification/refinement workflow, particles from the Non-Uniform refinement job were imported into RELION and subject to 3D classification without alignment using a mask encompassing only the transmembrane domain density. The 259,702 particles belonging to the highest resolution class from this alignment-free classification run were re-imported into CryoSPARC and subject to another round of Non-Uniform refinement. A final

round of local refinement using a mask encompassing the transmembrane domain:G protein complex was performed, and the subsequent reconstruction was used for generating a composite map with the ECD-TR1402 reconstruction in ChimeraX using the *vop maximum* command. Directional FSC curves were calculated for each independent reconstruction.

M22-bound TSHR-G_s complex

The M22-bound TSHR-G_s complex was concentrated to 13.8 μM and 3 μL was applied onto a glow discharged 300 mesh 1.2/1.3 gold grid covered in a holey gold film and plunge frozen into liquid ethane utilizing identical blotting procedures as with the TSH complex. 25,030 super-resolution videos were recorded with the same 300 kV microscope and detector used in the TR1402-bound TSHR-G_s complex collection. 60-frame videos were recorded across a defocus range of -0.8 to - 2.0 μm at a nominal magnification of 81,000x (physical pixel size of 0.844 Å/pixel) for 4.4s, resulting in a total exposure dose of 60 e-/Å².

25,030 super-resolution videos were motion corrected, binned to physical pixel size and dose weighted, post acquisition, using UCSF MotionCor2. Contrast transfer function estimation was performed with the cryoSPARC “Patch CTF Estimation” job prior to automated particle picking with a gaussian template. After curating for high quality micrographs as previously described, 9,078,725 particles were extracted in a 512-pixel box and Fourier cropped to 128 pixels. Due to the size of the dataset, image processing was performed in three cohorts. In the first cohort, 2,978,754 particles were subjected to two rounds of 2D classification followed by *Ab initio* reconstruction of selected particles. Distinct conformations of the TSHR ECD in “up” and “side” states were apparent from 2D classification. Particles were selected for one round of 3D

classification without alignment irrespective of their ECD conformation, then split into a second separate round 3D classification in “up” or “side” classes only. Particles from the best looking and highest resolution “side” class were extracted in an unbinned 512 pixel box and subject to non-uniform refinement. The resulting reconstruction was of low quality, with uninterpretable 7TM domains that contoured at thresholds similar to the detergent micelle. As a result, particle images belonging to “side” class were not further processed. Particles classified into a high-quality ECD “up” class were extracted into a 512 pixel box and subject to non-uniform refinement. The remaining 16,159 micrographs were processed in a similar workflow of multiple rounds of 2D classification followed by 3D classification with alignment. Unbinned particles from each cohort in qualitatively good ECD “up” classes were subject to independent non-uniform refinement jobs. The three refined sets of particles were then combined and subject to unmasked, alignment-free 3D classification in RELION. All classes exhibited high quality 7TM domain features. Thus, particles in classes containing distinct ECD:M22 VH/VL density were selected, combined, and re-imported into cryoSPARC for 1 round of focused refinement using masked regions separately encompassing the TSHR ECD:M22 domain or the TSHR transmembrane:G protein complex. Directional FSC curves for each focused refinement reconstruction were calculated, and each map was loaded into ChimeraX for generation of a composite map via the *vop maximum* command.

Org 274179-0-bound TSHR complex

The Org 274179-0- bound TSHR- Δ Cpep complex was concentrated to 55 μ M and 3 μ L was plunge frozen on holey gold grids as previously for other TSHR samples described after a 30 second hold and a 3 second blotting time at a blotting force of 0. Cryo-EM data were collected on a 300 keV Titan Krios microscope equipped with a K3 detector and a BioQuantum post-column energy filter

with a slit width of 20 eV. 10,003 super-resolution movies were recorded at a nominal magnification of 105,000x (physical pixel size of 0.86 Å/pixel) for 1.35 s each and fractionated across 50 frames. The defocus range for this collection was -0.8 to -2.4 μm.

10,003 TSHR-ΔCpeptide-Org 274179-0 videos were motion corrected, binned to physical pixel size and dose weighted, post acquisition, using UCSF MotionCor2. CTF estimation and reference-free “blob” particle autopicking were performed in cryoSPARC. Micrographs were curated following similar thresholds as previously described prior to extraction of 5,922,556 particles in a 288-pixel box that was subsequently Fourier cropped to 72 pixels. To generate a suitable reference for 3D classification, 2 rounds of 2D classification were performed. From 2D classification, the particles belonging to classes that exhibited “micelle”-like density were then subjected for *Ab initio* 3D reconstruction. Then, 3D classification with alignment was performed on the entire set of initially picked particles using one class suggestive of a micelle:ECD reconstruction and three classes generated from prematurely truncated *Ab initio* jobs, as previously described. Particles classified into the micelle:ECD class were subject to two further rounds of the same classification workflow prior to being extracted without Fourier cropping. Finally, the 357,869 remaining particles were subject to non-uniform refinement. The reconstruction quality of the TSHR-Org 274179-0 complex was poor, and was not improved upon further rounds of unbinned 3D classification with, or without, alignment. 2D classification of the particles in the initial non-uniform refinement job revealed clear density for the ECD with more diffuse alignment on the 7TM domain.

CS-17-bound TSHR complex

3 μL of purified CS-17-bound TSHR- ΔCpep :Org 274179-0 complex was concentrated to 40.8 μM and similarly plunge frozen into liquid ethane after a 30 second hold, a blotting time of 2 seconds, and a blotting force of 0. The TSHR- $\Delta\text{Cpeptide}$:Org 274179-0 complex at 55 μM was plunge frozen on holey gold grids as previously described after a 30 second hold and a 3 second blotting time at a blotting force of 0.

9,244 super-resolution videos were recorded with identical collection parameters on the same Titan Krios microscope as described for the TSHR:Org 274179-0 complex.

Next, super-resolution videos of the TSHR- $\Delta\text{Cpeptide}$:CS-17 Fab fragment complex were motion corrected, dose-weighted, and binned on-the-fly with MotionCor2 during data collection, as previously described. CTF estimation, “blob” particle autopicking, and Fourier cropped particle extraction were all performed in cryoSPARC, after curating for high quality micrographs. 2,742,080 picked particles were subject to 4 rounds of 2D classification followed by *Ab initio* reconstruction to generate reference maps for 3D classification. Then, all picked particles were subject to 4 rounds of 3D classification with alignment, two rounds at a 4x-Fourier cropped box size of 112 pixels, and two rounds at a 2x-Fourier cropped box size. 114,707 particles were then subject to 3D classification with alignment using four identical input reference volumes. The shifts and poses of the 41,054 particles belonging to the highest resolution reconstruction from this final 3D classification were then refined using the non-uniform refinement job type. In an attempt to improve reconstruction quality in the TSHR- ΔCpep :CS-17 Fab binding interface, 2 rounds of focused refinement were performed. The first focused refinement utilized a mask encompassing

only the ECD and CS-17 variable domains and was followed by a refinement with the entire TSHR- Δ Cpep:CS-17 complex masked. Both focused refinements used the same rotation/shift prior restrictions as in the TSHR-G_s processing workflow. However, in comparison to the initial non-uniform refinement reconstruction, the focused refinements yielded lower quality 7TM domain reconstructions. As a result, the pre-focused refinement reconstruction half maps were used for directional FSC calculation and subsequent atomic model building.

Model Building and refinement

We first modeled the M22-bound TSHR-G_s complex in a composite cryo-EM map (**Figure 1.16**). For TSHR, we started with an AlphaFold⁶³ model of full-length human TSHR, which had high structural agreement with a previously determined X-ray crystal structure of the M22 Fab bound to the TSHR ECD (PDB ID: 3G04⁴⁵). After truncating unresolved regions, the 7TM domain in this model was fit into the composite cryo-EM map of M22-bound TSHR-G_s with ChimeraX v1.3⁶². This template model was rebuilt in Coot⁶⁴ for rigid body fitting of the extracellular domain. For the M22 Fab, we used the PDB ID: 3G04 as a starting structure. For the G protein (miniG_s, G β , and G γ) we used PDB ID: 7LJC⁶⁵ as a starting template. Finally, for Nb35, we used PDB ID: 3SN6¹⁸ as a starting template. For DPPC, restraints were generated using the Prodrgr server⁶⁶ and the aliphatic tails of the lipid were manually docked into cryo-EM density using Coot. Each of these components was individually fit into cryo-EM density with ChimeraX. We subsequently iteratively refined the model with manual refinement in Coot and ISOLDE⁶⁷ and real space refinement in Phenix⁶⁸.

To model TSH- and TR1402-bound TSHR, we started with the M22-activated structure described above and fit it into a composite map of the TSH-TSHR-G_s complex using ChimeraX (**Figure 1.7**). The fit of the TSHR ECD was further optimized by rigid body fitting in Coot and Phenix. A key distinction between the M22- and TSH/TR1402- activated structures of TSHR is in the hinge region. For TSH- and TR1402- activated TSHR, we were able to resolve residues 291-302 and 387-396 in the hinge domain. We first built a model for TR1402 bound to the TSHR ECD using AlphaFold2, which revealed a partially ordered hinge region consistent with our cryo-EM structure. This model was fit into the TR1402-TSHR-G_s complex composite map in ChimeraX, and then iteratively refined with manual changes in Coot and ISOLDE and real space refinement in Phenix. As the TSH-bound TSHR complex was of lower resolution in the extracellular region, we used the TR1402 structure as a starting point and iteratively refined TSHR ECD and TSH using the composite map of the TSH-TSHR-G_s complex.

For the CS-17/Org 274179-0 bound TSHR complex, we started with AlphaFold2 predictions for the TSHR and for the CS-17 Fab which were fit into the cryo-EM map with ChimeraX. The TSHR ECD was fit into the cryo-EM density by rigid body refinement using Coot. Similarly, the constant regions of the CS-17 Fab were fit into the cryo-EM density by rigid body refinement in Coot. The resulting model was refined in ISOLDE with manual changes in both ISOLDE and Coot followed by real space refinement in Phenix. All maps and models were validated using MolProbity⁶⁹.

Modeling TSH, M22, and K1-70 binding poses on inactive state TSHR (as in **Figure 1.4**) was performed in ChimeraX with the matchmaker command to align the active state ECD residues C29-F405 to the inactive state structure as a reference model. The same operation was performed

for modeling CS-17 onto the active state TSHR ECD. PDB files for each OPM alignment are available upon request.

Identification of lipid in M22-bound TSHR-G_s

Lipids were extracted using a modified version of the Bligh-Dyer method⁷⁰. Briefly, samples of PC(16:0/16:0) (DPPC) for the standard curve were manually shaken for 30 s in a glass vial (VWR) with 1 mL PBS, 1 mL methanol and 2 mL chloroform containing the internal standard Cer(d18:1-d7/18:1). The resulting mixtures were vortexed for 15 s and centrifuged at 2400 x g for 6 min to induce phase separation. The organic (bottom) layer was retrieved using a glass pipette, dried under a gentle stream of nitrogen, and reconstituted in 2:1 chloroform:methanol for LC/MS analysis. Targeted lipidomic analysis was performed on a Dionex Ultimate 3000 LC system (Thermo) coupled to a TSQ Quantiva mass spectrometer (Thermo). Data was acquired in positive ionization mode. Solvent A consisted of 95:5 water:methanol, Solvent B was 70:25:5 isopropanol:methanol:water. Solvents A and B contained 5 mM ammonium formate with 0.1% formic acid. A XBridge (Waters) C8 column (5 μ m, 4.6 mm \times 50 mm) was used. The gradient was held at 0% B between 0 and 5 min, raised to 20% B at 5.1 min, increased linearly from 20% to 100% B between 5.1 and 35 min, held at 100% B between 35 and 40 min, returned to 0% B at 40.1 min, and held at 0% B until 50 min. Flow rate was 0.1 mL/min from 0 to 5 min, 0.3 mL/min from 5.1 to 50 min. MS analyses were performed using electrospray ionization in positive ion mode, with spray voltages of 3.5 kV, ion transfer tube temperature of 325 C, and vaporizer temperature of 200 C. Sheath, auxiliary, and sweep gases were 40, 10 and 1, respectively. Internal standard Cer(d18:1-d7/18:1) was detected using the following transitions: 571.6>271.2, 571.6>289.3, 571.6>259.3. DPPC was detected using the following transition: 734.6>184.1, 734.6>125.1,

734.6>86.1. Other PCs were detected using the specific parent ion m/z and the same diagnostic fragments as DPPC. Chromatography and peak integration of lipid targets were verified with Skyline⁷¹. Peak areas were used in data reporting, data was normalized using internal standards.

Native mass spectrometry (nMS) of K1-70 fab fragments

Purified Fab fragments were prepared for native mass spectrometry (nMS) by buffer exchange into 200 mM ammonium acetate with a Zeba spin desalting column. Samples, normalized to 5 μ M, were injected directly in a Q Exactive Extended Mass Range mass spectrometer using nanoES electrospray capillaries. The instrument parameters were: capillary voltage 1.1 kV; S-lens RF 100%; quadrupole selection 300-10,000 m/z , collisional activation in the HCD cell 150 °C, trapping gas pressure setting 8; temperature 250 °C, instrument resolution 35,000. Data was analyzed with UniDec 4.2.2⁷².

Mass spectrometry of human TSH

Native human TSH (National Hormone and Pituitary Program) was resuspended in water, reduced with 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP, Sigma-Aldrich) and alkylated with 10 mM chloroacetamide (Sigma-Aldrich). Protein was digested at 37°C in 100 mM Triethylammonium bicarbonate buffer, pH 8.5, with trypsin (Promega) for 1 hour. Digestion was quenched with formic acid, 5% (v/v) final concentration.

The digested TSH was analyzed on a Orbitrap Eclipse mass spectrometer (Thermo). The digest was injected directly onto a 25 cm, 100 μ m ID column packed with BEH 1.7 μ m C18 resin (Waters). The sample was separated at a flow rate of 300 nl/min on an Easy nLC 1200 (Thermo). Buffer A and B were 0.1% formic acid in water and 0.1% formic acid in 90% acetonitrile,

respectively. A gradient of 1-25% B over 100 min, an increase to 40% B over 20 min, an increase to 90% B over 10 min and held at 90% B for a final 10 min was used for 140 min total run time. Column was re-equilibrated with 15 ul of buffer A prior to the injection of sample. Peptides were eluted directly from the tip of the column and nanosprayed directly into the mass spectrometer by application of 2.5 kV voltage at the back of the column. The Eclipse was operated in a data dependent mode. Full MS scans were collected in the Orbitrap at 120k resolution with a mass range of 350 to 1800 m/z and an AGC target of $4e^5$. The cycle time was set to 3 sec, and within this 3 sec the most abundant ions per scan were selected for HCD MS/MS in the Orbitrap with an AGC target of $5e^4$, resolution of 30k and minimum intensity of 50000. Maximum fill times were set to 50 ms and 54 ms for MS and MS/MS scans respectively. Quadrupole isolation at 2 m/z was used, monoisotopic precursor selection was enabled and dynamic exclusion was used with exclusion duration of 60 sec. Protein and peptide identification were done with Byonic node (Protein Metrics Inc) within Proteome Discoverer 2.5 (Thermo).

G_s signaling assays

To measure ligand-dependent activation of G_s by TSHR we measured cAMP production. For each TSHR construct (wild-type, ΔC-peptide, K262C, N483C, K262C-N483C, Y385F, Y385A, E409A, I486F, A644K, A647K), a 20 mL suspension culture of Expi293 cells was co-transfected with a pCDNA3.1 plasmid expressing TSHR and a luciferase-based cAMP biosensor (pGlosensor-22F; Promega). DNA quantities for each construct were titrated and resultant receptor expression levels were determined by flow cytometry using an Alexa647-conjugated anti-FLAG antibody. Transfection quantities of all constructs resulting in similar expression levels to WT TSHR were selected, and were co-transfected with a fixed 3:1 DNA mass ratio of receptor construct to cAMP

biosensor as used for WT TSHR transfections. Cells were harvested 24 h post-transfection, resuspended in Expi293 expression media (Gibco) supplemented with 10% DMSO, and gradually frozen to -80°C in a Mr. Frosty Freezing container for future use. To perform the assay, frozen Expi293 cells were rapidly thawed in a 37°C water bath and resuspended in fresh Expi293 expression medium. Cells were diluted to a final concentration of 500,000 cells/mL in Expi293 expression medium plus 2% (v/v) Glosensor assay reagent (Promega) and incubated for 75 minutes at room temperature with gentle rotation. Expi293 cells were then plated into a white 384-well plate (Greiner) to a final density of 10,000 cells per well. Separately, a 5x ligand stock plate (13-point half-log dilution series, 3-wells per condition) was made in Hank's balanced salt solution + 0.1% (w/v) bovine serum albumin. The ligand plate was stamped into new 384-well plates to achieve a final 1X concentration upon addition of Glosensor-reagent incubated, co-transfected TSHR:cAMP biosensor cells. Immediately after cell addition, luminescence was measured in 0.1 s intervals for 10 minutes using a CLARIOstar instrument at an emission wavelength of 580 nm ± 80 nm band-pass (BMG LabTech). The resulting dose response curves from the 5-minute read point were fit to a nonlinear regression three-parameter log(agonist) vs. response fit in GraphPad Prism 9 (Dotmatics).

Molecular dynamics simulations

Simulation setup

We performed simulations of four distinct conditions, using the M22-bound active-state TSHR and the CS17-bound inactive-state TSHR structures as starting points. The four conditions are the following: (A) from active, with M22 bound; (B) from active, with M22 removed; (C) from

inactive, with CS17 bound; (D) from inactive, with CS17 removed. For each simulation condition, we performed 10 independent 1- μ s simulations.

For conditions A and B, the lipid in the transmembrane binding pocket was modeled as DPPC in Maestro (Schrödinger), and the heterotrimeric G protein, Nb35, and N-acetylglucosamine (NAG) residues were removed from the structure. Missing loops and amino acid side chains were modeled using Prime (Schrödinger). We did not model the \sim 100 residue TSHR hinge region. Neutral acetyl and methylamide groups were added to cap the N- and C-termini, respectively, of the protein chains. Titratable residues were kept in their dominant protonation state at pH 7, except for residues E506^{3,37} and D633^{6,44}, which were neutral. Histidine residues were modeled as neutral, with a hydrogen atom bound either to the epsilon nitrogen or the delta nitrogen, depending on which optimized the local hydrogen-bonding network upon visual inspection. Dowser⁷³ was used to add water molecules to protein cavities. We used the OPM PPM webserver⁴⁸ to align the proteins, after which the aligned structures were inserted into a pre-equilibrated palmitoylcholine (POPC) membrane bilayer using Dabble⁷⁴. Sodium and chloride ions were added to neutralize each system at a concentration of 150 mM.

The final system of condition A comprised 308,127 atoms, including 592 lipid molecules and 70,982 water molecules (initial system dimensions: 150 Å x 150 Å x 135 Å). The final system of condition B comprised 131,486 atoms, including 246 lipid molecules and 29,762 water molecules (initial system dimensions: 100 Å x 100 Å x 130 Å). The final system of condition C comprised 306,262 atoms, including 576 lipid molecules and 71,057 water molecules (initial system dimensions: 150 Å x 150 Å x 135 Å). The final system of condition D comprised 131,806 atoms,

including 245 lipid molecules and 29,907 water molecules (initial system dimensions: 100 Å x 100 Å x 130 Å).

Simulation protocols

For each simulation, initial atom velocities were assigned randomly and independently. We employed the CHARMM36m force field for protein molecules, the CHARMM36 parameter set for lipid molecules and salt ions, and the associated CHARMM TIP3P model for water^{75,76}. Simulations were run using the AMBER20 software⁷⁷ under periodic boundary conditions with the Compute Unified Device Architecture (CUDA) version of Particle-Mesh Ewald Molecular Dynamics (PMEMD) on one GPU⁷⁸.

After energy minimization, the systems were first heated over 12.5 ps from 0 K to 100 K in the NVT ensemble using a Langevin thermostat with harmonic restraints of $10.0 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ on the non-hydrogen atoms of the lipids, protein, and ligand. Initial velocities were sampled from a Boltzmann distribution. The systems were then heated to 310 K over 125 ps in the NPT ensemble. Equilibration was performed at 310 K and 1 bar in the NPT ensemble, with harmonic restraints on the protein and ligand non-hydrogen atoms tapered off by $1.0 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ starting at $5.0 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ in a stepwise manner every 2 ns for 10 ns, and finally by $0.1 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ every 2 ns for an additional 18 ns. All restraints were completely removed during production simulation. Production simulations were performed at 310 K and 1 bar in the NPT ensemble using the Langevin thermostat and Monte Carlo barostat. Lengths of bonds to hydrogen atoms were constrained using SHAKE, and the simulations were performed using a timestep of 4.0 fs while using hydrogen mass repartitioning⁷⁹. Non-bonded interactions were cut off at 9.0 Å, and long-

range electrostatic interactions were calculated using the particle-mesh Ewald (PME) method with an Ewald coefficient (β) of approximately 0.31 Å and B-spline interpolation of order 4. The PME grid size was chosen such that the width of a grid cell was approximately 1 Å. Snapshots of each trajectory were saved every 200 ps.

Simulation analysis protocols

The AmberTools17 CPPTRAJ package⁸⁰ was used to reimage trajectories at 1 ns per frame. Visual Molecular Dynamics (VMD)⁸¹ was used for visualization and analysis.

For the analysis of ECD conformation shown in **Figure 1.4** and **Figure 1.17**, we performed a projection analysis as described in Latorraca, N. et al.⁸². Briefly, we monitor ECD conformation by representing the C α atoms of ECD residues 30–260 as a single vector containing the Cartesian coordinates. After aligning on transmembrane (TM) helices TM1–TM4 of the receptor (C α atoms of residues 415–441, 448–476, 492–525, and 536–558), we project this vector in a given simulation frame onto the line connecting the experimentally determined inactive ECD conformation (from the CS17-bound structure reported in this paper) and experimentally determined active ECD conformation (from the M22-bound structure reported in this paper). We then report the position of the projected point on the line, using the convention that the inactive ECD conformation is at 0 and positive values indicate change toward the active conformation. The experimentally determined active ECD conformation is assigned a value equal to its RMSD from the inactive conformation (in this case, 33.5Å).

1.9 Acknowledgements

We thank Rui Yan at the HHMI Janelia CryoEM Facility and Dan Toso at Cal-Cryo at QB3-Berkeley for help in microscope operation and data collection, Cole Bracken for mass spectrometry troubleshooting, and Hazel Shan for sample preparation for lipid mass spectrometry. This work was supported by National Institutes of Health (NIH) grants DP5OD023048 (A.M.), 1R35GM140847 (Y.C.), P30CA014195, R01GM102491 (A.S.) and a Human Frontier Science Program Long-Term Fellowship LT000916/2018-L (C.-M.S.). Cryo-EM equipment at UCSF is partially supported by NIH grants S10OD020054 and S10OD021741. Some of this work was performed at the Stanford-SLAC Cryo-EM Center (S2C2), which is supported by the National Institutes of Health Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (U24 GM129541). This research was, in part, supported by the National Cancer Institute's National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E. Some of this work was supported by the Mass Spectrometry Core of the Salk Institute with funding from NIH-NCI CCSG: P30 014195 and the Helmsley Center for Genomic Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Y.C. is an Investigator of Howard Hughes Medical Institute. A.M. acknowledges support from the Pew Charitable Trusts, the Esther and A. & Joseph Klingenstein Fund and the Searle Scholars Program.

1.10 Author Contributions

B.F. cloned, expressed, and biochemically optimized the purification of all TSHR constructs for structural studies. B.F. expressed and purified G β ₁ γ ₂, Nb35, M22 Fab, CS-17 IgG, and K1-70 IgGs, and performed enzymatic digestion and further purification to form CS-17 and K1-70 Fab

fragments. B.F. performed complexing and identified optimal cryo-EM grid preparation procedures, screened samples, and collected 300 keV datasets. B.F. determined high resolution cryo-EM maps by extensive image processing under the guidance of A.M. and Y.C. B.F. and A.M. built and refined models of TSHR complexes. B.F. and C.B.B. generated receptor constructs and determined expression levels by flow cytometry and performed and analyzed data from signaling studies. C.B.B. assisted with cloning of TSHR constructs, expression, purification, and labeling of CS-17 and M22 Fab fragments, cloning and generation of baculoviruses for expression of G β γ ₂. C.-M.S. performed and analyzed MD simulations under the supervision of R.O.D. I.S. performed initial biochemical optimization of TSHR complexes with M22 Fab and worked with K.Z. to collect negative stain and cryo-EM data. N.H. prepared samples for, performed, and analyzed native mass spectrometry experiments and prepared control samples for lipid identification experiments. J.K.D. analyzed native human TSH glycosylation by mass spectrometry. Y.M. generated TSHR and M22 Fab expression constructs and performed pilot biochemical purification of TSHR complexes. A.M.P. performed and analyzed data from the lipid identification experiments with guidance from A.S. Figures were generated and the manuscript written by B.F, C.B.B., and A.M., with edits from Y.C. and with approval from all authors. The overall project was supervised by Y.C. and A.M.

1.11 References

1. Molecular Basis of Thyroid Hormone Action. (1983) doi:10.1016/b978-0-12-527560-6.x5001-x.
2. Vassart, G. & Dumont, J. E. The thyrotropin receptor and the regulation of thyrocyte function and growth. *Endocr. Rev.* 13, (1992).
3. Smith, T. J. & Hegedüs, L. Graves' Disease. *N. Engl. J. Med.* 375, 1552–1565 (2016).
4. Fekete, C. & Lechan, R. M. Central regulation of hypothalamic-pituitary-thyroid axis under physiological and pathophysiological conditions. *Endocr. Rev.* 35, 159–194 (2014).
5. Laurent, E., Mockel, J., Van Sande, J., Graff, I. & Dumont, J. E. Dual activation by thyrotropin of the phospholipase C and cyclic AMP cascades in human thyroid. *Mol. Cell. Endocrinol.* 52, 273–278 (1987).
6. Taylor, P. N. et al. Global epidemiology of hyperthyroidism and hypothyroidism. *Nat. Rev. Endocrinol.* 14, 301–316 (2018).
7. Zimmermann, M. B. & Boelaert, K. Iodine deficiency and thyroid disorders. *Lancet Diabetes Endocrinol* 3, 286–295 (2015).
8. Mincer, D. L. & Jialal, I. Hashimoto Thyroiditis. in *StatPearls* (StatPearls Publishing, 2021).
9. Lane, L. C., Cheetham, T. D., Perros, P. & Pearce, S. H. S. New Therapeutic Horizons for Graves' Hyperthyroidism. *Endocr. Rev.* 41, (2020).
10. Flack, M. R., Froehlich, J., Bennet, A. P., Anasti, J. & Nisula, B. C. Site-directed mutagenesis

- defines the individual roles of the glycosylation sites on follicle-stimulating hormone. *J. Biol. Chem.* 269, 14015–14020 (1994).
11. Matzuk, M. M., Keene, J. L. & Boime, I. Site specificity of the chorionic gonadotropin N-linked oligosaccharides in signal transduction. *J. Biol. Chem.* 264, 2409–2414 (1989).
 12. Grossmann, M., Weintraub, B. D. & Szekudlinski, M. W. Novel insights into the molecular mechanisms of human thyrotropin action: structural, physiological, and therapeutic implications for the glycoprotein hormone family. *Endocr. Rev.* 18, 476–501 (1997).
 13. Jiang, X. et al. Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. *Proc. Natl. Acad. Sci. U. S. A.* 109, 12491–12496 (2012).
 14. Duan, J. et al. Structures of full-length glycoprotein hormone receptor signalling complexes. *Nature* (2021) doi:10.1038/s41586-021-03924-2.
 15. Chazenbalk, G. D. et al. Evidence that the thyrotropin receptor ectodomain contains not one, but two, cleavage sites. *Endocrinology* 138, 2893–2899 (1997).
 16. Chen, C.-R., Salazar, L. M., McLachlan, S. M. & Rapoport, B. Deleting the Redundant TSH Receptor C-Peptide Region Permits Generation of the Conformationally Intact Extracellular Domain by Insect Cells. *Endocrinology* 156, 2732–2738 (2015).
 17. Nehmé, R. et al. Mini-G proteins: Novel tools for studying GPCRs in their active conformation. *PLoS One* 12, e0175642 (2017).
 18. Rasmussen, S. G. F. et al. Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature* 477, 549–555 (2011).

19. Szkudlinski, M. W., Teh, N. G., Grossmann, M., Tropea, J. E. & Weintraub, B. D. Engineering human glycoprotein hormone superactive analogues. *Nat. Biotechnol.* 14, 1257–1263 (1996).
20. Reinfelder, J. et al. Effects of recombinant human thyroid-stimulating hormone superagonists on thyroidal uptake of ¹⁸F-fluorodeoxyglucose and radioiodide. *Thyroid* 21, 783–792 (2011).
21. Ulloa-Aguirre, A., Timossi, C., Damián-Matsumura, P. & Dias, J. A. Role of glycosylation in function of follicle-stimulating hormone. *Endocrine* 11, 205–215 (1999).
22. Fan, Q. R. & Hendrickson, W. A. Structure of human follicle-stimulating hormone in complex with its receptor. *Nature* 433, 269–277 (2005).
23. Costagliola, S. et al. Tyrosine sulfation is required for agonist recognition by glycoprotein hormone receptors. *EMBO J.* 21, 504 (2002).
24. Kosugi, S., Ban, T., Akamizu, T. & Kohn, L. D. Site-directed mutagenesis of a portion of the extracellular domain of the rat thyrotropin receptor important in autoimmune thyroid disease and nonhomologous with gonadotropin receptors. Relationship of functional and immunogenic domains. *J. Biol. Chem.* 266, 19413–19418 (1991).
25. Caltabiano, G. et al. The specificity of binding of glycoprotein hormones to their receptors. *Cell. Mol. Life Sci.* 65, 2484–2492 (2008).
26. Moyle, W. R. et al. Co-evolution of ligand-receptor pairs. *Nature* 368, 251–255 (1994).
27. Laphorn, A. J. et al. Crystal structure of human chorionic gonadotropin. *Nature* 369, 455–461 (1994).

28. Wu, H., Lustbader, J. W., Liu, Y., Canfield, R. E. & Hendrickson, W. A. Structure of human chorionic gonadotropin at 2.6 Å resolution from MAD analysis of the selenomethionyl protein. *Structure* 2, 545–558 (1994).
29. Grossmann, M. et al. Substitution of the seat-belt region of the thyroid-stimulating hormone (TSH) beta-subunit with the corresponding regions of choriogonadotropin or follitropin confers luteotropic but not follitropic activity to chimeric TSH. *J. Biol. Chem.* 272, 15532–15540 (1997).
30. Dias, J. A., Zhang, Y. & Liu, X. Receptor binding and functional properties of chimeric human follitropin prepared by an exchange between a small hydrophilic intercysteine loop of human follitropin and human lutropin. *J. Biol. Chem.* 269, 25289–25294 (1994).
31. Vischer, H. F., Granneman, J. C. M. & Bogerd, J. Opposite Contribution of Two Ligand-Selective Determinants in the N-Terminal Hormone-Binding Exodomain of Human Gonadotropin Receptors. *Mol. Endocrinol.* 17, 1972–1981 (2003).
32. Chen, F., Wang, Y. & Puett, D. Role of the invariant aspartic acid 99 of human choriogonadotropin beta in receptor binding and biological activity. *J. Biol. Chem.* 266, 19357–19361 (1991).
33. Smits, G. et al. Glycoprotein hormone receptors: determinants in leucine-rich repeats responsible for ligand specificity. *EMBO J.* 22, 2692–2703 (2003).
34. van Koppen, C. J. et al. Mechanism of action of a nanomolar potent, allosteric antagonist of the thyroid-stimulating hormone receptor. *Br. J. Pharmacol.* 165, 2314–2324 (2012).
35. Van Sande, J. et al. In Chinese hamster ovary K1 cells dog and human thyrotropin receptors

- activate both the cyclic AMP and the phosphatidylinositol 4,5-bisphosphate cascades in the presence of thyrotropin and the cyclic AMP cascade in its absence. *Eur. J. Biochem.* 229, 338–343 (1995).
36. Chen, C.-R., McLachlan, S. M. & Rapoport, B. A monoclonal antibody with thyrotropin (TSH) receptor inverse agonist and TSH antagonist activities binds to the receptor hinge region as well as to the leucine-rich domain. *Endocrinology* 150, 3401–3408 (2009).
 37. Mueller, S., Jaeschke, H., Günther, R. & Paschke, R. The hinge region: an important receptor component for GPCR function. *Trends Endocrinol. Metab.* 21, 111–122 (2010).
 38. Mizutori, Y., Chen, C.-R., McLachlan, S. M. & Rapoport, B. The Thyrotropin Receptor Hinge Region Is Not Simply a Scaffold for the Leucine-Rich Domain but Contributes to Ligand Binding and Signal Transduction. *Molecular Endocrinology* vol. 22 1171–1182 (2008).
 39. Brüser, A. et al. The Activation Mechanism of Glycoprotein Hormone Receptors with Implications in the Cause and Therapy of Endocrine Diseases. *J. Biol. Chem.* 291, 508 (2016).
 40. Schulze, A. et al. The intramolecular agonist is obligate for activation of glycoprotein hormone receptors. *FASEB J.* 34, 11243–11256 (2020).
 41. Parma, J. et al. Somatic mutations causing constitutive activity of the thyrotropin receptor are the major cause of hyperfunctioning thyroid adenomas: identification of additional mutations activating both the cyclic adenosine 3',5'-monophosphate and inositol phosphate-Ca²⁺ cascades. *Mol. Endocrinol.* 9, 725–733 (1995).
 42. Vlaeminck-Guillem, V., Ho, S.-C., Rodien, P., Vassart, G. & Costagliola, S. Activation of the cAMP Pathway by the TSH Receptor Involves Switching of the Ectodomain from a Tethered

- Inverse Agonist to an Agonist. *Mol. Endocrinol.* 16, 736–746 (2002).
43. Human monoclonal thyroid stimulating autoantibody. *Lancet* 362, 126–128 (2003).
 44. Sanders, J. et al. Crystal structure of the TSH receptor in complex with a thyroid-stimulating autoantibody. *Thyroid* 17, 395–410 (2007).
 45. Evans, M. et al. Monoclonal autoantibodies to the TSH receptor, one with stimulating activity and one with blocking activity, obtained from the same blood sample. *Clin. Endocrinol.* 73, 404–412 (2010).
 46. Sanders, P. et al. Crystal structure of the TSH receptor (TSHR) bound to a blocking-type TSHR autoantibody. *J. Mol. Endocrinol.* 46, 81–99 (2011).
 47. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* 40, D370–6 (2012).
 48. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* 15, 346–366 (2019).
 49. Erbel, P. J. A., Haseley, S. R., Kamerling, J. P. & Vliegthart, J. F. G. Studies on the relevance of the glycan at Asn-52 of the alpha-subunit of human chorionic gonadotropin in the alphabeta dimer. *Biochem. J* 364, 485–495 (2002).
 50. Amr, S. et al. Activities of deglycosylated thyrotropin at the thyroid membrane receptor-adenylate cyclase system. *J. Endocrinol. Invest.* 8, 537–541 (2014).
 51. Fares, F. A., Levi, F., Reznick, A. Z. & Kraiem, Z. Engineering a potential antagonist of

- human thyrotropin and thyroid-stimulating antibody. *J. Biol. Chem.* 276, 4543–4548 (2001).
52. Feng, X., Müller, T., Mizrachi, D., Fanelli, F. & Segaloff, D. L. An intracellular loop (IL2) residue confers different basal constitutive activities to the human lutropin receptor and human thyrotropin receptor through structural communication between IL2 and helix 6, via helix 3. *Endocrinology* 149, 1705–1717 (2008).
53. Zhang, M. et al. Intrinsic differences in the response of the human lutropin receptor versus the human follitropin receptor to activating mutations. *J. Biol. Chem.* 282, 25527–25539 (2007).
54. Skiba, M. A. & Kruse, A. C. Autoantibodies as Endogenous Modulators of GPCR Signaling. *Trends Pharmacol. Sci.* 42, 135–150 (2021).
55. Girdlestone, C. & Hayward, S. The DynDom3D Webserver for the Analysis of Domain Movements in Multimeric Proteins. *J. Comput. Biol.* 23, 21–26 (2016).
56. Sanders, J. et al. Characteristics of a human monoclonal autoantibody to the thyrotropin receptor: sequence structure and function. *Thyroid* 14, 560–570 (2004).
57. Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microsc. Microanal.* 9, 1182–1183 (2003).
58. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332 (2017).
59. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290–296 (2017).

60. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* 17, 1214–1221 (2020).
61. Scheres, S. H. W. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* 415, 406–418 (2012).
62. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82 (2021).
63. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021).
64. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132 (2004).
65. Zhuang, Y. et al. Mechanism of dopamine binding and allosteric modulation of the human D1 dopamine receptor. *Cell Res.* 31, 593–596 (2021).
66. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* 60, 1355–1363 (2004).
67. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* 74, 519–530 (2018).
68. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221 (2010).
69. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular

- crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66, 12–21 (2010).
70. Bligh, E. G. A Rapid Method of Total Lipid Extraction and Purification. (1959).
 71. MacLean, B. et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966–968 (2010).
 72. Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* 87, 4370–4376 (2015).
 73. Zhang, L. & Hermans, J. Hydrophilicity of cavities in proteins. *Proteins: Structure, Function, and Genetics* vol. 24 433–438 (1996).
 74. Betz, R. Dabble. (2017) doi:10.5281/zenodo.836914.
 75. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14, 71–73 (2017).
 76. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* 114, 7830–7843 (2010).
 77. Case, D. A. et al. Amber 2021. (University of California, San Francisco, 2021).
 78. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* 9, 3878–3888 (2013).
 79. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* 11, 1864–1874

(2015).

80. Roe, D. R. & Cheatham, T. E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 9, 3084–3095 (2013).
81. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–8, 27–8 (1996).
82. Latorraca, N. R. et al. How GPCR Phosphorylation Patterns Orchestrate Arrestin-Mediated Signaling. *Cell* 183, 1813–1825.e18 (2020).
83. Asarnow, D., Palovcak, E. & Cheng, Y. UCSF pyem v0. 5. Zenodo <https://doi.org/10.5281/zenodo.3576630>, (2019).

Chapter 2

An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike

Contributing Authors

Michael Schoof^{1,2*†}, Bryan Faust^{1,2,3,4*}, Reuben A. Saunders^{1,5*}, Smriti Sangwan^{1,2*}, Veronica Rezelj^{6*}, Nick Hoppe^{3,4}, Morgane Boone^{1,2}, Christian B. Billesbølle^{3,4}, Cristina Puchades⁴, Caleigh M. Azumaya⁴, Huong T. Kratochvil⁴, Marcell Zimanyi^{1,2}, Ishan Deshpande^{3,4}, Jiahao Liang³, Sasha Dickinson⁴, Henry C. Nguyen⁴, Cynthia M. Chio⁴, Gregory E. Merz⁴, Michael C. Thompson⁴, Devan Diwanji⁴, Kaitlin Schaefer⁴, Aditya A. Anand^{1,2}, Niv Dobzinski^{1,2}, Beth Shoshana Zha⁷, Camille R. Simoneau^{8,9,10}, Kristoffer Leon^{8,9,10}, Kris M. White^{11,12}, Un Seng Chio⁴, Meghna Gupta⁴, Mingliang Jin⁴, Fei Li⁴, Yanxin Liu⁴, Kaihua Zhang⁴, David Bulkley⁴, Ming Sun⁴, Amber M. Smith⁴, Alexandra N. Rizo⁴, Frank Moss⁴, Axel F. Brilot⁴, Sergei Pourmal⁴, Raphael Trenker⁴, Thomas Pospiech⁴, Sayan Gupta¹³, Benjamin Barsi-Rhyne³, Vladislav Belyy^{1,2}, Andrew W. Barile-Hill¹⁴, Silke Nock^{1,2}, Yuwei Liu^{1,2}, Nevan J. Krogan^{4,5,8,9}, Corie Y. Ralston¹³, Danielle L. Swaney^{4,5,8,9}, Adolfo García-Sastre^{11,12,15,16}, Melanie Ott^{8,9,10}, Marco Vignuzzi⁶, QCRG Structural Biology Consortium^{4‡}, Peter Walter^{1,2†}, Aashish Manglik^{3,4,8,17†}

1. Howard Hughes Medical Institute, University of California at San Francisco, San Francisco, CA, USA.
2. Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, CA, USA.
3. Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA, USA.

4. Quantitative Biosciences Institute (QBI) Coronavirus Research Group Structural Biology Consortium, University of California, San Francisco, CA, USA.
5. Department of Cellular and Molecular Pharmacology, University of California at San Francisco, San Francisco, CA, USA.
6. Viral Populations and Pathogenesis Unit, CNRS UMR 3569, Institut Pasteur, 75724 Paris Cedex 15, France.
7. Department of Pulmonary, Critical Care, Allergy and Sleep Medicine, University of California San Francisco, San Francisco, CA, USA.
8. Quantitative Biosciences Institute (QBI), University of California San Francisco, San Francisco, CA, USA.
9. J. David Gladstone Institutes, San Francisco, CA, USA.
10. Department of Medicine, University of California San Francisco, San Francisco, CA, USA.
11. Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
12. Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
13. Molecular Biophysics and Integrated Bioimaging and the Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.
14. Cytiva Life Sciences, Marlborough, MA, USA.
15. Department of Medicine, Division of Infectious Diseases, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
16. The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
17. Department of Anesthesia and Perioperative Care, University of California at San Francisco, San Francisco, CA, USA.

*These authors contributed equally to this work.

2.1 Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus enters host cells via an interaction between its Spike protein and the host cell receptor angiotensin-converting enzyme 2 (ACE2). By screening a yeast surface-displayed library of synthetic nanobody sequences, we developed nanobodies that disrupt the interaction between Spike and ACE2. Cryo-electron microscopy (cryo-EM) revealed that one nanobody, Nb6, binds Spike in a fully inactive conformation with its receptor binding domains locked into their inaccessible down state, incapable of binding ACE2. Affinity maturation and structure-guided design of multivalency yielded a trivalent nanobody, mNb6-tri, with femtomolar affinity for Spike and picomolar neutralization of SARS-CoV-2 infection. mNb6-tri retains function after aerosolization, lyophilization, and heat treatment, which enables aerosol-mediated delivery of this potent neutralizer directly to the airway epithelia.

2.2 Introduction

Over the past two decades, three zoonotic β -coronaviruses have entered the human population, causing severe respiratory symptoms with high mortality^{1,2,3}. The COVID-19 pandemic is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the most readily transmissible of these three coronaviruses^{4,5,6,7}. No preventive treatment has been approved for any coronavirus to date, and the timeline for an effective and broadly available vaccine for SARS-CoV-2 remains uncertain. The development of new therapeutic and prophylactic approaches thus remains essential. Coronavirus virions are bounded by a membrane that contains the homotrimeric transmembrane glycoprotein Spike, which is responsible for virus entry into the host cell^{8,9}. The surface-exposed portion of Spike is composed of two domains, S_1 and S_2 ¹⁰. S_1 binds the host cell

receptor angiotensin-converting enzyme 2 (ACE2), whereas S₂ catalyzes fusion of the viral and host cell membranes^{11,12,13}. Contained within S₁ is the receptor binding domain (RBD), which directly binds to ACE2, and the N-terminal domain (NTD). The RBD is attached to the body of Spike by a flexible region and can exist in an inaccessible down state or an accessible up state¹⁴.¹⁵ Binding to ACE2 requires the RBD to occupy the up state and enables cleavage by host proteases, triggering a conformational change in S₂ required for viral entry¹⁶. In SARS-CoV-2 virions, Spike exchanges between an active, open conformation with at least one RBD in the up state and an inactive, closed conformation with all RBDs in the down state^{8,9}.

2.3 Results

We isolated single-domain antibodies (nanobodies) that neutralize SARS-CoV-2 by screening a yeast surface-displayed library of $>2 \times 10^9$ synthetic nanobody sequences for binders to the Spike ectodomain¹⁷. We used a mutant form of SARS-CoV-2 Spike (Spike^{S2P}) as the antigen¹⁵. Spike^{S2P} lacks one of the two proteolytic cleavage sites between the S₁ and S₂ domains and introduces two mutations and a trimerization domain to stabilize the prefusion conformation. We labeled Spike^{S2P} with biotin or with fluorescent dyes and selected nanobody-displaying yeast over multiple rounds, first by magnetic bead binding and then by fluorescence-activated cell sorting (**Figure 2.1**).

Three rounds of selection yielded 21 distinct nanobodies that bound Spike^{S2P} and showed decreased binding in the presence of a dimeric construct of the ACE2 extracellular domain (ACE2-Fc). These nanobodies fall into two classes. Class I binds the RBD and competes directly with ACE2-Fc (**Figure 2.1**). A prototypical example of this class is nanobody Nb6, which binds to Spike^{S2P} and to RBD alone with a dissociation constant (K_D) of 210 and 41 nM, respectively

(**Figure 2.1** and **Table 2.1**). Class II, exemplified by nanobody Nb3, binds to Spike^{S2P} ($K_D = 61$ nM) but displays no binding to RBD alone (**Figure 2.1** and **Table 2.1**). In the presence of excess ACE2-Fc, binding of Nb6 and other class I nanobodies is blocked entirely, whereas binding of Nb3 and other class II nanobodies is moderately decreased (**Figure 2.1**). These results suggest that class I nanobodies target the RBD to block ACE2 binding, whereas class II nanobodies target other epitopes. Indeed, surface plasmon resonance (SPR) experiments demonstrate that class I and class II nanobodies can bind Spike^{S2P} simultaneously (**Figure 2.1**).

Class I nanobodies show a consistently faster association rate constant (k_a) for nanobody binding to the isolated RBD than to Spike^{S2P} (**Table 2.1**), which suggests that RBD accessibility influences the K_D . We next tested the efficacy of class I and class II nanobodies to inhibit binding of fluorescently labeled Spike^{S2P} to ACE2-expressing human embryonic kidney (HEK) 293 cells (**Figure 2.1** and **Table 2.1**). Class I nanobodies Nb6 and Nb11 emerged as two of the most potent clones, with half-maximal inhibitory concentration (IC_{50}) values of 370 and 540 nM, respectively. Class II nanobodies showed little to no activity in this assay. We prioritized two class I nanobodies, Nb6 and Nb11, that combine potent Spike^{S2P} binding with relatively small differences in k_a between binding to Spike^{S2P} or RBD. For class II nanobodies, we prioritized Nb3 because of its relative yield during purification (**Table 2.1**).

To define the binding sites of Nb6 and Nb11, we determined their cryo-electron microscopy (cryo-EM) structures bound to Spike^{S2P} (**Figure 2.2**; **Figures 2.6-2.8**; and **Table 2.2**). Both nanobodies recognize RBD epitopes that overlap the ACE2 binding site (**Figure 2.2**). For Nb6 and Nb11, we resolved nanobody binding to both the open and closed conformations of Spike^{S2P}. We obtained a

3.0-Å map of Nb6 bound to closed Spike^{S2P}, which enabled modeling of the Nb6-Spike^{S2P} complex (**Figure 2.2**), including the complementarity-determining regions (CDRs). We also obtained lower-resolution maps for Nb6 bound to open Spike^{S2P} (3.8 Å), and Nb11 bound to open and closed Spike^{S2P} (4.2 and 3.7 Å, respectively). For these lower resolution maps, we could define the nanobody's binding orientation but not accurately model the CDRs.

Nb6 bound to closed Spike^{S2P} straddles the interface between two adjacent RBDs. Most of the contacting surfaces are contributed by CDR1 and CDR2 of Nb6 (**Figure 2.2**). CDR3 contacts the adjacent RBD positioned counterclockwise when viewed from the top (**Figure 2.2**). The binding of one Nb6 therefore stabilizes two adjacent RBDs in the down state and likely preorganizes the binding site for a second and third Nb6 molecule to stabilize the closed Spike conformation. By contrast, Nb11 bound to down-state RBDs only contacts a single RBD (**Figure 2.2**).

The structure of Nb6 bound to closed Spike^{S2P} enabled us to engineer bivalent and trivalent nanobodies predicted to lock all RBDs in the down state. We inserted flexible Gly-Ser linkers of either 15 or 20 amino acids to span the 52-Å distance between adjacent Nb6 monomers bound to down-state RBDs in closed Spike^{S2P} (**Figure 2.9**). These linkers are too short to span the 72-Å distance between Nb6 molecules bound to open Spike. Moreover, steric clashes would prevent binding of three RBDs in open Spike with a single up-state RBD even with longer linker length (**Figure 2.9**). By contrast, the minimum distance between adjacent Nb11 monomers bound to either open or closed Spike^{S2P} is 68 Å. We predicted that multivalent binding by Nb6 constructs would display substantially slowed dissociation rates owing to enhanced avidity.

In SPR experiments, both bivalent Nb6 with a 15–amino acid linker (Nb6-bi) and trivalent Nb6 with two 20–amino acid linkers (Nb6-tri) dissociate from Spike^{S2P} in a biphasic manner. The dissociation phase can be fitted to two components: a fast phase with kinetic rate constants k_{d1} of $2.7 \times 10^{-2} \text{ s}^{-1}$ for Nb6-bi and $2.9 \times 10^{-2} \text{ s}^{-1}$ for Nb6-tri, which are close to that observed for monovalent Nb6 ($k_d = 5.6 \times 10^{-2} \text{ s}^{-1}$), and a slow phase that is dependent on avidity ($k_{d2} = 3.1 \times 10^{-4} \text{ s}^{-1}$ for Nb6-bi and $k_{d2} < 1.0 \times 10^{-6} \text{ s}^{-1}$ for Nb6-tri) (**Figure 2.3**). The relatively similar k_d for the fast phase suggests that a fraction of the observed binding for the multivalent constructs is nanobody binding to a single Spike^{S2P} RBD. By contrast, the slow dissociation phase of Nb6-bi and Nb6-tri indicates engagement of two or three RBDs. We observed no dissociation for the slow phase of Nb6-tri over 10 min, indicating an upper boundary for k_{d2} of $1 \times 10^{-6} \text{ s}^{-1}$ and subpicomolar affinity. This measurement remains an upper boundary estimate because the measurement is limited by the intrinsic dissociation rate of Spike^{S2P} from the SPR chip imposed by the chemistry used to immobilize Spike^{S2P}. The true dissociation rate, therefore, may be considerably lower.

Biphasic dissociation could be explained by a slow interconversion between up- and downstate RBDs, with conversion to the more stable down state required for multivalent binding: A single domain of Nb6-tri engaged with an up-state RBD would dissociate rapidly. The system would then reequilibrate as the RBD flips into the down state, eventually allowing Nb6-tri to trap all RBDs in closed Spike^{S2P}. To test this directly, we varied the association time for Nb6-tri binding to Spike^{S2P}. Indeed, we observed an exponential decrease in the percentage of fast-phase dissociation with a half-life ($t_{1/2}$) of 65 s (**Figure 2.3**), which, we surmise, reflects the time scale of conversion between the RBD up and down states in Spike^{S2P}. Taken together, dimerization and trimerization of Nb6 afforded 750-fold and >200,000-fold gains in K_D , respectively.

Unable to determine the binding site of Nb3 by cryo-EM, we turned to radiolytic hydroxyl radical footprinting. We exposed apo- or Nb3-bound Spike^{S2P} to synchrotron x-ray radiation to label solvent-exposed amino acids with hydroxyl radicals, which we subsequently quantified by mass spectrometry of protease-digested Spike^{S2P} ¹⁸. Two neighboring surface residues on the S1 NTD of Spike (Met¹⁷⁷ and His²⁰⁷) were protected in the presence of Nb3 at a level consistent with prior observations of antibody-antigen interactions by hydroxyl radical footprinting¹⁹ (**Figure 2.10**). Previously discovered coronavirus neutralizing antibodies bind an epitope within the NTD of Spike with Fab fragments that are noncompetitive with the host cell receptor^{20,21}. Further SPR experiments demonstrated that Nb3 can bind Spike^{S2P} simultaneously with monovalent ACE2 (**Figure 2.11**). We hypothesized that the multivalent display of Nb3 on the surface of yeast may account for the partial decrease in Spike^{S2P} binding observed in the presence of ACE2-Fc. Indeed, a trivalent construct of Nb3 with 15–amino acid linkers (Nb3-tri) inhibited Spike^{S2P} binding to ACE2 cells with an IC₅₀ of 41 nM (**Figure 2.11**). How Nb3-tri disrupts Spike-ACE2 interactions remains unclear.

We next tested the neutralization activity of monovalent and trivalent versions of our top class I (Nb6 and Nb11) and class II (Nb3) nanobodies against SARS-CoV-2 pseudotyped lentivirus using a previously described assay²². Nb6 and Nb11 inhibited pseudovirus infection with IC₅₀ values of 2.0 and 2.4 mM, respectively. Nb3 inhibited pseudovirus infection with an IC₅₀ of 3.9 mM (**Figure 2.3** and **Table 2.1**). Nb6-tri shows a 2000-fold enhancement of inhibitory activity, with an IC₅₀ of 1.2 nM, whereas trimerization of Nb11 and Nb3 resulted in more modest gains of 40- and 10-fold (51 and 400 nM), respectively (**Figure 2.3**). We confirmed these neutralization activities with a

viral plaque assay using live SARS-CoV-2 virus infection of VeroE6 cells. Here, Nb6-tri proved exceptionally potent, neutralizing SARS-CoV-2 with an average IC₅₀ of 160 pM (**Figure 2.3**). Nb3-tri neutralized SARS-CoV-2 with an average IC₅₀ of 140 nM (**Figure 2.3**).

We further optimized the potency of Nb6 by selecting a saturation mutagenesis library targeting all three CDRs. Two rounds of selection identified high-affinity clones with two penetrant mutations: I27Y (Ile²⁷→Tyr) in CDR1 and P105Y (Pro¹⁰⁵→Tyr) in CDR3. We incorporated these mutations into Nb6 to generate matured Nb6 (mNb6), which binds with 500-fold increased affinity to Spike^{S2P} (**Figure 2.4**). mNb6 inhibits both pseudovirus and live SARS-CoV-2 infection with low nanomolar potency, a ~200-fold improvement compared with Nb6 (**Figure 2.4** and **Table 2.1**).

A 2.9-Å cryo-EM structure shows that mNb6 binds to closed Spike^{S2P} (**Figure 2.4** and **Figure 2.12**). mNb6 induces a slight rearrangement of the down-state RBDs as compared with Spike^{S2P} bound to Nb6, inducing a 9° rotation of the RBD away from the central threefold-symmetry axis. This deviation likely arises from a different interaction between CDR3 and Spike^{S2P}, which nudges the RBDs into a new resting position (**Figure 2.4**). Although the I27Y substitution optimizes local contacts between CDR1 in its original binding site on the RBD, the P105Y substitution leads to a marked rearrangement of CDR3 in mNb6 (**Figure 2.4**). This conformational change yields a different set of contacts between mNb6 CDR3 and the adjacent RBD. An x-ray crystal structure of mNb6 alone revealed dramatic conformational differences in CDR1 and CDR3 between free and Spike^{S2P}-bound mNb6 (**Figure 2.4** and **Table 2.3**). Although differences in loop conformation in the crystal structure may arise from crystal lattice contacts, they are suggestive of

conformational heterogeneity for unbound mNb6 and induced-fit rearrangements upon binding to Spike^{S2P}.

The binding orientation of mNb6 is similar to that of Nb6, suggesting that multivalent design would likewise enhance binding affinity. Unlike Nb6-tri, trivalent mNb6 with a 20–amino acid linker (mNb6-tri) bound to Spike^{S2P} with no observable fast-phase dissociation and no measurable dissociation over 10 minutes, yielding an upper bound for the dissociation rate constant k_d of $1.0 \times 10^{-6} \text{ s}^{-1}$ ($t_{1/2} > 8$ days) and a K_D of $<1 \text{ pM}$ (**Figure 2.4**). mNb6-tri displays further gains in potency in both pseudovirus and live SARS-CoV-2 infection assays with IC_{50} values of 120 pM (5.0 ng/ml) and 54 pM (2.3 ng/ml), respectively (**Figure 2.4** and **Table 2.1**). Given the subpicomolar affinity observed by SPR, it is likely that these viral neutralization potencies reflect the lower limit of the assays. mNb6-tri is therefore an exceptionally potent SARS-CoV-2 neutralizing molecule.

We next tested whether viral neutralization by the class I nanobody mNb6 is potentially synergistic with the class II nanobody Nb3-tri. In pseudovirus neutralization assays, we observed an additive effect when combining Nb3-tri with mNb6 (**Figure 2.13**). However, the potency for mNb6 viral neutralization was unchanged with increasing concentrations of Nb3-tri, suggesting minimal synergy between these two nanobodies.

We next tested Nb6 and its derivatives for stability. Circular dichroism revealed melting temperatures of 66.9° , 62.0° , 67.6° , and 61.4°C for Nb6, Nb6-tri, mNb6, and mNb6-tri, respectively (**Figure 2.14**). Moreover, mNb6 and mNb6-tri were stable to lyophilization and to

aerosolization, showing no aggregation by size exclusion chromatography, and preserved high-affinity binding to Spike^{S2P} (**Figure 2.5** and **Figure 2.14**). Finally, mNb6-tri retains potent inhibition of pseudovirus and live SARS-CoV-2 infection after aerosolization, lyophilization, or heat treatment for 1 hour at 50°C (**Figure 2.5** and **Figure 2.14**).

2.4 Discussion

Strategies to prevent SARS-CoV-2 entry into the host cell aim to block the ACE2-RBD interaction. Although high-affinity monoclonal antibodies are leading the way as potential therapeutics^{20,23,24,25,26,27,28,29,30}, they are expensive to produce by mammalian cell expression and need to be intravenously administered by health care professionals³¹. Large doses are needed for prophylactic use because only a small fraction of systemic antibodies cross the epithelial cell layers lining the airways³². By contrast, nanobodies can be inexpensively produced in bacteria or yeast. The inherent stability of nanobodies enables aerosolized delivery directly to the nasal and lung epithelia³³. Indeed, aerosol delivery of a trimeric nanobody targeting respiratory syncytial virus (ALX-0171) was recently demonstrated to be effective in substantially decreasing measurable viral load in hospitalized infants³⁴. Finally, potential immunogenicity of camelid-derived nanobodies can be mitigated by established humanization strategies³⁵.

Nanobody multimerization has been shown to improve target affinity by avidity^{33, 36}. In the case of Nb6 and mNb6, structure-guided design of a multimeric construct that simultaneously engages all three RBDs yielded profound gains in potency. Furthermore, because RBDs must be in the up state to engage with ACE2, conformational control of RBD accessibility serves as an added

neutralization mechanism³⁰. Indeed, when mNb6-tri engages with Spike, it prevents ACE2 binding both by directly occluding the binding site and by locking the RBDs into an inactive conformation.

Our discovery of class II neutralizing nanobodies demonstrates potentially new mechanisms of disrupting Spike function. The pairing of class I and class II nanobodies in a prophylactic or therapeutic cocktail could provide both potent neutralization and prevention of escape variants²³. The combined stability, potency, and diverse epitope engagement of our anti-Spike nanobodies therefore provide a distinctive potential prophylactic and therapeutic strategy to limit the continued toll of the COVID-19 pandemic.

2.5 Figures

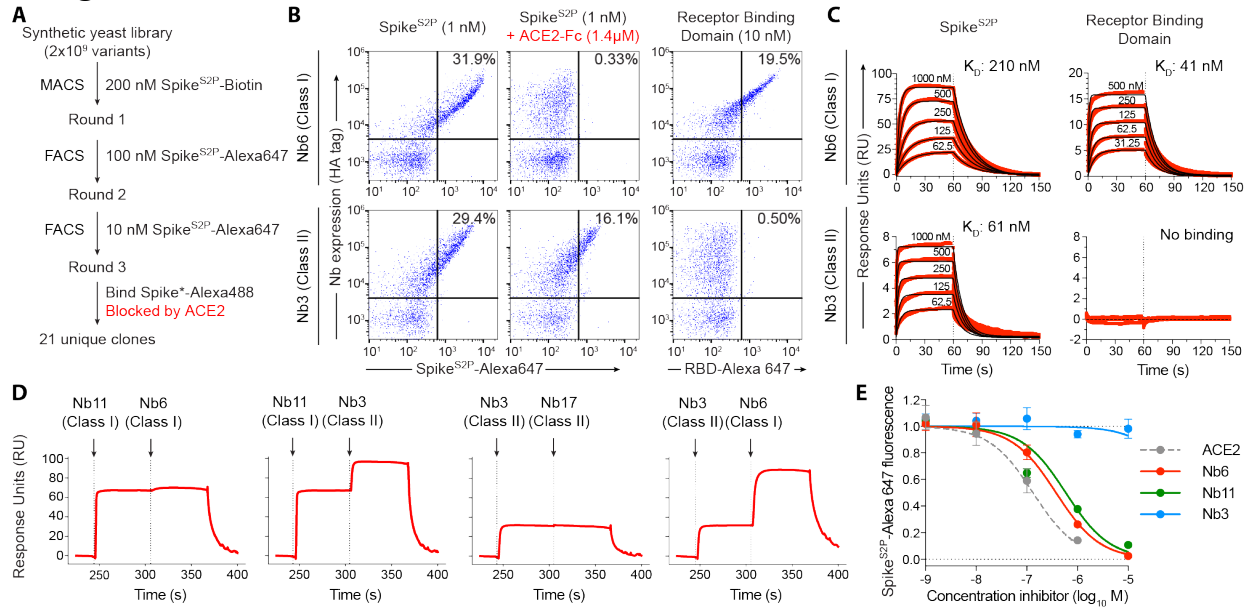


Figure 2.1 Discovery of two distinct classes of anti-Spike nanobodies

A, Selection strategy for identification of anti-Spike nanobodies that disrupt Spike-ACE2 interactions using magnetic bead selections (MACS) or fluorescence activated cell sorting (FACS). **B**, Flow cytometry of yeast displaying Nb6 (a Class I nanobody) or Nb3 (a Class II nanobody). Nb6 binds Spike^{S2P}-Alexa 647 and receptor binding domain (RBD-Alexa 647). Nb6 binding to Spike^{S2P} is completely disrupted by an excess (1.4 μ M) of ACE2-Fc. Nb3 binds Spike^{S2P}, but not the RBD. Nb3 binding to Spike^{S2P} is partially decreased by ACE2-Fc. **C**, SPR of Nb6 and Nb3 binding to either Spike^{S2P} or RBD. Red traces are raw data and global kinetic fits are shown in black. Nb3 shows no binding to RBD. **D**, SPR experiments with immobilized Spike^{S2P} show that Class I and Class II nanobodies can bind Spike^{S2P} simultaneously. By contrast, two Class I nanobodies or Class II nanobodies do not bind simultaneously. **E**, Nanobody inhibition of 1 nM Spike^{S2P}-Alexa 647 binding to ACE2 expressing HEK293T cells. $n = 3$ (ACE2, Nb3) or 5 (Nb6, Nb11) biological replicates. All error bars represent s.e.m.

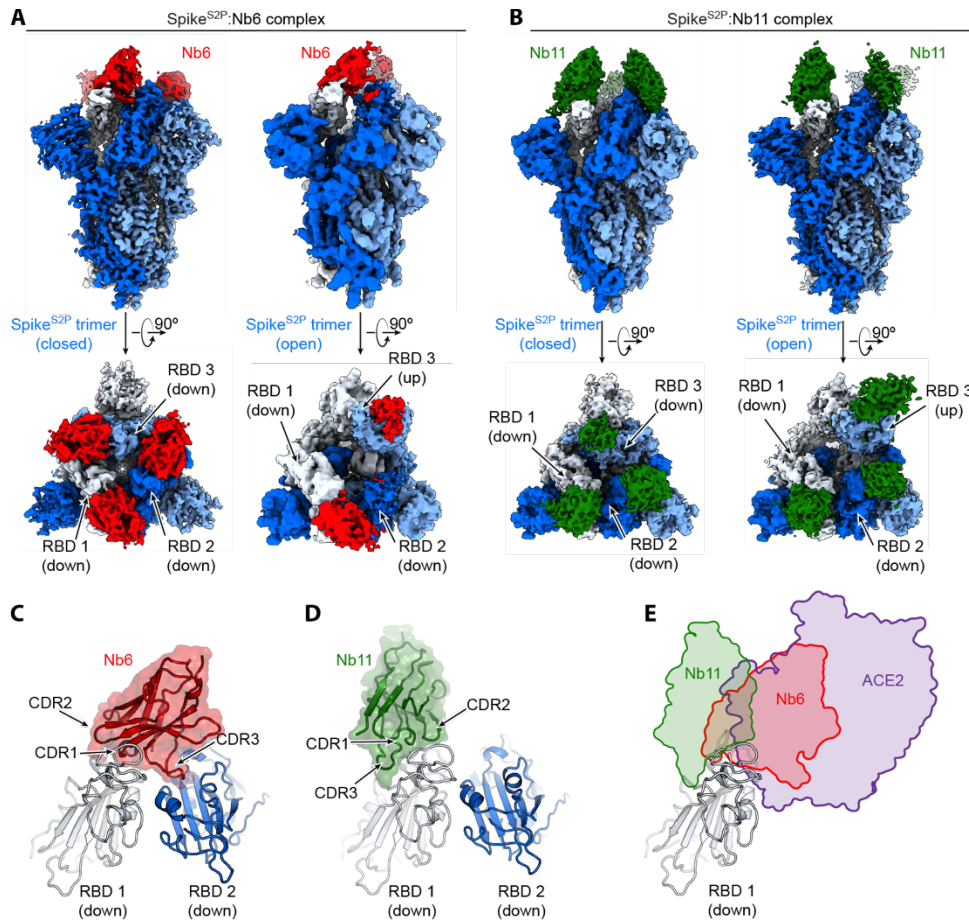


Figure 2.2 Cryo-EM structures of Nb6 and Nb11 bound to Spike.

A, Cryo-EM maps of Spike^{S2P}-Nb6 complex in either closed (left) or open (right) Spike^{S2P} conformation. **B**, Cryo-EM maps of Spike^{S2P}-Nb11 complex in either closed (left) or open (right) Spike^{S2P} conformation. The top views show receptor binding domain (RBD) up- or down-states. **C**, Nb6 straddles the interface of two down-state RBDs, with CDR3 reaching over to an adjacent RBD. **D**, Nb11 binds a single RBD in the down-state (displayed) or similarly in the up-state. No cross-RBD contacts are made by Nb11 in either RBD up- or down-state. **E**, Comparison of RBD epitopes engaged by ACE2 (purple), Nb6 (red), or Nb11 (green). Both Nb11 and Nb6 directly compete with ACE2 binding.

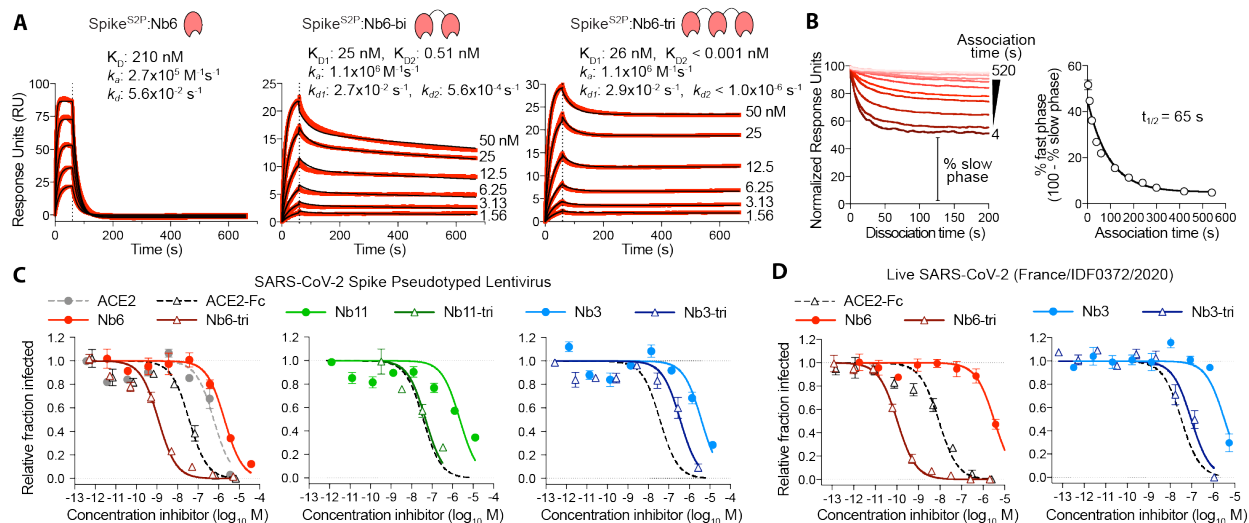


Figure 2.3 Multivalency improves nanobody affinity and inhibitory efficacy.

A, SPR of Nb6 and multivalent variants. Red traces show raw data and black lines show global kinetic fit for Nb6 and independent fits for association and dissociation phases for Nb6-bi and Nb6-tri. **B**, Dissociation phase SPR traces for Nb6-tri after variable association time ranging from 4 to 520 s. Curves were normalized to maximal signal at the beginning of the dissociation phase. Percent fast phase is plotted as a function of association time (right) with a single exponential fit. $n = 3$ independent biological replicates. **C**, Inhibition of pseudotyped lentivirus infection of ACE2 expressing HEK293T cells. $n = 3$ biological replicates for all but Nb11-tri ($n = 2$) **D**, Inhibition of live SARS-CoV-2 virus. Representative biological replicate with $n = 3$ (right panel) or 4 (left panel) technical replicates per concentration. $n = 3$ biological replicates for all but Nb3 and Nb3-tri ($n = 2$). All error bars represent s.e.m.

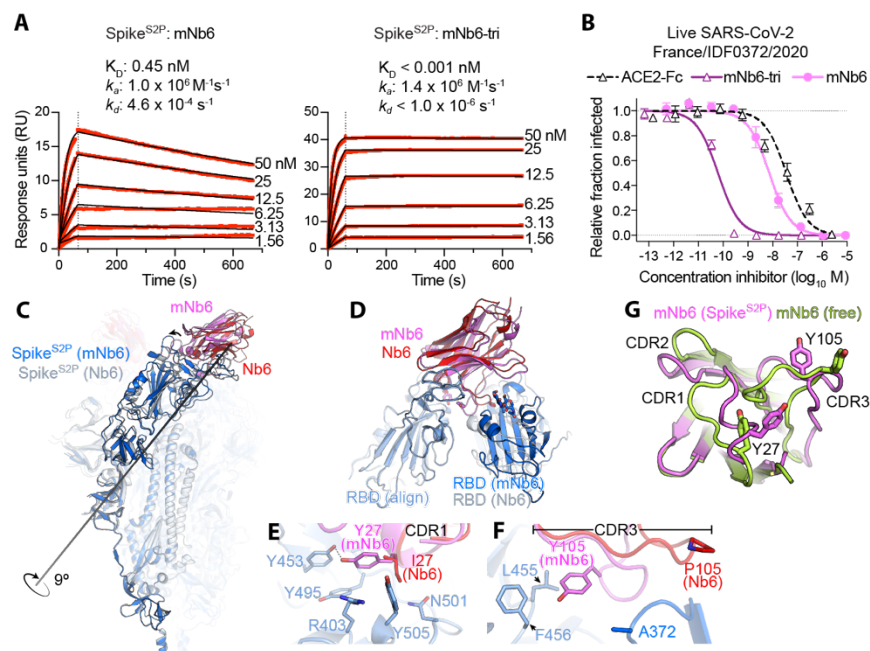


Figure 2.4 Affinity maturation of Nb6 yields a picomolar SARS-CoV-2 neutralizing molecule.

A, SPR of mNb6 and mNb6-tri binding to immobilized Spike^{S2P}. Red traces show raw data and black lines show global kinetic fit. No dissociation was observed for mNb6-tri over 10 minutes. **B**, mNb6 and mNb6-tri inhibit SARS-CoV-2 infection of VeroE6 cells in a plaque assay. Representative biological replicate with $n = 4$ technical replicates per concentration. $n = 3$ biological replicates for all samples. All error bars represent s.e.m. **C**, Comparison of closed Spike^{S2P} bound to mNb6 and Nb6. Rotational axis for RBD movement is highlighted. **D**, Comparison of receptor binding domain (RBD) engagement by Nb6 and mNb6. One RBD was used to align both structures (RBD align), demonstrating changes in Nb6 and mNb6 position and the adjacent RBD. **E**, CDR1 of Nb6 and mNb6 binding to the RBD. As compared to I27 in Nb6, Y27 of mNb6 hydrogen bonds to Y453 and optimizes pi-pi and pi-cation interactions with the RBD. **F**, CDR3 of Nb6 and mNb6 binding to the RBD demonstrating a large conformational rearrangement of the entire loop in mNb6. **G**, Comparison of mNb6 complementarity determining regions in either the cryo-EM structure of the Spike^{S2P}-mNb6 complex or an X-ray crystal structure of mNb6 alone.

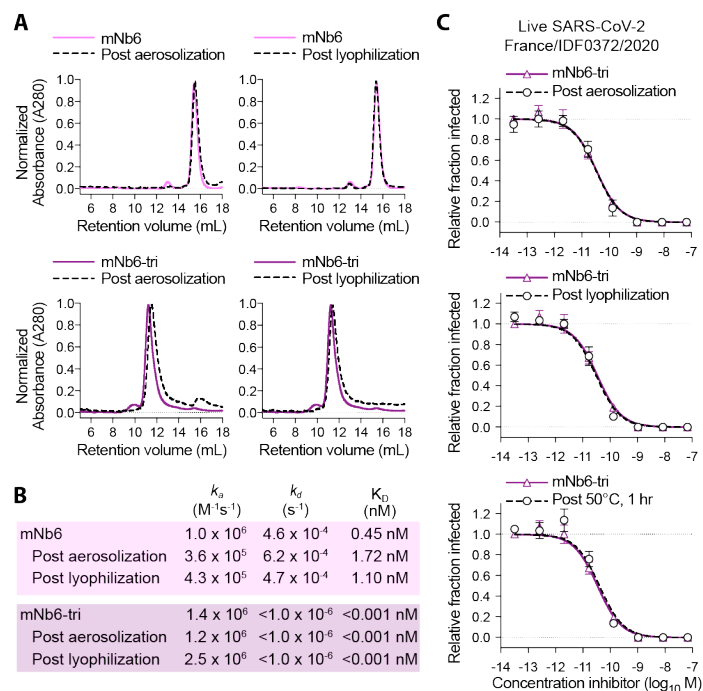


Figure 2.5 mNb6 and mNb6-tri retain activity after aerosolization, lyophilization, and heat treatment.

A, Size exclusion chromatography of nanobodies after lyophilization or aerosolization. **B**, Summary table of SPR kinetics data and affinities for aerosolized or lyophilized mNb6 and mNb6-tri. **C**, Inhibition of SARS-CoV-2 infection of VeroE6 cells by mNb6-tri after aerosolization, lyophilization, or heat treatment at 50°C for 1 hour. Representative biological replicate with $n = 2$. Technical replicates $n = 3$ per concentration.

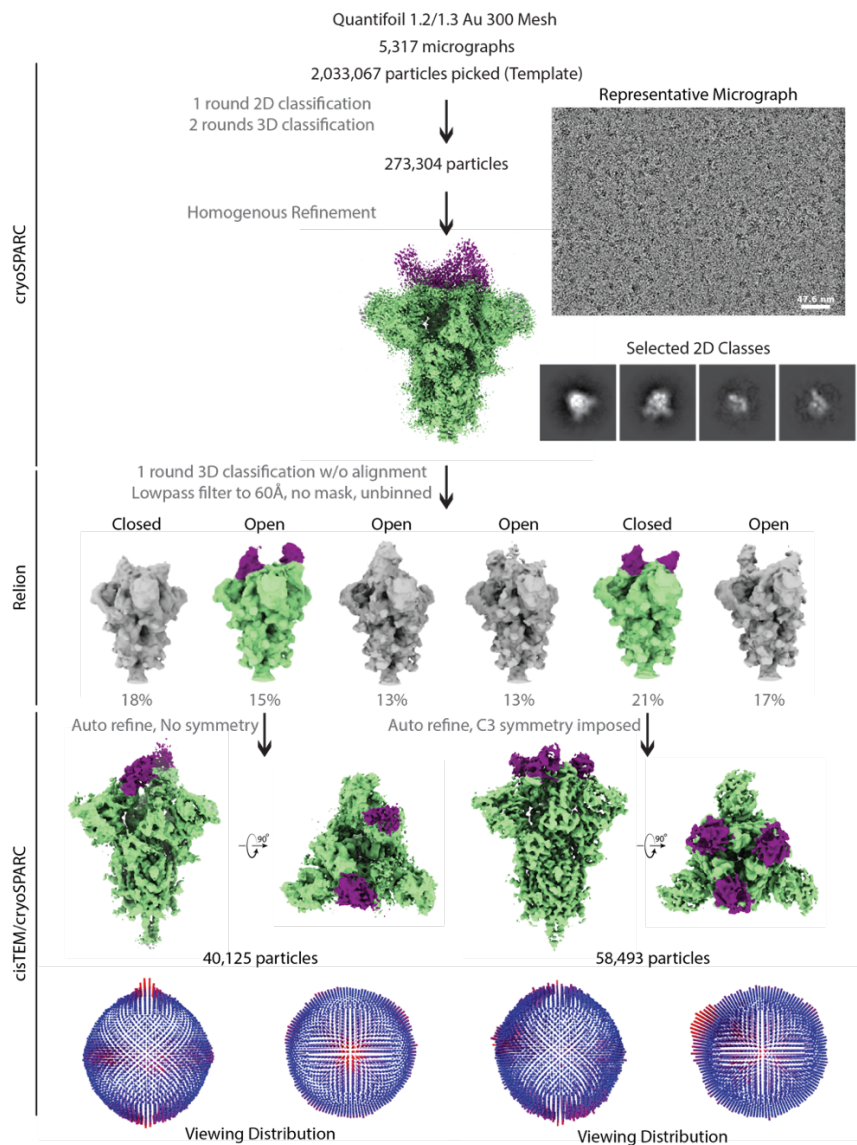


Figure 2.6 Cryo-EM workflow for Nb6

A flowchart representation of the classification workflow for Spike^{S2P}-Nb6 complexes yielding open and closed Spike^{S2P} conformations. From top to bottom, particles were template picked with a set of 20 Å low-pass filtered 2D backprojections of apo-Spike^{S2P} in the closed conformation. Extracted particles in 2D classes suggestive of various Spike^{S2P} views were subject to a round of heterogenous refinement in cryoSPARC with two naïve classes generated from a truncated Ab initio job, and a 20 Å low-pass filtered volume of apo-Spike^{S2P} in the closed conformation. Particles in the Spike^{S2P} 3D class were subject to 25 iterations of 3D classification into 6 classes without alignment in RELION, using the same input volume from cryoSPARC 3D classification, low pass filtered to 60 Å, T = 8. Particles in classes representing the open and closed Spike^{S2P} conformations were imported into cisTEM for automatic refinement. Viewing distribution plots were generated with pyEM, and visualized with ChimeraX. Half maps from refinement were imported into cryoSPARC for local resolution estimation as shown in **Figure 2.8**.

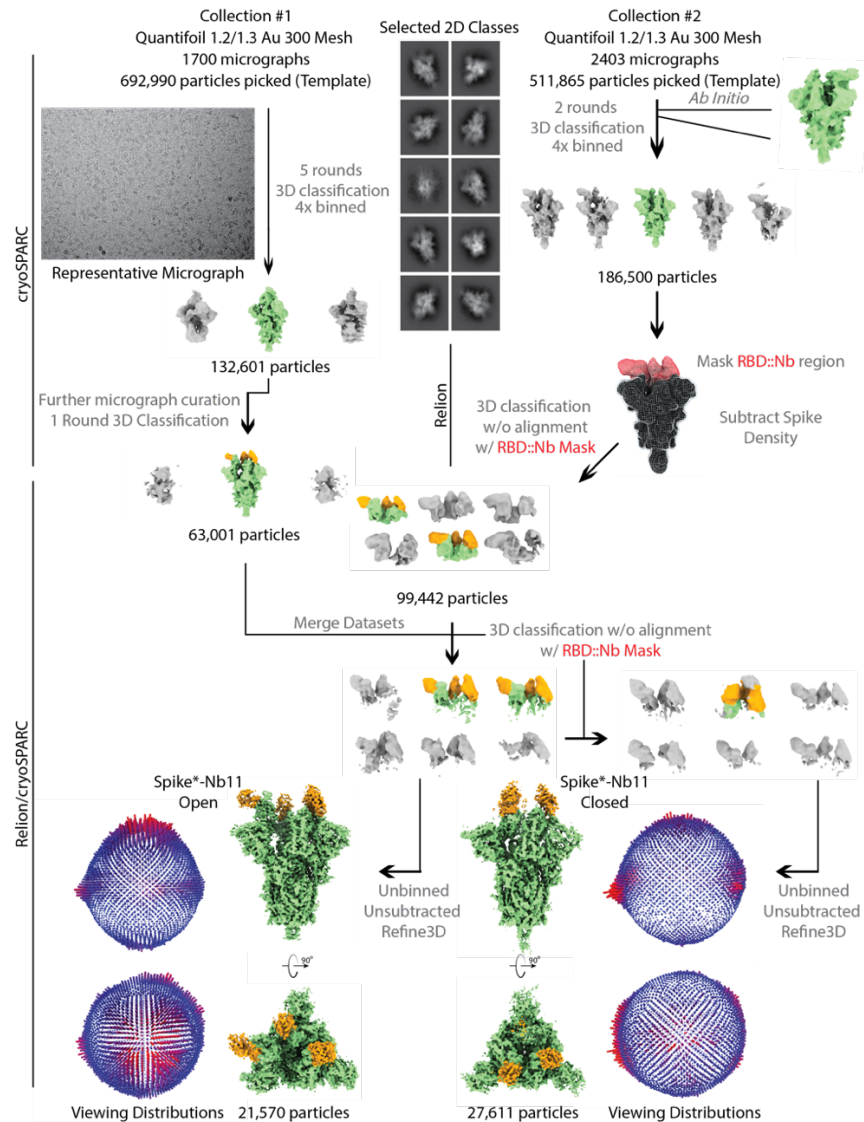


Figure 2.7 Cryo-EM workflow for Nb11

Classification workflow for Spike^{S2P}-Nb11 complexes yielding open and closed Spike^{S2P} conformations. Particles were template picked from two separate collections with a set of 20 Å low-pass filtered 2D backprojections of apo-Spike^{S2P} in the closed conformation. Extracted particles were Fourier cropped to 128 pixels prior to extensive heterogeneous refinement in cryoSPARC, using a 20 Å low-pass filtered volume of apo-Spike^{S2P} in the closed conformation and additional naïve classes for removal of non-Spike^{S2P} particles. After cryoSPARC micrograph curation and heterogeneous refinement, Spike^{S2P} density corresponding to all regions outside of the ACE2-RBD::Nanobody interface were subtracted. A mask around the ACE2-RBD::Nanobody interface was generated, and used for multiple rounds of 3D classification without alignment in RELION. Particles in classes representing open and closed Spike^{S2P} conformations were selected, unmasked and unbinned prior to refinement in RELION. Viewing distribution plots were generated with pyEM, and visualized with ChimeraX. Half maps from refinement were imported into cryoSPARC for local resolution estimation as shown in **Figure 2.8**.

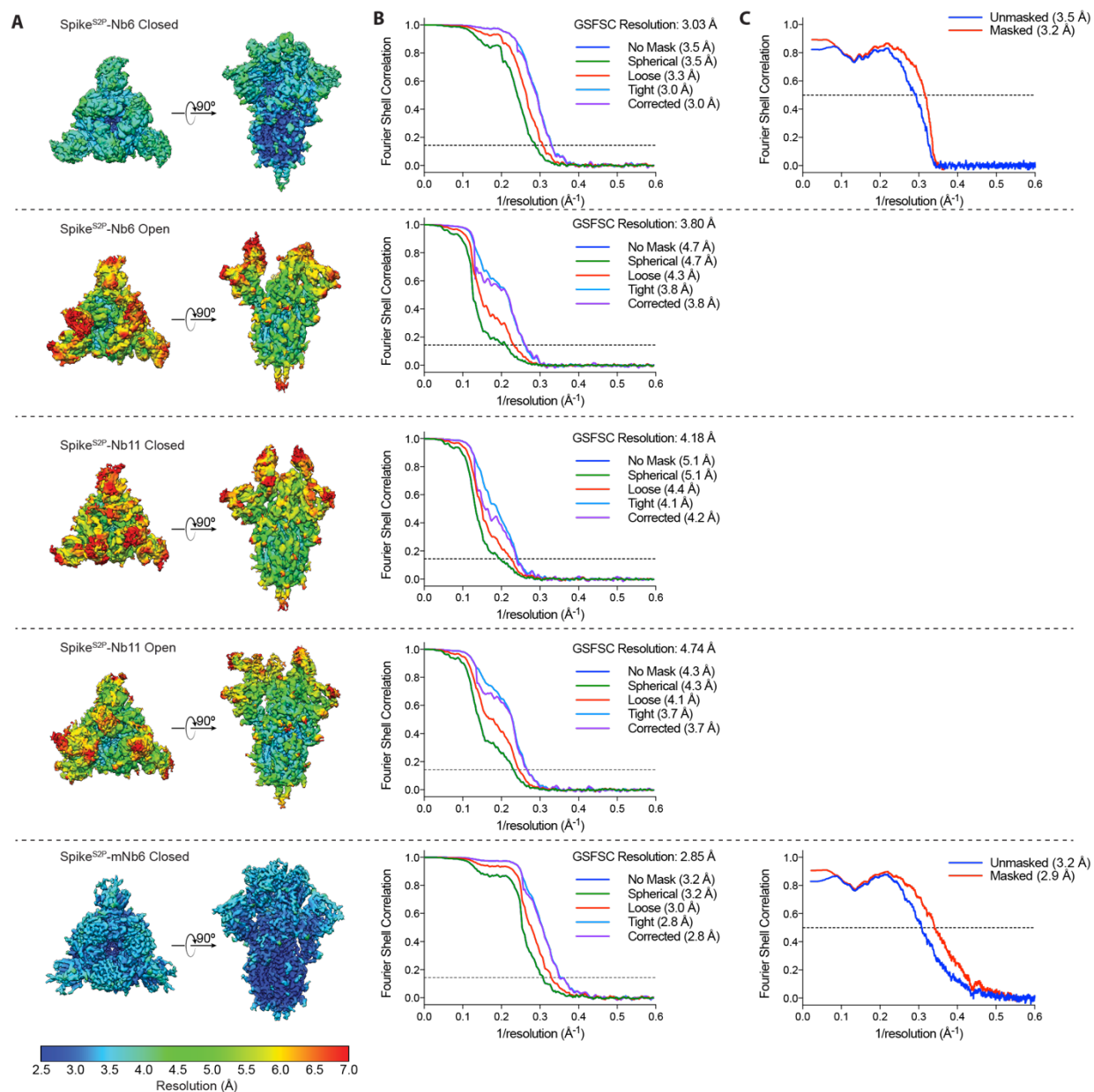


Figure 2.8 Resolution of cryo-EM maps and models.

A, Local resolution estimates of Spike^{S2P} complexes as generated in cryoSPARC. All maps (except mNb6) are shown with the same enclosed volume. All maps are colored on the same scale, as indicated. **B**, Gold standard Fourier Shell Correlation (GSFSC) plots for cryo-EM maps calculated within cryoSPARC. Resolution values in parentheses represent values at FSC = 0.143 (dashed line). **C**, Model-map correlation calculated in Phenix. Resolution values in parentheses represent values at FSC = 0.5 (dashed line).

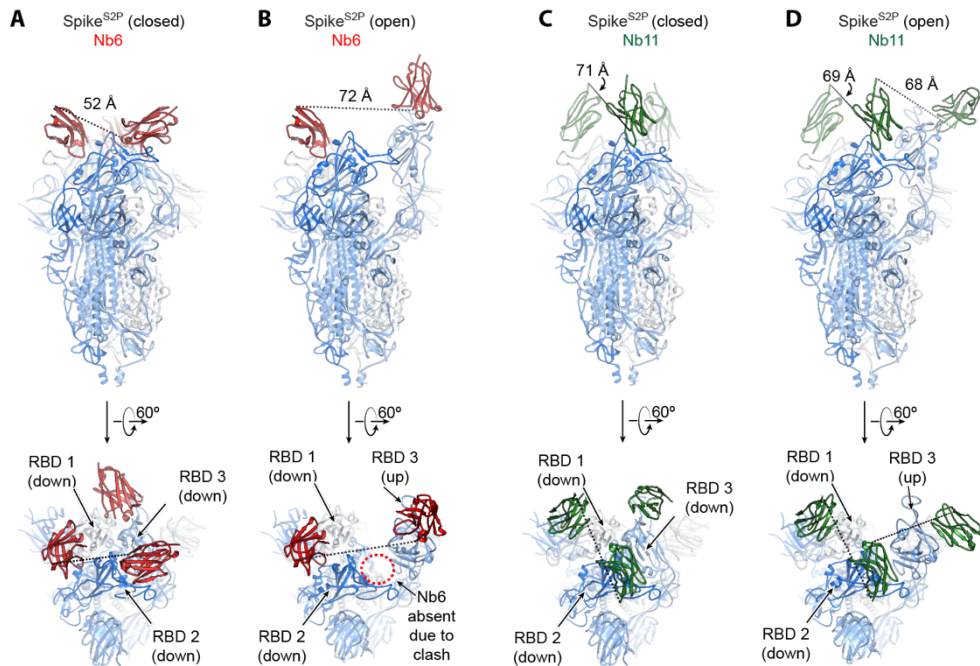


Figure 2.9 Modeling of distances for multimeric nanobody design.

A, Model of Spike^{S2P}:Nb6 complex in the closed state. The minimal distance between adjacent Nb6 N- and C-termini is 52 Å (dashed line). **B**, Model of Spike^{S2P}:Nb6 complex in the open state with Nb6 docked into the cryo-EM density for up-state RBD. Minimal distance between N- and C-termini of both nanobodies is 72 Å. Nb6 cannot bind RBD2 in open Spike^{S2P}, as this would sterically clash with RBD3. **C**, Model of Spike^{S2P}:Nb11 complex in the closed state. The minimal distance between adjacent Nb6 N- and C-termini is 71 Å (dashed line). **D**, Model of Spike^{S2P}:Nb11 complex in the open state. The minimal distance between adjacent Nb6 N- and C-termini is 68 Å between Nb11 bound to RBD2 in the down-state and RBD3 in the up-state. For **B**, the model of Nb6 from **A** was docked into the cryo-EM map to enable modeling of distance between N- and C-termini. For **C** and **D**, a generic nanobody was docked into cryo-EM maps to model the distance between N- and C-termini.

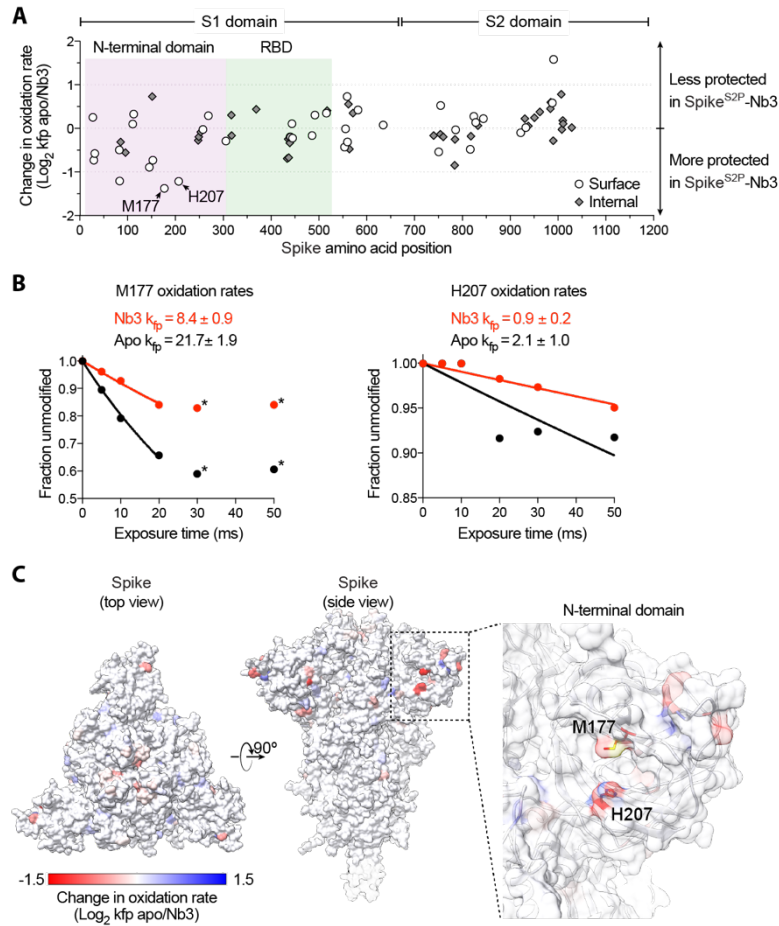


Figure 2.10 Radiolytic hydroxyl radical footprinting of Spike^{S2P}.

A, Change in oxidation rate between Spike^{S2P} and Nb3-Spike^{S2P} complexes at all residues. A cluster of highly protected residues in the Spike^{S2P}-Nb3 complex is observed in the N-terminal domain. **B**, Oxidation rate plots of the two (M177, H207) most heavily protected residues upon Nb3 binding to Spike^{S2P}. Data points labeled with an asterisk are excluded from rate calculations as these values fall outside of the first order reaction, likely due to extensive oxidation-mediated damage. **C**, Change in oxidation rate mapped onto Spike in the all RBD down conformation.

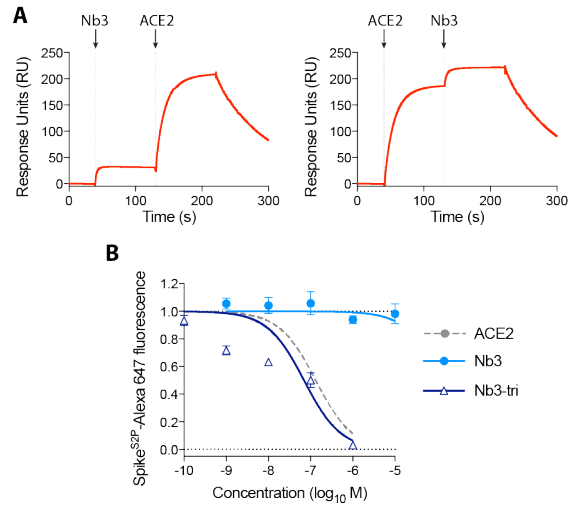


Figure 2.11 Multivalent Nb3 construct inhibits Spike^{S2P}:ACE2 interaction.

A, SPR experiments with immobilized Spike^{S2P} show that Nb3 and monovalent ACE2 can bind Spike^{S2P} simultaneously. The order of Nb3 and monovalent ACE2 does not affect the binding of the second reagent. Nb3 therefore does not inhibit Spike^{S2P} binding to monovalent ACE2. **B**, Nanobody inhibition of 1 nM Spike^{S2P}-Alexa 647 binding to ACE2 expressing HEK293T cells by either monovalent or trivalent Nb3. $n = 2$ biological replicates for Nb3-tri. All error bars represent s.e.m.

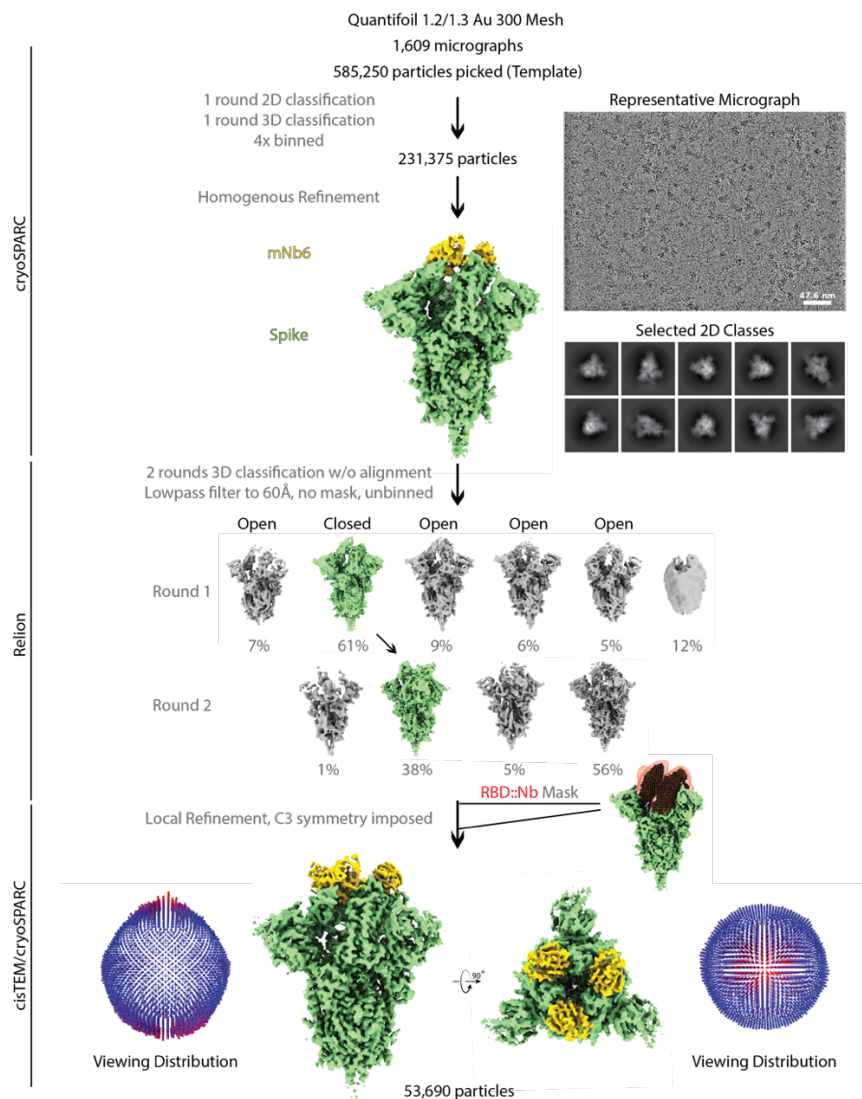


Figure 2.12 CryoEM workflow for mNb6.

Classification workflow for the Spike^{S2P}-mNb6 complex yielding a closed Spike^{S2P} conformation. From top to bottom, particles were template picked from two separate collections with a set of 20 Å low-pass filtered 2D backprojections of apo-Spike^{S2P} in the closed conformation. Extracted particles were Fourier cropped to 96 pixels prior to 2D classification. Particles in Spike^{S2P} 2D classes were selected for a round of heterogeneous refinement in cryoSPARC using a 20 Å low-pass filtered volume of apo-Spike^{S2P} in the closed conformation and additional naïve classes for removal of non-Spike^{S2P} particles. In RELION, particles in the Spike^{S2P} 3D class were subject to two rounds of 3D classification without alignment into 6 classes using the same input volume from cryoSPARC 3D classification, low pass filtered to 60 Å, T = 8. Unbinned particles in the Spike^{S2P}-closed conformation were exported into cisTEM for automatic refinement, followed by local refinement using a mask around the RBD::Nanobody interface. Viewing distribution plots were generated with pyEM, and visualized with ChimeraX. Half maps from refinement were imported into cryoSPARC for local resolution estimation as shown in **Figure 2.8**.

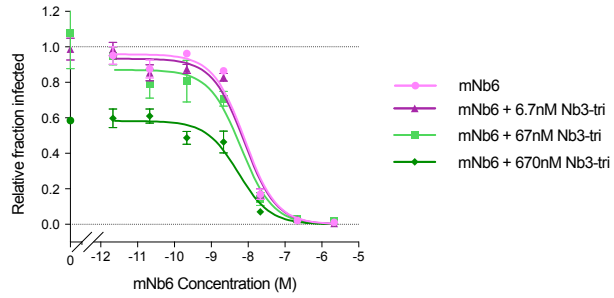


Figure 2.13 mNb6 and Nb3-tri are additive for viral neutralization.

Inhibition of pseudotyped lentivirus infection of ACE2 expressing HEK293T cells by mNb6 with increasing concentrations of Nb3-tri. mNb6 neutralization is additive with Nb3-tri, as demonstrated by inhibitory activity at a sub-saturating dose of Nb3-tri. However, the potency of mNb6 is unchanged by Nb3-tri, suggesting no synergistic effect on viral neutralization.

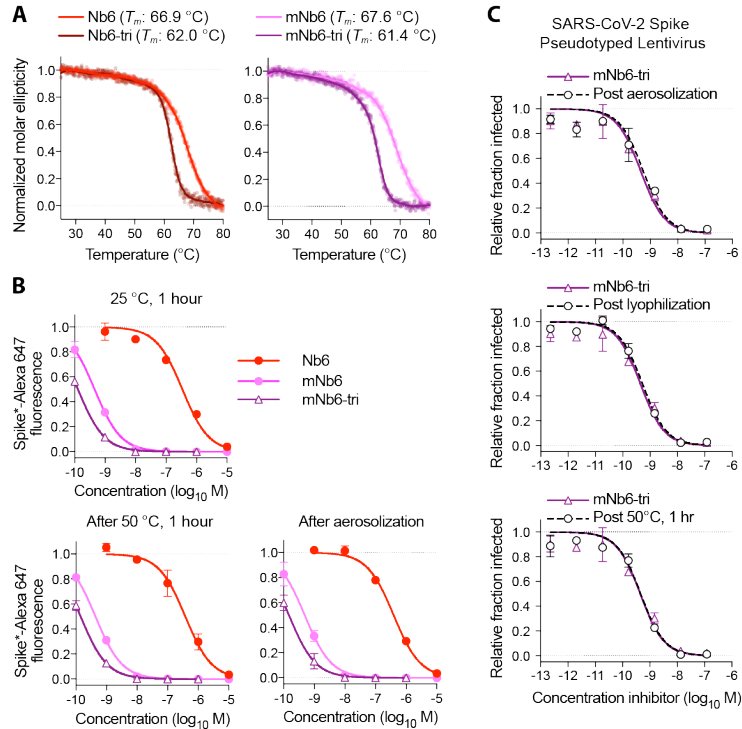


Figure 2.14 Stability of Nb6 and its derivatives.

A, Thermal denaturation of nanobodies assessed by circular dichroism measurement of molar ellipticity at 204 nm. Apparent melting temperatures (T_m) for each nanobody are indicated. **B**, Nanobody inhibition of 1 nM Spike^{S2P}-Alexa 647 binding to ACE2 expressing HEK293T cells after incubation at either 25 °C or 50 °C for 1 hour or after aerosolization. **C**, Inhibition of pseudotyped lentivirus infection of ACE2 expressing HEK293T cells by mNb6-tri after aerosolization, lyophilization, or heat treatment at 50°C for 1 hour.

2.6 Tables

Table 2.1 Nanobody affinities and efficacies in neutralization assays.

Nanobody	Class	Spike ^{S2P} Binding (SPR)			RBD Binding (SPR)			Spike ^{S2P} Competition IC ₅₀ (s.e.m) (M) ^a	SARS-CoV-2 Pseudovirus IC ₅₀ (s.e.m.) (M) ^b	Live SARS-CoV-2 IC ₅₀ (s.e.m.) (M) ^c
		k_a (M ⁻¹ s ⁻¹)	k_d (s ⁻¹)	K _D (M)	k_a (M ⁻¹ s ⁻¹)	k_d (s ⁻¹)	K _D (M)			
Nb2	I	9.0x10 ⁵	5.3x10 ⁻¹	5.9x10 ⁻⁷	1.0x10 ⁶	9.9x10 ⁻¹	9.7x10 ⁻⁷	8.3x10 ⁻⁶ (1.7x10 ⁻⁶)	NP	NP
Nb3 ^d	II	1.8x10 ⁶	1.1x10 ⁻¹	6.1x10 ⁻⁸	NB			NC	3.9x10 ⁻⁶ (7.9x10 ⁻⁷)	3.0x10 ⁻⁶ (3.2x10 ⁻⁷)
Nb6	I	2.7x10 ⁵	5.6x10 ⁻²	2.1x10 ⁻⁷	2.1x10 ⁶	8.7x10 ⁻²	4.1x10 ⁻⁸	3.7x10 ⁻⁷ (4.9x10 ⁻⁸)	2.0x10 ⁻⁶ (3.5x10 ⁻⁷)	3.3x10 ⁻⁶ (7.2x10 ⁻⁷)
Nb8	I	1.4x10 ⁵	8.1x10 ⁻¹	5.8x10 ⁻⁶	6.6x10 ⁵	3.3x10 ⁻¹	5.1x10 ⁻⁷	4.8x10 ⁻⁶ (4.9x10 ⁻⁷)	NP	NP
Nb11	I	1.2x10 ⁶	1.6x10 ⁻¹	1.4x10 ⁻⁷	3.2x10 ⁶	2.4x10 ⁻¹	7.6x10 ⁻⁸	5.4x10 ⁻⁷ (1.2x10 ⁻⁷)	2.4x10 ⁻⁶ (5.4x10 ⁻⁷)	NP
Nb12	I	1.2x10 ²	2.0x10 ⁻⁴	1.6x10 ⁻⁶	Biphasic	Biphasic	Biphasic	2.5x10 ⁻⁷ (5.5x10 ⁻⁸)	1.2x10 ⁻⁶ (9.0x10 ⁻⁷)	NP
Nb15	I	1.7x10 ⁵	2.3x10 ⁻¹	1.3x10 ⁻⁶	6.0x10 ⁵	2.2x10 ⁻¹	3.6x10 ⁻⁷	2.2x10 ⁻⁶ (2.5x10 ⁻⁷)	6.7x10 ⁻⁶ (3.6x10 ⁻⁶)	NP
Nb16	I	1.1x10 ⁵	1.3x10 ⁻¹	1.3x10 ⁻⁶	NP			9.5x10 ⁻⁷ (1.1x10 ⁻⁷)	NP	NP
Nb17 ^d	II	7.3x10 ⁵	2.0x10 ⁻¹	2.7x10 ⁻⁷	NB			NC	7.6x10 ⁻⁶ (1.0x10 ⁻⁶)	NP
Nb18 ^d	II	1.4x10 ⁵	6.4x10 ⁻³	4.5x10 ⁻⁸	NB			5.2x10 ⁻⁵ (1.5x10 ⁻⁵)	NP	NP
Nb19	I	2.4x10 ⁴	1.1x10 ⁻¹	4.5x10 ⁻⁶	1.0x10 ⁵	8.9x10 ⁻²	8.8x10 ⁻⁷	4.1x10 ⁻⁶ (4.9x10 ⁻⁷)	2.4x10 ⁻⁵ (7.7x10 ⁻⁶)	NP
Nb24	I	9.3x10 ⁵	2.7x10 ⁻¹	2.9x10 ⁻⁷	2.4x10 ⁶	3.5x10 ⁻¹	1.5x10 ⁻⁷	7.5x10 ⁻⁷ (1.0x10 ⁻⁷)	NP	NP
ACE2	N/A	2.7x10 ⁵	1.2x10 ⁻²	4.4x10 ⁻⁸	NP	NP	NP	1.7x10 ⁻⁷ (6.6x10 ⁻⁸)	6.2x10 ⁻⁷ (1.7x10 ⁻⁷)	NP
mNb6	I	1.0x10 ⁶	4.5x10 ⁻⁴	4.5x10 ⁻¹⁰	1.1x10 ⁶	6.4x10 ⁻⁴	5.6x10 ⁻¹⁰	1.3x10 ⁻⁹ (4.1x10 ⁻¹⁰)	6.3x10 ⁻⁹ (1.6x10 ⁻⁹)	1.2x10 ⁻⁸ (2.5x10 ⁻⁹)
Nb3-bi	II	NP	NP	NP	NP	NP	NP	NP	3.6x10 ⁻⁷ (1.5x10 ⁻⁷)	1.8x10 ⁻⁷ (1.2x10 ⁻⁸)
Nb3-tri	II	Biphasic	Biphasic	Biphasic	NP	NP	NP	4.1x10 ⁻⁸ (1.6x10 ⁻⁸)	4.0x10 ⁻⁷ (1.6x10 ⁻⁷)	1.4x10 ⁻⁷ (4.9x10 ⁻⁸)
Nb6-bi	I	1.1x10 ⁶	2.7x10 ⁻² 5.6x10 ⁻⁴	2.5x10 ⁻⁸ 5.1x10 ⁻¹⁰	NP	NP	NP	NP	6.3x10 ⁻⁸ (1.5x10 ⁻⁸)	NP
Nb6-tri (15 aa)	I	1.2x10 ⁶	2.1x10 ⁻² <1.0x10 ⁻⁶	1.8x10 ⁻⁸ <1.0x10 ⁻¹²	NP	NP	NP	8.6x10 ⁻¹⁰ (1.8x10 ⁻¹⁰)	1.0x10 ⁻⁹ (1.9x10 ⁻¹⁰)	NP
Nb6-tri (20 aa)	I	1.1x10 ⁶	2.9x10 ⁻² <1.0x10 ⁻⁶	2.6x10 ⁻⁸ <1.0x10 ⁻¹²	NP	NP	NP	1.5x10 ⁻⁹ (5.2x10 ⁻¹⁰)	1.2x10 ⁻⁹ (2.5x10 ⁻¹⁰)	1.6x10 ⁻¹⁰ (2.6x10 ⁻¹¹)
Nb11-tri	I	Biphasic	Biphasic	Biphasic	NP	NP	NP	NP	5.1x10 ⁻⁸ (1.6x10 ⁻⁸)	NP
ACE2-Fc	N/A	NP	NP	NP	NP	NP	NP	5.3x10 ⁻⁹ (2.5x10 ⁻⁹)	4.0x10 ⁻⁸ (8.8x10 ⁻⁹)	2.6x10 ⁻⁸ (8.5x10 ⁻⁹)
mNb6-tri (15 aa)	I	1.6x10 ⁶	<1.0x10 ⁻⁶	<1.0x10 ⁻¹²	NP	NP	NP	4.3x10 ⁻¹⁰ (1.4x10 ⁻¹⁰)	2.9x10 ⁻¹⁰ (4.2x10 ⁻¹¹)	NP
mNb6-tri (20 aa)	I	1.4x10 ⁶	<1.0x10 ⁻⁶	<1.0x10 ⁻¹²	NP	NP	NP	4.0x10 ⁻¹⁰ (1.7x10 ⁻¹⁰)	1.2x10 ⁻¹⁰ (2.8x10 ⁻¹¹)	5.4x10 ⁻¹¹ (1.0x10 ⁻¹¹)

^aAverage values from n = 5 biological replicates for Nb6, Nb11, Nb15, Nb19 are presented, all others were tested with n = 3 biological replicates.

^bAverage values from n = 2 biological replicates for Nb12, Nb17, and Nb11-tri are presented, all others were tested with n = 3 biological replicates.

^cAverage values from n = 2 biological replicates for Nb3, Nb3-bi, and Nb3-tri. n = 3 biological replicates for all others.

^dNb3, Nb17, and Nb18 expresses at 41.3, 4.0, and 2.2 milligrams per liter of *E. coli* culture, respectively. Nb3 is monodisperse on size exclusion chromatography over a GE S200 Increase 10x300 column, while Nb17 and Nb18 are polydisperse.

NB – no binding

NC – no competition

NP – not performed

Table 2.2 Cryo-electron microscopy data collection, refinement, and validation statistics

Sample: Spike ^{S2P} conformation: EMDB: PDB:	Spike ^{S2P} -Nb6		Spike ^{S2P} -Nb11		Spike ^{S2P} -mNb6
	Open EMD-22908	Closed EMD-22907 7KKK	Open EMD-22911	Closed EMD-22909	Closed EMD-22910 7KKL
Data collection and processing	Titan Krios/Gatan K3 with Gatan Bioquantum Energy Filter				
Microscope/Detector	SerialEM, 3x3 image shift				
Imaging software and collection	105,000				
Magnification	300				
Voltage (kV)	66				
Electron exposure (e ⁻ /Å ²)	8				
Dose rate (e ⁻ /pix/sec)	0.55				
Frame exposure (e ⁻ /Å ²)	-0.8 to -2.0				
Defocus range (µm)	0.834 (physical)				
Pixel size (Å)	5,317		4,103		1,609
Micrographs					
Reconstruction					
Autopicked particles (template-based in cryosparc)	2,033,067		1,204,855		585,250
Particles in final refinement	40,125 (cisTEM)	58,493 (cisTEM)	21,570 (cisTEM)	27,611 (RELION)	53,690 (cisTEM)
Symmetry imposed	C1	C3	C1	C1	C3
Map sharpening <i>B</i> factor (Å ²)		-90			-140
Map resolution, global FSC (Å)					
FSC 0.5, unmasked/masked	7.8/4.6	4.1/3.4	7.0/4.4	7.6/5.3	3.9/3.3
FSC 0.143, unmasked/masked	4.7/3.8	3.5/3.0	4.3/3.7	5.1/4.2	3.2/2.9
Refinement					
Initial model used (PDB code)	6VXX, 3P0G				6VXX, 3P0G
Model resolution (Å)					
FSC 0.5, unmasked/masked		3.5/3.2			3.2/2.9
Model composition					
Non-hydrogen atoms		26871			27024
Protein residues		3360			3360
Glycans (NAG)		54			63
<i>B</i> factors (Å ²)					
Protein		93			62
Ligand		77			88
R.m.s. deviations					
Bond lengths (Å)		0.020			0.020
Bond angles (°)		1.70			1.64
Validation					
MolProbity score		0.72			0.76
Clashscore		0.66			0.83
Poor rotamers (%)		0.21			0.41
EMRinger score		3.30			4.22
CaBLAM score		2.31			1.85
Ramachandran plot					
Favored (%)		98.15			98.55
Allowed (%)		1.85			1.45
Disallowed (%)		0			0

Table 2.3. X-ray crystallography data collection and refinement statistics

	mNb6 (PDB: 7KKJ)
Data collection	
Space group	$P2_1$
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	44.56, 71.25, 46.43
α , β , γ (°)	90, 114.93, 90
Total Reflections	104,195
Unique Reflections	16,204 (1582)
Molecules in asymmetric unit	2
Resolution (Å)	71.25 - 2.05 (2.09 - 2.05) ^a
R_{sym} or R_{merge}	0.13 (0.94) ^b
R_{pim}	0.055 (0.396)
$I / \sigma I$	7.2 (0.9)
Completeness (%)	97.8 (96.6)
Redundancy	6.4 (5.7)
CC (1/2) (%)	99.8 (64.4)
Refinement	
Resolution (Å)	42.1 – 2.05
No. reflections	16,204
$R_{\text{work}} / R_{\text{free}}$ (%)	21.2 / 24.8
No. atoms	
Protein	1880
Ligand/ion	21
Water	131
<i>B</i> -factors	
Protein	33
Ligand/ion	58
Water	39
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.14
Ramachandran plot	
Allowed (%)	97.85
Generous (%)	2.15
Disallowed (%)	0

^a Values in parentheses correspond to the highest resolution shell.

^b $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$

Table 2.4. Nanobody expression plasmids

Plasmid	Nanobody	Plasmid backbone	Resistance Marker
pPW3544	Nb2	pet-26b(+)	kanamycin
pPW3545	Nb3	pet-26b(+)	kanamycin
pPW3546	Nb6	pet-26b(+)	kanamycin
pPW3547	Nb8	pet-26b(+)	kanamycin
pPW3548	Nb11	pet-26b(+)	kanamycin
pPW3549	Nb12	pet-26b(+)	kanamycin
pPW3550	Nb15	pet-26b(+)	kanamycin
pPW3551	Nb16	pet-26b(+)	kanamycin
pPW3552	Nb17	pet-26b(+)	kanamycin
pPW3553	Nb18	pet-26b(+)	kanamycin
pPW3554	Nb19	pet-26b(+)	kanamycin
pPW3555	Nb24	pet-26b(+)	kanamycin
pPW3557	Trivalent Nb6, 20AA length GS linker	pet-26b(+)	kanamycin
pPW3558	Trivalent Nb3, 15AA length GS linker	pet-26b(+)	kanamycin
pPW3559	Trivalent Nb11, 15AA length GS linker	pet-26b(+)	kanamycin
pPW3560	Bivalent Nb3, 15AA length GS linker	pet-26b(+)	kanamycin
pPW3561	Bivalent Nb6, 15AA length GS linker	pet-26b(+)	kanamycin
pPW3563	Trivalent mNb6, 20AA length GS linker	pet-26b(+)	kanamycin
pPW3564	mNb6	pet-26b(+)	kanamycin

2.7 Materials and Methods

Expression and purification of SARS-CoV-2 Spike, RBD, and ACE2.

We used a previously described construct to express and purify the pre-fusion SARS-CoV-2 Spike ectodomain¹⁵ (Spike^{S2P}). ExpiCHO or Expi293T cells (ThermoFisher) were transfected with the Spike^{S2P} construct per the manufacturer's instructions for the MaxTiter protocol and harvested between 3-9 days after transfection. Clarified cell culture supernatant was loaded onto Ni-Excel beads (Cytiva) followed by extensive washes in 20 mM HEPES pH 8.0, 200 mM sodium chloride, and 10 mM imidazole and elution in the same buffer supplemented with 500 mM imidazole. Spike^{S2P} was concentrated using a 100 kDa MWCO spin concentrator (Millipore) and further purified by size exclusion chromatography over a Superose 6 Increase 10/300 column (GE Healthcare) in 20 mM HEPES pH 8.0 and 200 mM sodium chloride. All purification steps were performed at room temperature. The resulting fractions for trimeric Spike^{S2P} were pooled and either used directly for cryo-EM studies or concentrated and flash frozen in liquid nitrogen with 15% glycerol for other biochemical studies. We used a previously described construct to express and purify the SARS-CoV-2 Receptor binding domain³⁷ (RBD). Expi293T cells (ThermoFisher) were transfected with the RBD construct per the manufacturer's instructions and harvested between 3-6 days after transfection. Clarified cell culture supernatant was loaded onto Ni-Excel beads (Cytiva) or a His-Trap Excel column (GE Healthcare) followed by washes in 20 mM HEPES pH 8.0, 200 mM sodium chloride, and 10 mM imidazole and elution using the same buffer supplemented with 500 mM imidazole. RBD was concentrated using a 30 kDa MWCO spin concentrator (Millipore) and further purified by size exclusion chromatography over a Superdex 200 Increase 10/300 GL column (GE Healthcare) in 20 mM HEPES pH 8.0 and 200 mM sodium chloride. The resulting fractions were pooled, concentrated, and flash frozen in liquid nitrogen

with 10% glycerol. For biochemical and yeast display experiments, Spike^{S2P} and RBD were labeled with freshly prepared stocks of Alexa 647-NHS, Alexa 488-NHS, or Biotin-NHS (ThermoFisher) with a 5-fold stoichiometry for 1 hour at room temperature followed by quenching of NHS with 10 mM Tris pH 8.0 for 60 minutes. Labeled proteins were further purified by size exclusion chromatography, concentrated using a spin concentrator (Millipore), and flash frozen in liquid nitrogen with 10-15% glycerol. We used an ACE2-ECD (18-614) Fc fusion expression plasmid to express and purify Fc tagged ACE2-ECD³⁸. Expi293T cells (ThermoFisher) were transfected with the ACE2-Fc construct per the manufacturer's instructions and harvested between 5-7 days after transfection. Clarified cell culture supernatant was loaded onto a MabSelect Pure 1 mL Column (GE Healthcare). Column was washed with Buffer A (20 mM HEPES pH 7.5, 150 mM NaCl) and protein was eluted with Buffer B (100 mM Sodium Citrate pH 3.0, 150 mM NaCl) into a deep well block containing 1 M HEPES pH 7.5 to neutralize the acidic elution. ACE2-Fc was concentrated using a 30 kDa MWCO spin concentrator (Millipore) and further purified by size exclusion chromatography over a Superdex 200 Increase 10/300 GL column (GE Healthcare) in SEC Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5% v/v Glycerol). The resulting fractions were pooled, concentrated, and flash frozen in liquid nitrogen. To obtain monomeric ACE2, 1:50 (w/w) His-tagged TEV protease was added to ACE2-Fc and incubated at 4 °C overnight. This mixture was then purified by size exclusion chromatography in SEC Buffer. Monomeric ACE2 fractions were pooled and washed with His-resin (1 mL of 50% slurry) to remove excess TEV. The resulting supernatant was pooled, concentrated, and flash frozen in liquid nitrogen.

Identification of anti SARS-CoV2 Spike nanobodies

To identify nanobodies against the SARS-CoV-2 Spike ECD, we used a yeast surface displayed library of synthetic nanobody sequences that recapitulate amino acid position specific-variation in natural llama immunological repertoires. This library encodes a diversity of $>2 \times 10^9$ variants, and uses a synthetic stalk sequence for nanobody display, as described previously in a modified vector encoding nourseothricin (NTC) resistance¹⁷. For the first round of selection, 2×10^{10} yeast induced in YPG (Yeast Extract-Peptone-Galactose) supplemented with NTC were washed repeatedly in selection buffer (20 mM HEPES, pH 7.5, 150 mM sodium chloride, 0.1% (w/v) low biotin bovine serum albumin, BSA) and finally resuspended in 10 mL of selection buffer containing 200 nM biotinylated-Spike^{S2P}. Yeast were incubated for 30 minutes at 25 °C, then washed repeatedly in cold selection buffer, and finally resuspended in 10 mL of cold selection buffer containing 200 μ L of Miltenyi anti-Streptavidin microbeads. After 30 minutes of incubation at 4 °C, yeast were again washed with cold selection buffer. Spike^{S2P} binding yeast were captured on a Miltenyi MACS LS column and recovered in YPD (Yeast Extract-Peptone-Dextrose) medium supplemented with NTC. For round 2, 4×10^8 induced yeast from Round 1 were incubated with 100 nM Spike^{S2P} labeled with Alexa647 in 1 mL of selection buffer for 1 hr at 25 °C. After extensive washes with cold selection buffer, Spike^{S2P} binding yeast were isolated by fluorescence activated cell sorting (FACS) on a Sony SH800 instrument. A similar approach was used for round 3, with substitution of 10 nM Spike^{S2P} labeled with Alexa647. Post round 3 yeast were plated on YPD+NTC solid media and 768 individual colonies were induced with YPG+NTC media in 2 mL deep well plates. Each individual clone was tested for binding to 4 nM Spike^{S2P}-Alexa488 by flow cytometry on a Beckman Coulter Cytotflex. To identify nanobodies that disrupt Spike-ACE2 interactions, Spike^{S2P}

binding was repeated in the presence of 0.5-1 μ M ACE2-Fc. Out of 768 clones, we identified 21 that strongly bind Spike^{S2P} and are competitive with ACE2 (**Table 2.4**).

Expression and purification of nanobodies

Nanobody sequences were cloned into the pET26-b(+) expression vector using In-Fusion HD cloning (Takara Bio), transformed into BL21(DE3) E. coli (New England BioLabs), grown in Terrific Broth at 37 °C until OD 0.7-0.8, followed by gene induction using 1 mM IPTG for 18- 22 hours at 25°C. E. Coli were harvested and resuspended in SET Buffer (200 mM Tris, pH 8.0, 500 mM sucrose, 0.5 mM EDTA, 1X cOmplete protease inhibitor (Roche)) for 30 minutes at 25 °C before a 45 minute osmotic shock with a two-fold volume addition of water. NaCl, MgCl₂, and imidazole were added to the lysate to 150 mM, 2 mM, and 40 mM respectively before centrifugation at 17-20,000xg for 15 minutes to separate cell debris from the periplasmic fraction. For every liter of bacterial culture, the periplasmic fraction was then incubated with 4 mL of 50% HisPur Ni-NTA resin (Thermo Scientific) which had been equilibrated in Nickel Wash Buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 40 mM imidazole). This mixture was incubated for 1 hr with rotation at RT before centrifugation at 50xg to collect the resin. The resin was then washed with 5 volumes of Nickel Wash buffer 3 times, each time using centrifugation to remove excess wash buffer. Bound proteins were then eluted using three washes with Elution Buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 500 mM imidazole). The eluted protein was concentrated using a 3.5 kDa MWCO centrifugal filter unit (Amicon) before injection onto a Superdex 200 Increase 10/300 GL column equilibrated with 20 mM HEPES, pH 7.5, 150 mM NaCl. Nanobody constructs were concentrated again using a 3.5k MWCO centrifugal filter unit, and flash frozen in liquid nitrogen.

Affinity determination by surface plasmon resonance

Nanobody (Nb) affinity determination experiments were performed on Biacore T200 and 8K instruments (Cytiva Life Sciences) by capturing the StreptagII-tagged Spike^{S2P} at 10 µg/mL on a StreptactinXT-immobilized (Iba Life Sciences) CM5 Series S sensor chip (Cytiva Life Sciences) to achieve maximum response (R_{max}) of approximately 30 response units (RUs) upon nanobody binding. 2-fold serial dilutions of purified nanobody from 1 µM to 31.25 nM (for monovalent constructs) or from 50 nM to 1.56 nM (for affinity matured and multimeric constructs) were flowed over the captured Spike^{S2P} surface at 30 µL/minute for 60 seconds followed by 600 seconds of dissociation flow. Following each cycle, the chip surface was regenerated with 3 M guanidine hydrochloride. Separately, biotinylated SARS-CoV-2 RBD at 8 µg/mL was loaded onto a preconditioned Series S Sensor Chip CAP chip (Cytiva Life Sciences) to achieve an R_{max} of approximately 60 RUs upon nanobody binding. 2-fold serial dilutions in the same running buffer and sample series (parent or affinity matured clone) as the Spike^{S2P} runs were flowed over the RBD surface at 30 µL/minute for 60 seconds followed by 600 seconds of dissociation flow. Chip surface regeneration was performed with a guanidine hydrochloride/sodium hydroxide solution. The resulting sensorgrams for all monovalent clones were fit to a 1:1 Langmuir binding model using the Biacore Insight Evaluation Software (Cytiva Life Sciences) or the association/dissociation model in GraphPad Prism 8.0. For determination of kinetic parameters for Nb6-bi and Nb6-tri binding, the dissociation phase was fit to a biexponential decay constrained to two dissociation rate constants shared between each concentration. The association phase was fit separately using an association kinetics model simultaneously fitting the association rate constant for each concentration. For nanobody competition experiments, Spike^{S2P} was loaded onto a StreptactinXT-immobilized CM5 sensor chip as previously described. As in the kinetics

experiments, the primary nanobody was flowed over the captured Spike^{S2P} surface for 60 seconds at 30 μ L/minute to achieve saturation. Immediately following this, a second injection of a mixture of primary and variable nanobody at the same concentration as in the primary injection was performed.

ACE2 cellular surface binding competition assays

A dilution series of nanobody was generated in PBE (PBS + 0.5% (w/v) BSA + 2 mM EDTA) and mixed with Spike^{S2P}-Alexa647 or RBD-Alexa647. ACE2 expressing HEK293T cells were dissociated with TrypLE Express (ThermoFisher) and resuspended in PBE²². The cells were mixed with the Spike^{S2P}-nanobody solution and incubated for 45 minutes, washed in PBE, and then resuspended in PBE. Cell surface Alexa647 fluorescence intensity was assessed on an Attune Flow Cytometer (ThermoFisher).

Affinity maturation of Nb6

A site saturation mutagenesis library of Nb6 was generated by assembly PCR of overlapping oligonucleotides encoding the Nb6 sequence. Individual oligos for each position in CDR1, CDR2, and CDR3 were designed with the degenerate “NNK” codon. The assembled gene product was amplified with oligonucleotides with overlapping ends to enable homologous recombination with the yeast surface display vector as previously described and purified with standard silica-based chromatography¹⁷. The resulting insert DNA was transformed into *Saccharomyces cerevisiae* strain BJ5465 (ATCC 208289) along with the yeast display vector pYDS2.0 to generate a library of 2×10^8 transformants. After induction in YPD+NTC medium at 20 °C for 2 days, 2×10^9 yeast were washed in selection buffer (20 mM HEPES, pH 8.0, 150 mM sodium chloride, 0.1% (w/v)

low biotin BSA) and incubated with 1 nM biotin-Spike^{S2P} for 1 hour at 25 °C. Yeast were subsequently washed in selection buffer, resuspended in 1 mL selection buffer, and incubated with 10 µL streptavidin microbeads (Miltenyi) for 15 min. at 4 °C. Yeast were washed again with cold selection buffer and Spike^{S2P}-binding yeast were isolated by magnetic separation using an LS column (Miltenyi). Recovered yeast were grown in YPD+NTC at 37 °C and induced in YPG+NTC at 20 °C. A second round of selection was performed as above, substituting 100 pM RBD-Alexa647 as the antigen. Yeast displaying high affinity clones were selected by magnetic separation using Anti-Cy5 microbeads (Miltenyi) and an LS column. Analysis of the library after the second round of selection revealed a population of clones with clear binding of 10 pM RBD-Alexa647. Therefore, 96 individual clones were screened for binding to 10 pM RBD-Alexa647 by flow cytometry. Sequence analysis of eight clones that showed robust binding to 10 pM RBD-Alexa647 revealed two consensus mutations, I27Y and P105Y, which were used to generate the affinity matured clone mNb6.

mNb6 crystallography and structure determination

Purified mNb6 was concentrated to 18.7 mg/mL and filtered using 0.1 µm hydrophilic PVDF filters (Millipore). mNb6 crystal screens were set up in 96 well plates in hanging drop format at 2:1 protein:reservoir in Index and AmSO4 screens (Hampton Research, Aliso Viejo, CA). Crystals in over 60 different screening conditions with various morphologies appeared overnight at ambient temperature and were obtained directly from the screens without further optimization. The crystals were cryoprotected by quick dipping in a solution containing 80% reservoir and 20% PEG400 or 20% Glycerol, then mounted in CrystalCap HT Cryoloops (Hampton Research, Aliso Viejo, CA) and flash cooled in a cryogenic nitrogen stream (100 K). All data were collected at the Advanced

Light Source (Berkeley, CA) beam line 8.3.1. A single crystal of mNb6 that grew in 0.1 M Tris.HCl pH 8.5, 1.0 M Ammonium sulfate diffracted to 2.05 Å. Integration, and scaling were performed with Xia2, using XDS for indexing and integration and XSCALE for scaling and merging³⁹. The structure was solved molecular replacement using PHASER using the structure of nanobody, Nb.b201 (PDB 5VNV) as search model^{17,40}. Model building was performed with COOT and refined with PHENIX and BUSTER^{41,42,43}.

Structures of Spike-nanobody complexes by cryo-EM

Sample preparation and microscopy

To prepare Spike^{S2P}-nanobody complexes, each nanobody was incubated on ice at a 3-fold molar excess to Spike^{S2P} at 2.5 μM for 10 minutes. 3 μL of Spike^{S2P}-nanobody complex was added to a 300 mesh 1.2/1.3R Au Quantifoil grid previously glow discharged at 15 mA for 30 seconds with a Pelco easiGlow Glow discharge cleaning system. Using Whatman No.1 qualitative filter paper, blotting was performed with a blot force of 0 for 4 seconds at 4°C and 100% humidity in a FEI Vitrobot Mark IV (ThermoFisher) prior to plunge freezing into liquid ethane.

For each complex, 120-frame super-resolution movies were collected with a 3x3 image shift collection strategy at a nominal magnification of 105,000x (physical pixel size: 0.834 Å/pix) on a Titan Krios (ThermoFisher) equipped with a K3 camera and a Bioquantum energy filter (Gatan) set to a slit width of 20 eV. Collection dose rate was 8 e⁻/pixel/second for a total dose of 66 e⁻/Å². Each collection was performed with semi-automated scripts in SerialEM⁴⁴.

Image Processing

For all datasets, dose fractionated super-resolution movies were motion corrected with MotionCor2⁴⁵. Contrast transfer function determination was performed with cryoSPARC patch CTF⁴⁶. Particles were picked with a 20 Å low-pass filtered apo Spike 2D templates generated from a prior data collection. Nb6-Spike^{S2P} and mNb6-Spike^{S2P} particles were extracted with a 384-pixel box, binned to 96 pixels and subject to single rounds of 2D and 3D classification prior to unbinning for homogenous refinement in cryoSPARC. Using pyEM modules, refined particles were then imported into Relion3.1 for 3D classification without alignment using the input refinement map low pass filtered to 40 Å^{47, 48}. Particles in classes representing the closed conformation of Spike were imported into cisTEM and subject to autorefinement followed by local refinement within a RBD:nanobody masked region⁴⁹. Following local refinement, a new refinement package symmetrized to the C3 axis was created for a final round of local refinement without masking. Final particle counts for each map are as follows: Nb6-Open: 40,125, Nb6-Closed: 58,493, mNb6: 53,690. Nb11-Spike^{S2P} particles were extracted with a 512-pixel box, binned to 128 pixels for multiple rounds of 3D classification as described in **Figure 2.7**. Following homogenous refinement, particles were exported to Relion3.1. Particle density roughly corresponding to RBD-nanobody complexes was retained post-particle subtraction. 3D classification without alignment was performed on the particle subtracted stacks. Particles in classes with robust RBD-nanobody density were selected, unsubtracted and refined in Relion followed by post-processing. 21,570 particles contributed to the final maps. Final particle counts for each map are as follows: Nb11-Open: 21,570, Nb11-Closed: 27,611. For all maps, final local resolution estimation and GSFSC determination was carried out in cryoSPARC. Viewing angle distribution plots were generated with pyEM modules and visualized with ChimeraX⁵⁰.

Structure modeling

Models of Nb6-Spike^{S2P} and mNb6-Spike^{S2P} were built using a previously determined structure of closed Spike^{S2P} 14 (PDB: 6VXX). A composite model incorporating resolved regions of the RBD was made using a previously determined X-ray crystal structure of the SARS-CoV-2 RBD⁵¹ (PDB: 6M0J). For Nb6, the beta2-adrenergic receptor nanobody Nb80 (PDB: 3P0G) was used as a template to first fit the nanobody into the cryo-EM density map for the Nb6-Spike^{S2P} complex⁵². Complementarity determining loops were then truncated and rebuilt using RosettaES⁵³. The higher resolution structure of mNb6 enabled manual building of nanobody CDR loops de novo, and therefore the Rosetta-based approach was not used for modeling. The final structures were inspected and manually adjusted in COOT and ISOLDE, followed by real space refinement in PHENIX^{41, 43, 54} and further refined and relaxed using Rosetta⁵⁵. Glycans were refined utilizing the glycan specific Rosetta protocol that incorporates prior knowledge on carbohydrate conformations to ensure lowest energy glycan geometries⁵⁶. Final glycan placement was inspected manually and using the Privateer software package distributed under CCP4^{57, 58}. Final protein models were analyzed with Molprobit⁵⁹, EMRinger⁶⁰, PHENIX, with statistics reported in **Table 2.2**. For models of Nb11-Spike^{S2P} complexes presented here, the closest nanobody by sequence in the PDB (beta2-adrenergic receptor Nb60, PDB ID: 5JQH) was fit by rigid-body refinement in COOT into the cryo-EM density map using only the framework regions⁶¹. While the lower resolution of these maps precluded confident assignment of loop conformations, the overall orientation of Nb11 relative to Spike^{S2P} was well constrained, enabling accurate modeling of distances between the N- and C- termini of two Nb11 molecules bound to Spike^{S2P}.

Radiolytic hydroxyl radical footprinting and mass-spectrometry of apo and Nb3-bound Spike^{S2P}

Spike^{S2P} and Nb3 samples were buffer exchanged into 10 mM phosphate buffer (pH 7.4) by extensive dialysis at 25 °C. A 1.5-fold molar excess of Nb3 was added to 5 μM Spike^{S2P} and the complex was incubated for >24 hr at 25 °C. For radiolytic footprinting, protein concentrations and beam parameters were optimized using an Alexa-488 fluorophore assay¹⁸. Apo Spike^{S2P} and Spike^{S2P}-Nb3 complex at concentrations of 1-3 μM were exposed to a synchrotron X-ray white beam at 6 timepoints between 0-50 ms at beamline 3.2.1 at the Advanced Light Source in Berkeley, CA and were quenched with 10 mM methionine amide immediately post-exposure. Glycans were removed by treatment with 5% SDS, 5 mM DTT at 95 °C for five minutes and subsequent PNGase (Promega) digestion at 37°C for 2 hours. Samples were buffer exchanged into ammonium bicarbonate (ABC) buffer (pH 8.0) using ZebaSpin columns (Thermo Fisher). Alkylation of cysteines was achieved by treatment with 8 M urea and 5 mM DTT at 37°C for 30 minutes followed by an incubation with 15 mM iodoacetamide at 25 °C in the dark for 30 minutes. All samples were further buffer exchanged to ABC pH 8.0 using ZebaSpin columns and digested with either Trypsin/Lys-C or Glu-C (Promega) at an enzyme:protein ratio of 1:20 (w/w) at 37 °C for 8 hours. Samples were lyophilized and resuspended in 1% formic acid at 200 fmol/μL concentration. For each MS analysis, 1 μL of sample was injected onto a 5 mm Thermo Trap C18 cartridge, and then separated over a 15 cm column packed with 1.9 μm Reprosil C18 particles (Dr. Maisch HPLC GmbH) by a nanoElute HPLC (Bruker). Separation was performed at 50 °C and a flow rate of 400 μL/min by the following gradient in 0.1% formic acid: 2% to 17% acetonitrile from 0 to 20 min, followed by 17% to 28% acetonitrile from 20 to 40 min. The eluent was electrospray ionized into a Bruker timsTOF Pro mass spectrometer and data was collected using data-dependent PASEF

acquisition. Database searching and extraction of MS1 peptide abundances was performed using the FragPipe platform with either trypsin or GluC enzyme specificity, and all peptide and protein identifications were filtered to a 1% false-discovery rate⁶². Searches were performed against a concatenated protein database of the Spike protein, common contaminant proteins, and the *Saccharomyces cerevisiae* proteome (downloaded July 23, 2020). Note, the *Saccharomyces cerevisiae* proteome was included to generate a sufficient population of true negative identifications for robust false discovery rate estimation of peptide and protein identifications. Lastly, the area under the curve MS1 intensities reported from FragPipe were summarized for each peptide species using MSstats⁶³.

The peak areas of extracted ion chromatograms and associated side-chain modifications were used to quantify modification at each timepoint. Increasing beamline exposure time decreases the fraction of unmodified peptide and can be represented as a site-specific dose-response plot. The rate of hydroxyl radical reactivity (k_{fp}) is dependent on both the intrinsic reactivity of each residue and its solvent accessibility and was calculated by fitting the dose-response to a pseudofirst order reaction scheme in Graphpad Prism Version 8. The ratio of k_{fp} between apo Spike^{S2P} and the Spike-Nb3 complex at specific residues gave information on solvent accessibility changes between the two samples. These changes were mapped onto the SARS-CoV-2 Spike¹¹ (PDB 6XR8). In some cases, heavily modified residues show a flattening of dose-response at long exposures which we interpret as radical induced damage. These over-exposed timepoints were excluded from the calculation of k_{fp} .

Pseudovirus assays for nanobody neutralization

ZsGreen SARS-CoV-2-pseudotyped lentivirus was generated according to a published protocol²². The day before transduction, 50,000 ACE2 expressing HEK293T cells were plated in each well of a 24-well plate. 10-fold serial dilutions of nanobody were generated in complete medium (DMEM + 10% FBS + PSG) and pseudotyped virus was added to a final volume of 200 μ L. Media was replaced with nanobody/pseudotyped virus mixture for four hours, then removed. Cells were washed with complete medium and then incubated in complete medium at 37 °C. Three days post-transduction, cells were trypsinized and the proportion of ZsGreen⁺ cells was measured on an Attune flow cytometer (ThermoFisher).

Authentic SARS-CoV-2 neutralization assay

SARS-CoV-2, isolate France/IDF0372/2020, was supplied by the National Reference Centre for Respiratory Viruses hosted by Institut Pasteur (Paris, France) and headed by Pr. Sylvie van der Werf. Viral stocks were prepared by propagation in Vero E6 cells in Dulbecco's modified Eagle's medium (DMEM) supplemented with 2% (v/v) fetal bovine serum (FBS, Invitrogen). Viral titers were determined by plaque assay. All plaque assays involving live SARS-CoV-2 were performed at Institut Pasteur Paris (IPP) in compliance with IPP's guidelines following Biosafety Level 3 (BSL-3) containment procedures in approved laboratories. All experiments were performed in at least three biologically independent samples. Neutralization of infectious SARS-CoV-2 was performed using a plaque reduction neutralization test in Vero E6 cells (CRL-1586, ATCC). Briefly, nanobodies (or ACE2-Fc) were eight-fold serially diluted in DMEM containing 2% (v/v) FBS and mixed with 50 plaque forming units (PFU) of SARS-CoV-2 for one hour at 37°C, 5% CO₂. The mixture was then used to inoculate Vero E6 cells seeded in 12-well plates, for one hour

at 37 °C, 5% CO₂. Following this virus adsorption time, a solid agarose overlay (DMEM, 10% (v/v) FBS and 0.8% agarose) was added. The cells were incubated for a further 3 days prior to fixation using 4% formalin and plaques visualized by the addition of crystal violet. The number of plaques in quadruplicate wells for each dilution was used to determine the half maximal inhibitory concentrations (IC₅₀) using 3- parameter logistic regression (GraphPad Prism version 8).

Nanobody stability studies

Nanobody thermostability by circular dichroism was assessed using a Jasco J710 CD spectrometer equipped with a Peltier temperature control. Individual nanobody constructs were diluted to 5 µM in phosphate buffered saline. Molar ellipticity was measured at 204 nm (2 nm bandwidth) between 25 °C and 80 °C with a 1 °C/min heating rate. The resulting molar ellipticity values were normalized and plotted in GraphPad Prism 8.0 after applying a nearest neighbor smoothing function. For nanobody competition experiments on ACE2 expressing HEK293T cells, nanobodies were incubated at either 25°C or 50°C for one hour. Alternatively, each nanobody was aerosolized with a portable mesh nebulizer producing 2-5 µm particles at a final concentration of 0.5 mg/mL. The resulting aerosol was collected by condensation into a 50 mL tube cooled on ice. Samples were then treated as indicated above to determine IC₅₀ values for binding to Spike^{S2P}-Alexa647 or used for pseudovirus neutralization studies as described above. Further experiments assessing mNb6 and mNb6-tri stability to aerosolization and lyophilization used a starting concentration of 0.5 mg/mL of each construct. Aerosolization was performed as described above. For lyophilization, nanobodies were first flash frozen in liquid nitrogen and the solution was dried to completion under vacuum. The resulting dried material was resuspended in 20 mM HEPES pH 7.5, 150 mM NaCl. Size exclusion chromatography of the unstressed, post-aerosolization, and

post-lyophilization samples were performed on a Superdex 75 Increase 10/300 column in 20 mM HEPES pH 7.5, 150 mM NaCl. SPR experiments to assess binding to Spike^{S2P} were performed as described above. For live SARS-CoV-2 virus experiments, aerosolized, lyophilized, or heat-treated samples were flash frozen in liquid nitrogen prior to shipping.

2.8 Acknowledgements

We thank the entire Walter and Manglik labs for facilitating the development and rapid execution of this large-scale collaborative effort. We thank S. Bernales and T. De Fougères for advice and helpful discussion and J. Weissman for input into the project and reagent and machine use. We thank J. Wells for providing the ACE2 ECD-Fc construct; J. McLellan for providing the Spike, RBD, and ACE2 constructs; and F. Krammer for providing an RBD construct. We thank J. Bloom for providing the ACE2-expressing HEK293T cells as well as the plasmids for pseudovirus work. We thank G. Meigs and other Beamline staff at ALS 8.3.1 for their help in data collection. We thank R. A. Albrecht for oversight of the conventional BSL3 biocontainment facility at the Icahn School of Medicine at Mount Sinai.

2.9 Author Contributions

M.Sc. purified Spike^{S2P}, RBD, and ACE2 proteins; performed yeast display selections to identify and affinity mature nanobodies; expressed and purified nanobodies; tested activity in cell-based assays; cloned, expressed, and purified multivalent nanobody constructs; and coordinated live virus experiments. B.F. purified and characterized Spike^{S2P} protein and candidate nanobodies; developed, performed, and analyzed SPR experiments for Spike^{S2P} and RBD-nanobody affinity determination; developed, performed, and analyzed SPR binning and experiments; determined

optimal freezing conditions for cryo-EM experiments; and processed, refined, and generated figures for Nb6, Nb11, and mNb6 EM datasets. R.A.S. expressed and purified ACE2 and nanobodies and developed and performed cell-based assays for inhibition of Spike^{S2P} binding and pseudovirus assays for determining nanobody efficacy. S.S. expressed and purified Spike^{S2P}, RBD, ACE2-Fc, and nanobodies; processed cryo-EM data; optimized RBD-nanobody complexes for crystallography; grew crystals of mNb6; collected diffraction data; and refined the x-ray crystal structure of mNb6. V.R. tested efficacy of nanobody constructs in live SARS-CoV-2 infection assays under the guidance of M.V. N.H. purified nanobodies; developed, performed, and analyzed SPR binning experiments; developed, performed, and analyzed variable Nb6-bi and Nb6-tri association experiments; and performed thermal melting stability assays for nanobody constructs. M.B. developed approaches to express and purify nanobodies from *Pichia pastoris* and developed, performed, and analyzed approaches to quantify nanobody efficacy in live virus assays. C.B.B. expressed and purified Spike^{S2P}, generated the affinity maturation library for Nb6, performed yeast display selections to identify mNb6, and built the synthetic yeast nanobody library with J.L. I.D. expressed and purified nanobody constructs. B.S.Z. performed live SARS-CoV-2 virus assays to test nanobody efficacy with guidance from QCRG Structural Biology Consortium member O. Rosenberg. C.R.S. and K.L. performed live SARS-CoV-2 virus assays to test nanobody efficacy with guidance from M.O. K.M.W. performed live SARS-CoV-2 virus assays to test nanobody efficacy with guidance from A.G.-S. A.W.B.-H. performed SPR experiments. A.A.A., N.D., B.B.-R., and Yu.L. assisted in cloning, expression, and purification of nanobody and pseudovirus constructs. V.B. performed single-molecule nanobody-Spike^{S2P} interaction studies. S.N. prepared media and coordinated lab usage during UCSF's partial shutdown. M.Z. and S.G. performed radiolytic footprinting experiments with guidance from C.Y.R. and analyzed mass spectrometry

data generated by D.L.S. Several members of the QCRG Structural Biology Consortium played an exceptionally important role for this project: C.M.A. and C.P. determined optimal freezing conditions for cryo-EM experiments, optimized data collection approaches, and collected cryo-EM datasets. A.F.B., A.N.R., A.M.S., F.M., D.B., and T.P. collected cryo-EM data on Spike^{S2P}-nanobody complexes. S.D., H.C.N., C.M.C., U.S.C., M.G., M.J., F.L., Ya.L. G.E.M., K.Z., and M.Su. analyzed cryo-EM data from 15 Spike^{S2P}-nanobody complex datasets. H.T.K. set up crystallization trials of various RBD-nanobody complexes and crystallized, collected diffraction data for, and refined the mNb6 structure. M.C.T. collected, processed, and refined the mNb6 structure. R.T., D.D., and K.S. expressed and purified Spike^{S2P}, and S.P. purified RBD. A.M. expressed and purified Spike^{S2P}, labeled Spike^{S2P} for biochemical studies, designed selection strategies for nanobody discovery, cloned nanobodies for expression, designed affinity maturation libraries and performed selections, analyzed SPR data, and performed nanobody stability studies. The overall project was supervised by P.W. and A.M.

QCRG Structural Biology Consortium Authors

In addition to those listed explicitly in the author contributions, the structural biology portion of this work was performed by the QCRG (Quantitative Biosciences Institute Coronavirus Research Group) Structural Biology Consortium. Listed below are the contributing members of the consortium listed by teams in order of team relevance to the published work. Within each team the team leads are italicized (responsible for organization of each team, and for the experimental design utilized within each team), then the rest of team members are listed alphabetically. *CryoEM grid freezing/collection team: Caleigh M. Azumaya, Cristina Puchades,*

Ming Sun, Julian R. Braxton, Axel F. Brilot, Meghna Gupta, Fei Li, Kyle E. Lopez, Arthur Melo, Gregory E. Merz, Frank Moss, Joana Paulino, Thomas H. Pospiech, Jr., Sergei Pourmal, Alexandra N. Rizo, Amber M. Smith, Paul V. Thomas, Feng Wang, Zanlin Yu. *CryoEM data processing team: Miles Sasha Dickinson, Henry C. Nguyen, Daniel Asarnow, Julian R. Braxton, Melody G. Campbell, Cynthia M. Chio, Un Seng Chio, Devan Diwanji, Bryan Faust, Meghna Gupta, Nick Hoppe, Mingliang Jin, Fei Li, Junrui Li, Yanxin Liu, Gregory E. Merz, Joana Paulino, Thomas H. Pospiech, Jr., Sergei Pourmal, Smriti Sangwan, Tsz Kin Martin Tsui, Raphael Trenker, Donovan Trinidad, Eric Tse, Kaihua Zhang, Fengbo Zhou. Crystallography team: Nadia Herrera, Huong T. Kratochvil, Ursula Schulze-Gahmen, Michael C. Thompson, Iris D. Young, Justin Biel, Ishan Deshpande, Xi Liu. Mammalian cell expression team: Christian Bache Billesballe, Melody G. Campbell, Devan Diwanji, Carlos Nowotny, Amber M. Smith, Jianhua Zhao, Caleigh M. Azumaya, Alisa Bowen, Nick Hoppe, Yen-Li Li, Phuong Nguyen, Cristina Puchades, Mali Safari, Smriti Sangwan, Kaitlin Schaefer, Raphael Trenker, Tsz Kin Martin Tsui, Natalie Whitis. Protein purification team: Daniel Asarnow, Michelle Moritz, Tristan W. Owens, Sergei Pourmal, Caleigh M. Azumaya, Cynthia M. Chio, Amy Diallo, Bryan Faust, Meghna Gupta, Kate Kim, Joana Paulino, Jessica K. Peters, Kaitlin Schaefer, Tsz Kin Martin Tsui. Bacterial expression team: Amy Diallo, Meghna Gupta, Erron W. Titus, Jenny Chen, Loan Doan, Sebastian Flores, Mingliang Jin, Huong T. Kratochvil, Victor L. Lam, Yang Li, Megan Lo, Gregory E. Merz, Joana Paulino, Aye C. Thwin, Stephanie Wankowicz, Zanlin Yu, Yang Zhang, Fengbo Zhou. Infrastructure team: David Bulkley, Arceli Joves, Almarie Joves, Liam McKay, Mariano Tabios, Eric Tse. Leadership team: Oren S Rosenberg, Kliment A Verba, David A Agard, Yifan Cheng, James S Fraser, Adam Frost, Natalia Jura, Tanja Kortemme, Nevan J Krogan, Aashish Manglik, Daniel R. Southworth, Robert M Stroud. The QCRG Structural Biology*

Consortium has received support from: Quantitative Biosciences Institute, Defense Advanced Research Projects Agency HR0011-19-2-0020 (to D.A.Agard and K.A.Verba; B. Shoichet PI), FastGrants COVID19 grant (K.A.Verba PI), Laboratory For Genomics Research (O.S.Rosenberg PI) and Laboratory for Genomics Research LGR-ERA (R.M.Stroud PI). R.M.Stroud is supported by NIH grants AI 50476, GM24485.

2.10 References

1. Ksiazek, T. G. *et al.* A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *New England Journal of Medicine* vol. 348 1953–1966 (2003).
2. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. & Fouchier, R. A. M. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine* vol. 367 1814–1820 (2012).
3. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* vol. 579 270–273 (2020).
4. Chan, J. F.-W. *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet* vol. 395 514–523 (2020).
5. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* vol. 395 497–506 (2020).
6. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* vol. 579 265–269 (2020).
7. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* vol. 382 727–733 (2020).
8. Ke, Z. *et al.* Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* **588**, 498–502 (2020).
9. Turoňová, B. *et al.* In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* vol. 370 203–208 (2020).
10. Bosch, B. J., van der Zee, R., de Haan, C. A. M. & Rottier, P. J. M. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the

- Fusion Core Complex. *Journal of Virology* vol. 77 8801–8811 (2003).
11. Cai, Y. *et al.* Distinct conformational states of SARS-CoV-2 spike protein. *Science* vol. 369 1586–1592 (2020).
 12. Wang, Q. *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* vol. 181 894–904.e9 (2020).
 13. Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* vol. 367 1444–1448 (2020).
 14. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* vol. 181 281–292.e6 (2020).
 15. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* vol. 367 1260–1263 (2020).
 16. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* vol. 181 271–280.e8 (2020).
 17. McMahon, C. *et al.* Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nature Structural & Molecular Biology* vol. 25 289–296 (2018).
 18. Gupta, S., Feng, J., Chan, L. J. G., Petzold, C. J. & Ralston, C. Y. Synchrotron X-ray footprinting as a method to visualize water in proteins. *Journal of Synchrotron Radiation* vol. 23 1056–1069 (2016).
 19. Zhang, Y., Wecksler, A. T., Molina, P., Deperalta, G. & Gross, M. L. Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* vol. 28 850–858 (2017).
 20. Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the Spike

- protein of SARS-CoV-2. *Science* vol. 369 650–655 (2020).
21. Zhou, H. *et al.* Structural definition of a neutralization epitope on the N-terminal domain of MERS-CoV spike glycoprotein. *Nat. Commun.* **10**, 3068 (2019).
 22. Crawford, K. H. D. *et al.* Protocol and Reagents for Pseudotyping Lentiviral Particles with SARS-CoV-2 Spike Protein for Neutralization Assays. *Viruses* vol. 12 513 (2020).
 23. Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* vol. 369 1014–1018 (2020).
 24. Cao, Y. *et al.* Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* vol. 182 73–84.e16 (2020).
 25. Ju, B. *et al.* Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* vol. 584 115–119 (2020).
 26. Liu, L. *et al.* Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* vol. 584 450–456 (2020).
 27. Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* vol. 583 290–295 (2020).
 28. Rogers, T. F. *et al.* Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* vol. 369 956–963 (2020).
 29. Zost, S. J. *et al.* Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* vol. 584 443–449 (2020).
 30. Tortorici, M. A. *et al.* Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **370**, 950–957 (2020).
 31. Ledford, H. Antibody therapies could be a bridge to a coronavirus vaccine — but will the

- world benefit? *Nature* vol. 584 333–334 (2020).
32. Leyva-Grado, V. H., Tan, G. S., Leon, P. E., Yondola, M. & Palese, P. Direct Administration in the Respiratory Tract Improves Efficacy of Broadly Neutralizing Anti-Influenza Virus Monoclonal Antibodies. *Antimicrobial Agents and Chemotherapy* vol. 59 4162–4172 (2015).
 33. Detalle, L. *et al.* Generation and Characterization of ALX-0171, a Potent Novel Therapeutic Nanobody for the Treatment of Respiratory Syncytial Virus Infection. *Antimicrobial Agents and Chemotherapy* vol. 60 6–13 (2016).
 34. Cunningham, S. *et al.* Nebulised ALX-0171 for respiratory syncytial virus lower respiratory tract infection in hospitalised children: a double-blind, randomised, placebo-controlled, phase 2b trial. *The Lancet Respiratory Medicine* vol. 9 21–32 (2021).
 35. Vincke, C. *et al.* General Strategy to Humanize a Camelid Single-domain Antibody and Identification of a Universal Humanized Nanobody Scaffold. *Journal of Biological Chemistry* vol. 284 3273–3284 (2009).
 36. Wrapp, D. *et al.* Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* vol. 181 1004–1015.e15 (2020).
 37. Stadlbauer, D. *et al.* SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. *Curr. Protoc. Microbiol.* **57**, e100 (2020).
 38. I. Lui, X. X. Zhou, S. A. Lim, S. K. Elledge, P. Solomon, N. J. Rettko, B. S. Zha, L. L. Kirkemo, J. A. Gramespacher, J. Liu, F. Muecksch, J. C. C. Lorenzi, F. Schmidt, Y. Weisblum, D. F. Robbiani, M. C. Nussenzweig, T. Hatziioannou, P. D. Bieniasz, O. S. Rosenburg, K. K. Leung, J. A. Wells, Trimeric SARS-CoV-2 Spike interacts with dimeric ACE2 with limited intra-Spike avidity. bioRxiv 2020.2005.2021.109157 [Preprint]. 21 May

- 2020; .
39. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* vol. 26 795–800 (1993).
 40. McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* vol. 40 658–674 (2007).
 41. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography* vol. 66 213–221 (2010).
 42. G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O. S. Smart, C. Vornrhein, T. O. Womack, BUSTER version 1.10.0 (Global Phasing Ltd., Cambridge, UK, 2017).
 43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* vol. 60 2126–2132 (2004).
 44. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* vol. 152 36–51 (2005).
 45. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* vol. 14 331–332 (2017).
 46. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* vol. 14 290–296 (2017).
 47. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7, (2018).

48. D. Asarnow, E. Palovcak, Y. Cheng, UCSF pyem v0.5 (Zenodo, 2019).
49. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *eLife* vol. 7 (2018).
50. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* vol. 27 14–25 (2018).
51. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* vol. 581 215–220 (2020).
52. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β 2 adrenoceptor. *Nature* vol. 469 175–180 (2011).
53. Frenz, B., Walls, A. C., Egelman, E. H., Veisler, D. & DiMaio, F. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nature Methods* vol. 14 797–800 (2017).
54. Croll, T. I. *ISOLDE*: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D Structural Biology* vol. 74 519–530 (2018).
55. Wang, R. Y.-R. *et al.* Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *Elife* **5**, (2016).
56. Frenz, B. *et al.* Automatically Fixing Errors in Glycoprotein Structures with Rosetta. *Structure* vol. 27 134–139.e3 (2019).
57. Agirre, J. *et al.* Privateer: software for the conformational validation of carbohydrate structures. *Nature Structural & Molecular Biology* vol. 22 833–834 (2015).
58. Winn, M. D. *et al.* Overview of theCCP4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography* vol. 67 235–242 (2011).

59. Chen, V. B. *et al.* *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* vol. 66 12–21 (2010).
60. Barad, B. A. *et al.* EMRinger: side chain–directed model and map validation for 3D cryo-electron microscopy. *Nature Methods* vol. 12 943–946 (2015).
61. Staus, D. P. *et al.* Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* vol. 535 448–452 (2016).
62. Yu, F. *et al.* Fast Quantitative Analysis of timsTOF PASEF Data with MSFragger and IonQuant. *Molecular & Cellular Proteomics* vol. 19 1575–1585 (2020).
63. Choi, M. *et al.* MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* vol. 30 2524–2526 (2014).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Bryan Faust

F090B4BA5B91415...

Author Signature

7/19/2022

Date