

NORTHWESTERN UNIVERSITY

High-Throughput Protein-Protein Interaction Screening Using Cell-Free Protein Synthesis

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemical and Biological Engineering

By

Andrew Colin Hunt

EVANSTON, ILLINOIS

September 2022

© Copyright by Andrew Colin Hunt 2022

All Rights Reserved

Abstract

Protein-protein interactions are ubiquitous in living systems, and mediate important cellular processes from decision making to immunity against pathogens. Furthermore, protein-protein interactions are key to many protein therapeutics, pathogen diagnostics, and numerous synthetic biology applications. As a result, there has been significant effort to develop methods to express potential protein interaction partners and evaluate their protein-protein interactions in high-throughput. However, despite decades of development, many methods are still bottlenecked by labor intensive and poorly scalable steps involving cell culture. In this dissertation, I describe my efforts to build a high-throughput protein expression platform leveraging cell-free protein synthesis (CFPS) and acoustic liquid handling robotics and my efforts to apply this method to the study and engineering of protein-protein interactions.

Towards this goal, I first developed methods to express and screen computationally designed protein heterodimers. In collaboration with researchers in the Baker lab, we sought to use these heterodimers to design protein-protein interaction-based logic gates to mediate post-translational control of biological systems. Leveraging CFPS and a nanoluciferase-based complementation reporter, I evaluated the pairwise interactions of computationally designed heterodimers and used their interaction map to construct induced dimerization, AND, OR, and NOR logic gates. I also showed that the gates can rapidly integrate information, with the induced dimerization gate exhibiting a 7-fold induction within 5 minutes of activation. In collaboration with others, we also found that protein-protein interaction-based logic gates also operate in both yeast and human T cells, highlighting this promising strategy for post-translational control of biological behavior.

I next turned my efforts to develop high throughput methods for the discovery of antibodies. I showed that a crude extract based CFPS system can support the expression and assembly of functional antibody fragments from linear DNA templates in 2 μ L CFPS reactions in 384-well plates. Using the AlphaLISA high-throughput protein-protein interaction assay, I analyzed the interactions of antibodies and their antigens in crude CFPS reactions without purification. To highlight the capabilities of the platform, in less than 24 hours I expressed and screened 120 previously reported neutralizing antibodies targeting the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike glycoprotein. The developed method would have enabled the discovery of 10 out of the 13 most potent antibodies screened, is end-to-

end automatable, and has more than 10x the throughput and is more than 3.5x faster than state-of-the-art antibody discovery workflows.

I next led an effort to develop computationally designed multivalent minibinders that inhibit the key protein-protein interaction that enables SARS-CoV-2 to enter host cells. Leveraging similar methods to those developed for the antibody screening workflow, I expressed and screened hundreds of multivalent minibinders and identified candidates that bound with high affinity to the SARS-CoV-2 spike glycoprotein of the original Wuhan-Hu-1 variant and all other tested variants of concern. The top design, TRI2-2, exhibits an apparent dissociation rate slower than 10^{-7} s^{-1} , simultaneously engages all three receptor binding domains of the spike glycoprotein, and potently neutralizes all tested SARS-CoV-2 variants including Delta (B.1.617.2) and Omicron (B.1.1.529 or BA.1). Furthermore, TRI2-2 confers protection against SARS-CoV-2 when administered intranasally in mice, indicating that TRI2-2 is a promising potential therapeutic for the treatment of the coronavirus disease 2019 (COVID-19).

In a subsequent collaborative effort, I adapted the TRI2-2 multivalent minibinder for the detection of SARS-CoV-2 infection in a nanomechanical biosensor. The developed sensors enable detection of the SARS-CoV-2 spike glycoprotein antigen in less than 5 minutes with a limit of detection more than two orders of magnitude better than state-of-the-art lateral flow antigen tests and on par with that of state-of-the-art nucleic acid amplifications tests like the quantitative reverse transcription polymerase chain reaction (RT-qPCR).

Taken as a whole, the work in this dissertation provides advancements in the high-throughput expression proteins and the analysis of protein-protein interactions.

Acknowledgments

First and foremost, I owe thanks to my advisor, Mike Jewett. I have learned immensely from you as a scientist, teacher, mentor, and person. Your vision of what is possible with synthetic biology drew me in and hasn't let me go. Thank you for all the work you do to bring this vision into reality. Thank you for creating an environment where people can thrive and support one another. I am so grateful that you gave me the freedom and resources to pursue the science that I was most interested in, especially considering where I started and the time it took us to get there. I would also like to thank my thesis committee: Julius Lucks, Danielle Tullman-Ercek, and Neil Kelleher for their advice as well as Josh Leonard for being my mentor in the teaching apprenticeship program. I also owe thanks to the National Defense Science and Engineering Graduate Fellowship which funded much of my PhD work.

I was fortunate to have a highly collaborative PhD and owe many thanks to the numerous researchers I collaborated with. First to David Baker, who agreed to collaborate with an unknown graduate student on an unknown project and find something interesting to all parties involved. It changed the course of my PhD for the better and I appreciate your patience and guidance throughout. In essence, you became a second advisor and I benefitted greatly from our interactions. From the Baker lab my thanks to Zibo Chen, Scott Boyken, Ryan Kibler, Basile Wicky, and Ajasja Ljubetič for our collaboration on protein logic gates and heterodimer profiling. Also from the Baker lab, a thank you to Longxing Cao, Kejia Wu, and Lauren Carter for our collaboration on the design of multivalent minibinders that neutralize SARS-CoV-2. On the same project, I also owe thanks to Michael Diamond, Brett Case, David Veessler, Lexi Walls, Young-Jun Park, and the many other folks I had the pleasure of working with. I also owe thanks to Vinayak Draid and Dilip Agarwal for our collaboration on the nanomechanical sensor project.

A PhD is a journey, and I could not have asked for better companions to adventure with. Bastian Vögeli, Antje Krüger, Jazzy Hershewe, Weston Kightlinger, Mike Hammerling, Dani Yoesep, Ashty Karim, Adam Silverman, Kosuke Seki, Grant Landwher, Jon Bogart, Steve Flemming, Blake Rasor, Grant Rybnicky, Eric Hammarlund, Do Soon Kim, and all the other members of the Jewett lab that I've crossed paths with, thank you for your continual support and friendship, and for making the journey as good as the destination. It has been a privilege to be part of the same communal superorganism as you all over the past six years. To Jazzy Hershewe, thank you for being my impromptu lab mentor, friend, and general co-

conspirator as we embarked on our PhDs together. To Bastian, thank you for being such a wonderful friend and mentor. It still amazes me that you gave months of your time during the pandemic to someone else's project, thank you for your help and guidance. Also thank you for DMing our D&D campaign and just generally always keeping the mood bright. To Antje, thank you for being an excellent benchmate, mentor, tennis partner, and friend. Like Bastian, I also owe you many thanks for your help with my projects during the pandemic. To Weston, thank you for your friendship that has now crossed state and national borders, and for your continual mentorship and perspective. To Mike Hammerling, thank you for being my desk mate, friend, mentor, and go to skeptic. Much of my perspective on the forces at play within science is a result of our conversations. To Ashty Karim, thank you for all that you've done for the lab and for me, I have learned much from your mentorship and friendship. To Kosuke, thank you for being my pandemic office mate, friend, and go-to person to complain about bad data and discuss new ideas. To Blake, thank you for your friendship and for being my grappling partner despite our weight differential. Whiskey nights and fight nights were always a highlight of the week. To Adam, thank you for your friendship and insight, I always know where to go when I need someone to tell me their honest opinion about my ideas. To the mongooses, Jon, Grant, Bastian, Kosuke, Steve, Eric, Kelsi, generally goofing around with you all made the end of my PhD some of the most enjoyable times I've had in Chicago and helped me to finish my PhD after the grueling life of the early pandemic.

I would not be where I am today without my family. To all my family, thank you for your love and support, it means the world to me. To my parents, you've enabled all my endeavors and taught me so much, I cannot fully express the depth of my gratitude in words. To my mother Lynne Hunt, thank you for being the most caring and supportive person I know. To my father Warren Hunt, thank you for teaching me how to question and giving me your love of building. To my grandfathers, Joseph Janni and Neil Hunt, you instilled in me curiosity about the world and fostered that curiosity as it grew, thank you.

I also want to acknowledge the friends supporting me along the way. To Alex Prybutok and Peter Tran, thanks for your friendship and support as we weathered the storm of not only our PhDs, but also a global pandemic. Alex, your friendship has been an anchor, thank you for always lending an ear. To those friends outside the Chicago area that kept me sane, Tobias Burger, Will Sullivan, Matteo D'Andrea, and Franciszka Stopa, thank you for your support over the years.

There are also many who were my teachers and mentors prior to graduate school that helped me reach the end of a PhD. Michaelann Tartis, Menake Piyasena, Jason Harper, Stefan Roberts, and Andrea Ambrosini, thank you for training me, supporting me, and believing in me. Michaelann, thank you for investing in a freshman student interested in research, I was fortunate for your lab to have been my introduction to research. Stefan, your work inspired my interest in proteins, and set me down my current path. Thank you to the teachers that fostered my interest in STEM, Agustin Kintanar, Barbara Gilbert, and Alan Vraspir. To my many coaches over the years, Laney Kolek, Monty Blodgett, Dean Jacoby, and Gretchen Hanks, thank you for teaching me grit and how to overcome failure.

“We have to continually be jumping off cliffs and developing our wings on the way down.”

— Kurt Vonnegut

Table of Contents

| | |
|---|-----------|
| CHAPTER 1 - INTRODUCTION | 34 |
| 1.1 MOTIVATION..... | 34 |
| 1.2 DISSERTATION ROADMAP | 36 |
| 1.3 BACKGROUND | 37 |
| 1.3.1 Cell free protein synthesis..... | 37 |
| 1.3.2 Templates for CFPS..... | 38 |
| 1.3.3 Quantifying the results of CFPS..... | 41 |
| 1.3.4 Expression of complex proteins using CFPS | 42 |
| 1.3.4.1 Folding chaperones | 43 |
| 1.3.4.2 Disulfide bond formation..... | 45 |
| 1.3.5 Enabling technologies for CFPS applications | 45 |
| 1.3.6 CFPS for protein-protein interaction screening and engineering | 48 |
| 1.3.6.1 Individual Screening | 48 |
| 1.3.6.2 Microarrays and Microfluidic Systems | 50 |
| 1.3.6.3 In Vitro Display Technologies | 51 |
| CHAPTER 2 - DE NOVO DESIGN OF PROTEIN LOGIC GATES..... | 55 |
| 2.1 ABSTRACT | 55 |
| 2.2 INTRODUCTION..... | 56 |
| 2.3 RESULTS | 57 |
| 2.3.1 Design and testing of an induced dimerization system | 57 |
| 2.3.2 Design and testing of multi-input logic gates..... | 60 |
| 2.3.3 Evaluation of CIPHR logic gates in different environments..... | 65 |
| 2.4 DISCUSSION..... | 67 |
| 2.5 MATERIALS AND METHODS | 69 |
| 2.5.1 Buffer and media recipe for protein expression and yeast-two-hybrid experiments | 69 |
| 2.5.2 Construction of synthetic genes..... | 69 |

| | |
|--|----|
| | 10 |
| 2.5.3 Protein expression | 70 |
| 2.5.4 Affinity purification | 71 |
| 2.5.5 Size-exclusion chromatography (SEC) | 71 |
| 2.5.6 Circular dichroism (CD) measurements | 71 |
| 2.5.7 CD data analysis and model fitting..... | 71 |
| 2.5.8 Small Angle X-ray Scattering (SAXS) | 73 |
| 2.5.9 Yeast two-hybrid assay for logic gates..... | 74 |
| 2.5.10 Native MS of individual proteins..... | 74 |
| 2.5.11 Native MS titration assay and data analysis | 74 |
| 2.5.12 Cell-Free extract preparation | 75 |
| 2.5.13 Cell-Free Protein Synthesis reactions..... | 76 |
| 2.5.14 Linear templates for Cell-Free Protein Synthesis..... | 77 |
| 2.5.15 NanoBiT reactions..... | 77 |
| 2.5.16 T cell husbandry..... | 78 |
| 2.5.17 Logic-gated control of TALE-KRAB repressors..... | 78 |
| 2.5.18 In vitro transcription of repressors..... | 79 |
| 2.5.19 T cell electroporation..... | 79 |
| 2.5.20 T Cell flow cytometry..... | 79 |
| 2.5.21 Progesterone-responsive yeast strain construction and growth media..... | 80 |
| 2.5.22 Yeast cell culture | 80 |
| 2.5.23 Progesterone induction | 80 |
| 2.5.24 Yeast cell flow cytometry..... | 81 |
| 2.5.25 Data Deposition..... | 81 |

| | |
|---|-----------|
| CHAPTER 3 - A HIGH-THROUGHPUT, AUTOMATABLE, CELL-FREE EXPRESSION AND SCREENING PLATFORM FOR ANTIBODY DISCOVERY | 82 |
| 3.1 ABSTRACT | 82 |
| 3.2 INTRODUCTION..... | 83 |
| 3.3 RESULTS | 84 |

| | |
|---|------------|
| | 11 |
| 3.3.3 Development of a cell-free DNA assembly and amplification workflow | 84 |
| 3.3.3 Development of a crude extract based CFPS system for antibody expression | 85 |
| 3.3.4 Integration of the AlphaLISA PPI screen into the workflow..... | 86 |
| 3.3.5 Evaluation of a large set of previously published antibodies..... | 87 |
| 3.4 DISCUSSION..... | 89 |
| 3.5 MATERIALS AND METHODS | 90 |
| 3.5.1 Antibody sdFab sequence design | 90 |
| 3.5.2 DNA assembly and linear expression template (LET) generation | 91 |
| 3.5.3 Cell extract preparation for cell-free protein synthesis..... | 95 |
| 3.5.4 DsbC and FkpA expression and purification | 95 |
| 3.5.5 Cell-free protein synthesis..... | 96 |
| 3.5.6 AlphaLISA reactions..... | 97 |
| 3.5.7 Data analysis..... | 99 |
| | |
| CHAPTER 4 - MULTIVALENT DESIGNED PROTEINS PROTECT AGAINST SARS-COV-2 VARIANTS | |
| OF CONCERN | 101 |
| | |
| 4.1 ABSTRACT | 101 |
| 4.2 INTRODUCTION..... | 102 |
| 4.3 RESULTS | 103 |
| 4.3.1 Impact of RBD mutations on minibinder binding..... | 103 |
| 4.3.2 Multivalent minibinder design and experimental optimization | 103 |
| 4.3.3 Structural studies of minibinders in complex with SARS-CoV-2 S..... | 106 |
| 4.3.4 Multivalent minibinder enables rapid detection of SARS-CoV-2 S..... | 108 |
| 4.3.5 Multivalent minibinders bind tightly to SARS-CoV-2 variants..... | 109 |
| 4.3.6 Multivalent minibinders potently neutralize circulating SARS-CoV-2 variants | 112 |
| 4.3.7 Multivalent minibinders resist viral escape..... | 113 |
| 4.3.8 Multivalent minibinder confers protection in human ACE2-expressing transgenic mice..... | 116 |
| 4.4 DISCUSSION..... | 116 |
| 4.5 MATERIALS AND METHODS | 118 |

| | |
|---|-----|
| | 12 |
| 4.5.1 Study design | 118 |
| 4.5.2 Cell lines and cell culture | 118 |
| 4.5.3 LCB1 (MON1) and LCB3 (MON3) optimization | 119 |
| 4.5.4 Multivalent fusion constructs design | 120 |
| 4.5.5 Self-assembling homotrimer design: scaffold selection and backbone generation..... | 120 |
| 4.5.6 Self-assembling homotrimer design: junction modification and flexibility insertion | 121 |
| 4.5.7 Deep mutational scanning profiles of minibinder escape..... | 121 |
| 4.5.8 Cell-free DNA assembly and CFPS template preparation | 122 |
| 4.5.9 CFPS cell extract preparation | 123 |
| 4.5.10 CFPS reactions and purification..... | 123 |
| 4.5.11 Expression and purification of competitor proteins for AlphaLISA experiments..... | 124 |
| 4.5.12 Expression and purification of SARS-CoV-2 S proteins..... | 125 |
| 4.5.13 Competition-based off-rate screening via AlphaLISA | 126 |
| 4.5.14 Negative stain electron microscopy | 127 |
| 4.5.15 Cryo-electron microscopy | 127 |
| 4.5.16 BRET sensor for SARS-CoV-2 S detection | 129 |
| 4.5.17 Minibinder/RBD interaction kinetics via bio-layer interferometry (BLI) | 131 |
| 4.5.18 Multivalent Minibinder/S6P variant interaction kinetics via surface plasmon resonance (SPR) | 131 |
| 4.5.19 Competition ELISA of minibinders and SARS-CoV-2 S6P variants for immobilized hACE2-Fc | 132 |
| 4.5.20 Pseudovirus production..... | 132 |
| 4.5.21 Pseudovirus neutralization | 133 |
| 4.5.21 SARS-CoV-2 neutralization..... | 133 |
| 4.5.22 Kidney organoid differentiation and infection with SARS-CoV-2..... | 134 |
| 4.5.23 Kidney organoid plaque assay..... | 134 |
| 4.5.24 Kidney organoid gene expression analysis..... | 135 |
| 4.5.25 Selection of escape mutants in SARS-CoV-2 S using VSV-SARS-CoV-2 chimera..... | 135 |

| | |
|---|------------|
| | 13 |
| 4.5.26 Protein production for animal studies..... | 136 |
| 4.5.27 Mouse studies..... | 136 |
| 4.5.28 Measurement of viral burden in mouse studies..... | 137 |
| 4.5.29 Pharmacokinetics of TRI2-2 in mouse studies..... | 137 |
| 4.5.30 Statistical analysis..... | 138 |
| | |
| CHAPTER 5 - RAPID AND SENSITIVE DETECTION OF ANTIGEN FROM SARS-COV-2 VARIANTS OF CONCERN BY A MULTIVALENT MINIBINDER-FUNCTIONALIZED NANOMECHANICAL SENSOR | 139 |
| 5.1 ABSTRACT..... | 139 |
| 5.2 INTRODUCTION..... | 140 |
| 5.3 RESULTS..... | 142 |
| 5.3.1 Comparison of antibody, AHB2, and TRI2-2 sensing of SARS-CoV-2 S trimer..... | 142 |
| 5.3.2 Detection of S trimer from SARS-CoV-2 variants of concern..... | 142 |
| 5.3.3 Sensing of SARS-CoV-2 in clinical samples and evaluation of cross-reactivity..... | 144 |
| 5.4 DISCUSSION..... | 145 |
| 5.5 MATERIALS AND METHODS..... | 146 |
| 5.5.1 Chemical reagents and materials..... | 146 |
| 5.5.2 Recombinant protein expression and purification..... | 146 |
| 5.5.3 Cantilever surface biofunctionalization..... | 148 |
| 5.5.4 Optical detection method..... | 148 |
| 5.5.5 Patient sample collection and Ct value calculation..... | 148 |
| 5.5.6 Ct Value Calculation and Variant Identification..... | 149 |
| | |
| CHAPTER 6 –FUTURE WORK AND PERSPECTIVES..... | 151 |
| 6.1 PROPOSED FUTURE WORK..... | 151 |
| 6.2 PERSPECTIVE..... | 151 |
| | |
| CHAPTER 7 - APPENDIX CONTAINING SUPPLEMENTARY INFORMATION FOR CHAPTER 2..... | 155 |
| 7.1 THERMODYNAMIC MODELING OF COOPERATIVITY..... | 155 |

| | |
|--|------------|
| | 14 |
| 7.2 SUPPLEMENTARY FIGURES | 159 |
| CHAPTER 8 - APPENDIX CONTAINING SUPPLEMENTARY INFORMATION FOR CHAPTER 3 | 187 |
| CHAPTER 9 - APPENDIX CONTAINING SUPPLEMENTARY INFORMATION FOR CHAPTER 4 | 200 |
| CHAPTER 10 - APPENDIX CONTAINING SUPPLEMENTARY INFORMATION FOR CHAPTER 5 | 232 |
| CHAPTER 11 REFERENCES | 236 |

List of Figures

| | |
|--|----|
| Figure 1.1. Templates used to drive transcription and or translation in CFPS systems. IVT: in vitro transcription; DNA: deoxyribonucleic acid; mRNA: messenger ribonucleic acid; PCR: polymerase chain reaction; RCA: rolling circle amplification. | 38 |
| Figure 1.2. Methods for evaluating PPIs in CFPS that don't require purification of the interaction partners. | 49 |
| Figure 1.3. Schemes depicting ribosome and mRNA display, the two most widely used display technologies leveraging CFPS. | 52 |
| Figure 2.1. Cooperativity of CIPHR logic gates. (A) Left: Backbone structure of A:A' heterodimer building block, with its hydrogen bond network in inset. Bottom: Shorthand representations used throughout figures. (B) Thermodynamic cycle describing the induced dimerization system. (C) Simulation of the induced dimerization system under thermodynamic equilibrium. A and B' monomers were held constant at 10 μ M each while titrating in various initial amounts of the A'-B dimerizer proteins. If binding is not cooperative (small c), the final amount of trimeric complexes decreases when the dimerizer protein is in excess. (D) Equilibrium denaturation experiments monitored by CD for designs with 6- and 12- amino acid (AA) linkers. Circles represent experimental data, and lines are fits to the 3-state unimolecular unfolding model. (E) Experimental SAXS profile of 1'-2' with a 6-residue linker (in black), fitted to the calculated profile of 1:1' heterodimer. (F) An induced dimerization system using a 6-residue linker. (G) Native MS titration of 2 against 1'-2' in the presence (red) or absence (blue) of 1. (H) Native MS titration of 1'-2' against 1 and 2. Dimer 1 and 2 refer to partial dimeric complexes consisting of the dimerizer binding to either of the monomers. For comparison, the thermodynamic model result with $c = 991,000$ is shown in cyan. (I) Schematic of testing of the induced dimerization system in yeast, with <i>in vivo</i> results in (J). Pg, progesterone. (K) A two-input AND gate schematic, with native MS titration results in (L). Trimer 1 and 2 refer to partial trimeric complexes of the two dimerizer proteins binding to either one of the monomers. (M) A three-input AND gate, with native MS titration results in (N). Tetramer 1 and 2 refer to partial tetrameric complexes of the three dimerizer proteins binding to either one of the monomers. All error bars are reported as standard deviations of n=3 independent replicates. | 58 |

Figure 2.2. CIPHR two input logic gates. (A) CIPHR gates are built from DHDs (top) with monomers or covalently connected monomers as inputs (left); some gates utilize only the designed cognate interactions (left side of middle panel), while others take advantage of observed binding affinity hierarchies (right side of middle panel). (B-C) Two-input AND (B) and OR (C) CIPHR logic gates based on orthogonal DHD interactions. (D-G) NOT (D), NOR (E), XNOR (F), and NAND (G) CIPHR logic gates made from multispecific and competitive protein binding. For each gate, black dots represent individual Y2H growth measurement corrected over background growth, with their average values shown in green bars. * indicates no yeast growth over background. 0s and 1s in the middle and right blocks represent different input states and expected outputs, respectively.62

Figure 2.3. Three-input CIPHR logic gates. (A) Schematic of a three-input AND gate. (B) Native MS results indicate proper activation of the 3-input AND gate only in the presence of all three inputs. (C) Schematic of a three-input OR gate. (D) Y2H results confirmed activation of the 3-input OR gate with either of the inputs. (E) Schematic of a DNF gate. (F) Y2H results confirmed proper activation of the gate. For each gate, black dots represent individual measurements with their average values shown in green bars. For Y2H-based measurements (D,F), the growth measurements are corrected over background growth.64

Figure 2.4. Transferability of CIPHR logic gates. (A) Four pairs of DHDs were modularly combined to construct CIPHR logic gates that can be used to control different functions: (B-E) catalytic activity of split luciferase, and (F-G) gene expression in primary human T cells. (B) Induced dimerization system, (C) AND gate, and (D) NOR gate coupled to NanoBiT split luciferase system, tested by *in vitro* translation and monitoring luminescence. (E) *In vitro* titration of the two inputs of the NOR gate in D while keeping 1'-smBiT and 2'-lgBiT fixed at 5 nM. (F) NOT gate and (G) OR gate using a split TALE-KRAB repression system to control expression of TIM3 proteins in primary human T cells, tested by flow cytometry.66

Figure 3.1. A high-throughput, cell-free antibody screening workflow. (A) Schematic of the steps involved in the cell-free antibody screening workflow. (B) Diagram of the AlphaLISA screen for neutralizing antibodies via competition with ACE2 for the SARS-CoV-2 RBD. (C) Evaluation of commercial neutralizing antibodies (nAbs) in the AlphaLISA ACE2 competition screen (n=3 independent replicates \pm SEM). (D) Comparison of the reported and measured potencies of commercial neutralizing antibodies.85

Figure 3.2. Performance of the cell-free antibody screening workflow evaluated on SARS-CoV-2 neutralizing antibodies. (A-F) AlphaLISA data are presented as the mean of 3 independent replicates. A dashed line indicates three standard deviations away from the background signal. (A-B) Heatmap of the binding of previously published antibodies measured using AlphaLISA to detect S trimer binding (\log_{10} scaled), RBD binding (\log_{10} scaled), and ACE2 competition (linearly scaled). AlphaLISA data are presented as the mean of 3 independent replicates. The lowest reported neutralization IC_{50} value is also plotted for comparison (\log_{10} scaled) and an X indicates no relevant data available (Table 8.2). (A) Heatmap of the binding of 36 diverse antibodies. (B) Heatmap of the binding of all 84 antibodies in the Brouwer *et al.* data set. (C-D) Parity plots comparing the AlphaLISA the 84 antibodies in the Brouwer *et al.* data set vs the published ELISA data. A dashed line indicates three standard deviations away from the background. (C) S trimer binding. (D) RBD binding. (E) Comparison of the S trimer and RBD AlphaLISA binding data. (F) Parity plot comparing the AlphaLISA ACE2 competition data for the 84 antibodies in the Brouwer *et al.* data set vs the published pseudovirus neutralization data. Antibodies that were reported to compete with ACE2 by Brouwer *et al.* are plotted in red. 88

Figure 4.1. Multivalent minibinders exhibit very slow dissociation rates upon binding to the prefusion SARS-CoV-2-S glycoprotein. Dissociation of the minibinder construct was monitored via competition with 100-fold molar excess of untagged MON1 using AlphaLISA (Mean \pm SEM, n = 3 technical replicates from a single experiment). (A) Dissociation of the minibinder construct complexed with the receptor-binding domain (RBD). (B) Dissociation of the minibinder construct complexed with the S trimer (S6P). 105

Figure 4.2. CryoEM structures of multivalent minibinders in complex with the SARS-CoV-2 S6P glycoprotein. (A) CryoEM map of TRI2-2 in complex with the S6P in two orthogonal orientations. (B) Zoomed-in view of the TRI2-2 and RBD complex obtained using focused 3D classification and local refinement. The RBD and MON2 built in the 2.9 Å resolution cryoEM map are shown in yellow and blue, respectively. (C) CryoEM map of FUS31-G10 bound to S6P. (D) CryoEM map of FUS231-P24 bound to S6P. (E) Negative-stain EM map of FUS231-G10 in complex with S6P. S and minibinder models were docked in the whole map by rigid body fitting for visualization. In all panels, the EM density is shown as a transparent gray surface, S protomers (PDB 7JZL) are rendered in yellow, cyan, and pink and minibinders (PDB 7JZU, 7JZM, and MON2 structure was determined in this study) are shown in orange. 108

Figure 4.3. FUS231-P12 enables detection of SARS-CoV-2 S trimer via BRET. (A) Schematic representation of the BRET sensor, teluc-FUS231-P12-mCyRFP3, to detect S trimer. (B) Luminescence emission spectra and image of the BRET sensor (100 pM) in the presence (orange trace, 100 pM) and absence (blue trace) of S2P. Emission color change was observed using a mobile phone camera (inset top right). (C) Titration of S2P with 100 pM sensor protein (Mean \pm SEM, n = 3 technical replicates from a single experiment). 109

Figure 4.4. Multivalency enhances both the breadth and potency of neutralization against SARS-CoV-2 variants by minibinders. (A) Dissociation of minibinder constructs from S6P variants after 24 hours was measured via competition with untagged TRI2-1 using AlphaLISA (mean, n = 3 technical replicates from a single experiment). Cells containing an X indicate insufficient signal in the absence of a competitor to quantify the fraction of protein bound. (B) Competition of minibinder constructs with ACE2 for binding S6P measured via ELISA (mean, n = 2 technical replicates representative of two independent experiments). (C) Neutralization of SARS-CoV-2 pseudovirus variants by minibinder constructs (mean, n = 2 technical replicates representative of two independent experiments). (D) Neutralization of authentic SARS-CoV-2 by minibinder constructs (mean \pm SEM, n = 4 technical replicates from two independent experiments for all but B.1.1.529 where n = 8 technical replicates from four independent experiments). (E) Summary of neutralization potencies of multivalent minibinder constructs against SARS-CoV-2 pseudovirus variants. N/A indicates an IC₅₀ value above the tested concentration range and an IC₅₀ greater than 50,000 pM. (F) Summary of neutralization potencies of multivalent minibinder constructs against authentic SARS-CoV-2 variants. N/A indicates an IC₅₀ value above the tested concentration range and an IC₅₀ greater than 30,000 pM. NT indicates not tested. (G) Quantification of replicating authentic B.1.351 virus in the presence of minibinder constructs (0.3 μ M) in human kidney organoids (mean \pm SEM, n = 4 biological replicates with 2-3 technical replicates per experiment, Kruskal-Wallis test with Dunn's post-hoc analysis: ** P < 0.01, *** P < 0.001). Dashed line indicates lower limit of detection of plaque assay. (H) Relative gene expression of SARS-CoV-2 envelope protein (SARS-CoV2-E) in kidney organoids post viral infection with and without multivalent minibinders (0.3 μ M) (mean \pm SEM, n = 4 biological replicas with 2-3 technical replicates per experiment, Kruskal-Wallis test with Dunn's post-hoc analysis: * P < 0.05, *** P < 0.001). 111

Figure 4.5. Top multivalent minibinder candidates are escape resistant. (A) Plaque assays were performed to isolate VSV-SARS-CoV-2 S chimera virus escape mutants against a control neutralizing antibody (2B04) and the FUS231-P12 and TRI2-2 multivalent minibinders. For each inhibitor tested, Vero cells were incubated with VSV-SARS-CoV-2 S chimera virus for one hour, followed by addition of the inhibitor protein at a fully neutralizing concentration, and further incubation to allow for replication and spread of resistant viruses. 36 independent selections were carried out for each minibinder compound in a single experiment; representative examples are shown in the images. Small plaques are indicative of inhibited viral spreading and large plaques, highlighted by black arrows, are indicative of viral escape mutants spreading. (B) Summary of the results of the viral escape screen. 114

Figure 4.6. Top multivalent minibinder candidates protect mice from SARS-CoV-2 challenge. (A) K18-hACE2-transgenic mice ($n = 6$ from two independent experiments) were dosed with $50 \mu\text{g}$ of the indicated minibinder by i.n. administration ($50 \mu\text{l}$ total) 24 h prior (D-1) to infection with 10^3 focus forming units (FFU) of SARS-CoV-2 variants B.1.1.7, Wash-B.1.351, or Wash-P.1 i.n. on Day 0. (B) Daily weight-change following inoculation (mean \pm SEM; two-way ANOVA with Sidak's post-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). (C) At 6 days post infection (dpi) animals ($n = 6$ from two independent experiments) were sacrificed and analyzed for SARS-CoV-2 viral RNA by RT-qPCR in the lung, heart, spleen, brain, or nasal wash (line at median; Kruskal-Wallis test with Dunn's post-hoc analysis: ns, not significant, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). (D) K18-hACE2-transgenic mice ($n = 6$ from two independent experiments) were dosed with $50 \mu\text{g}$ of the indicated minibinder by i.n. administration ($50 \mu\text{l}$ total) 24 h after (D+1) infection with 10^3 FFU of the SARS-CoV-2 Wash-B.1.351 or B.1.617.2 variant on Day 0. (E) Daily weight-change following inoculation (mean \pm SEM; two-way ANOVA with Sidak's post-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). (F) At 6 dpi (B.1.351) or 7 dpi (B.1.617.2), animals ($n = 6$ from two independent experiments) were sacrificed and analyzed for SARS-CoV-2 viral RNA by RT-qPCR in the lung, heart, spleen, brain, or nasal wash (line at median: two-tailed Mann-Whitney test: ns, not significant, * $P < 0.05$, ** $P < 0.01$). 115

Figure 5.1. Cantilever-based sensing of SARS-CoV-2 spike protein (S6P). (A) Cantilever sensing mechanism; deflection is observed upon binding of the SARS-CoV-2 spike trimer to the captured binder immobilized on the cantilever. (b-d) Deflection of cantilever sensors over time with anti-S antibody (B),

AHB2 (C), and TRI2-2 (D) immobilized on the cantilever (mean \pm SEM, n = 3). (E) Comparison of deflection between antibody, AHB2, and TRI2-2 cantilevers after 15 minutes of equilibration (mean \pm SEM, n = 3). For all plots, the dashed horizontal line indicates the deflection LOD cut-off (average of the combined negative control measurements + 3 standard deviations) and an absence of error bars indicates error within the marker. 142

Figure 5.2. Detection of purified S trimer representing current and historical SARS-CoV-2 variants of concern. (A-E) Deflection observed over time in response to purified S6P variants corresponding to the Alpha (A), Beta (B), Gamma (C), Delta (D), and Omicron (E) variants (mean \pm SEM, n = 3). (F) Comparison of cantilever deflection after 15 minutes of equilibration (mean \pm SEM, n = 3). For all plots, the dashed horizontal line indicates the deflection LOD cut-off (average of the combined negative control measurements + 3 standard deviations) and an absence of error bars indicates error within the marker. 143

Figure 5.3. Detection of SARS-CoV-2 in nasopharyngeal swabs from infected individuals. All samples were evaluated in triplicate using three distinct cantilevers. (A-C) Deflection of cantilevers observed over time in response to samples collected early in 2021 (A), from samples confirmed to contain the Alpha variant (B), and from samples confirmed to contain the Omicron variant (C). (D) Deflection of cantilevers measured against patient nasopharyngeal swabs RT-qPCR negative for SARS-CoV-2 (mean \pm SEM, n = 3). (E) Comparison of the RT-qPCR Ct value and cantilever deflection for the tested patient samples (mean \pm SEM, n = 3). (F) Deflection of cantilevers measured against recombinant purified MERS-CoV spike protein (1,000 ng/ml) and against a patient nasopharyngeal swab RT-qPCR positive for Influenza A (Ct = 19.33) (mean \pm SEM, n = 3). For all plots, the dashed horizontal line indicates the deflection LOD cut-off (average of the combined negative control measurements + 3 standard deviations) and an absence of error bars indicates error within the marker. 145

Figure 7.1. Additional information on the induced dimerization system. (A) Molecular implementation of the cooperative induced dimerization system, binding only occurs when all three components are present. (B) Size exclusion chromatography profiles of 1'-2' variants with 0, 2, 6, 12, and 24 amino acids as the flexible linker connecting 1' and 2'. 159

Figure 7.2. GdnHCl equilibrium denaturation experiments monitored by CD at 222 nm for 1'-2' designs with 0-, 2-, 6-, 12- and 24- residue linkers. For each plot, the top subplot shows the experimental data (circles) fitted to the 3-state unimolecular unfolding model (black line). The middle subplots show residuals to the fit, and the bottom subplots indicate the fraction of each species as a function of denaturant concentration. An approximation of helical content is also reported (second ordinate, based on the MRE value at 222 nm using the model of Muñoz & Serrano⁴⁹⁰). All fitted parameters are reported in Table 7.3. ΔG and m -values are also reported in the title of each plot. ΔG represent the free energy change extrapolated to buffer. The corresponding m -values represent the denaturant sensitivity of each transition, and relate to the $\Delta SASA$ associated with each event. 160

Figure 7.3. Native MS signal intensity of the dimerizer 1'-2' (6-residue linker), 1, and 2 against protein concentrations (see also Table 7.8). Equations for linear regression fits are displayed. All error bars are reported as standard deviations of $n=3$ independent replicates. 162

Figure 7.4. Induction of RFP (mScarlet) expression by the "Induced Dimerization" system in Fig. 2.1J, where the dimerizer species is measured by a separate, proxy "pZ3-YFP" (Venus) reporter. Dotted lines represent constitutive expression levels of YFP and RFP under the RPL18b promoter used to express the DBD and AD monomer species in Fig. 2.1J. Maximal induction of the dimerizer/YFP (128 nM Progesterone) is about 5x the value of pRPL18b-YFP, indicating that the cooperative binding functions across a wide range of concentrations. All error bars are reported as standard deviations of $n=3$ independent replicates. 163

Figure 7.5. Binding affinity gradient from individual Y2H experiments. (A) The 8:8' heterodimer binds more tightly than the homodimers of its monomers. (B) Binding affinity gradient among the monomers of 1:1', 9:9', and 10:10' pairs. Cell growth was measured on different days across the rows but on the same day within each row, and not corrected against background growth. 164

Figure 7.6. NanoBiT characterization of CIPHR logic gate components. (A) Specificity matrix for designs tested in the cell-free expression system. (B) Yeast two-hybrid specificity matrix for the same designs. (C) Kinetic measurements of induced dimerization gate at varying concentrations of the dimerizer input. (D) Kinetics of induced dimerization (with or without 335 nM dimerizer protein) in the first 30 minutes. (E) Kinetic measurements of NOR gate activation and subsequent deactivation via competitive binding upon addition of input (arrow). All error bars are reported as standard deviations of $n=3$ independent replicates. 165

Figure 7.7. T cell KRAB-TALE characterization of CIPHR logic gate components. (A) Specificity matrix for designs tested in T cells. (B) Yeast two-hybrid specificity matrix for the same designs. (C) Expression level of TIM3 in T cells not treated with CIPHR logic gates. 166

Figure 8.1. The cell-free DNA assembly and amplification workflow. (A) Schematic of the cell-free DNA assembly and amplification protocol for generating sfGFP linear expression template for CFPS. (B) Agarose gel of amplified LET PCR products of Gibson assembly reactions. Backbone only and insert only conditions included as negative controls. Labeled band indicates assembly and amplification of the correct length PCR product. (C) sfGFP yields in Origami™ B(DE3) CFPS from cell-free assembled linear expression templates and purified plasmid (n = 3 independent replicates ± SEM). (D) Schematic of the cell-free DNA assembly and amplification protocol for generating sdFab linear expression template for CFPS. (E) Agarose gel of amplified sdFab heavy chain (HC) LET PCR products. Labeled bands indicate assembly and amplification of the correct length PCR product. (F) Agarose gel of amplified sdFab light chain (LC) LET PCR products. Labeled bands indicate assembly and amplification of the correct length PCR product. 188

Figure 8.2. Development of an Origami™ B(DE3) CFPS system for the expression of antibodies and antibody fragments. (A-C) SDS PAGE of antibodies and antibody fragments manufactured in CFPS. Samples were fluorescently labeled with the FluoroTect™ reagent during protein synthesis. (A) Expression and assembly of full-length Trastuzumab (αHER2). The antigen-binding fragment (Fab) and fragment, crystallizable (FC) domains are labeled. (B) Expression and assembly of a panel of 13 Fabs. (C) Expression and assembly of a panel of 13 sdFabs. The leucine zipper heterodimer (LZA and LZB) assisting with Fab assembly are labeled. 189

Figure 8.3. AlphaLISA for profiling antibody protein-protein interactions in CFPS. (A-C) Evaluation of the effect of CFPS reagents on AlphaLISA. Concentrations are plotted as v/v fraction of the final concentration of the reagent in a CFPS reaction. Reagents were diluted in water at the concentration they normally reside at in CFPS. Reagents were tested in mixtures that were used to assemble CFPS reactions. The salt solution contains 8 mM magnesium glutamate, 10 mM ammonium glutamate, and 130 mM potassium glutamate. Master Mix contains 1.2 mM ATP, 0.85 mM GTP, 0.85 mM UTP, 0.85 mM CTP, 0.03 mg/mL folic acid, and 0.17 *E. coli* tRNA. Reagent Mix contains 0.4 mM NAD, 0.27 mM CoA, 4 mM oxalic acid, 1 mM putrescine, 1.5 mM spermidine, and 57 mM HEPES. Amino Acid Mix contains 2 mM of all 20 amino acids.

PEP is 30 μ M phosphoenolpyruvate. LET is 0.066 v/v fraction unpurified PCR mix containing the LET for sfGFP. (A) Evaluation of the effect of CFPS reagents on AlphaLISA detection chemistry using the TrueHits kit. Biotin and Streptavidin labeled beads associate directly with one another and serve as a control for reagents impacting the AlphaLISA measurement chemistry. (B) Evaluation of the effect of CFPS reagents on AlphaLISA detection of the SARS-CoV-2 RBD and ACE2 interaction measured by the Protein A donor bead and Ni Chelate acceptor bead. The Salt Solution and Amino Acids Mix inhibit immobilization of his-tagged proteins on the NiChelate bead. (C) Evaluation of the effect of CFPS reagents on AlphaLISA detection of the SARS-CoV-2 RBD and ACE2 interaction measured by the Protein A donor bead and anti-6xhis acceptor bead. The Salt Solution and Amino Acids Mix inhibit immobilization of his-tagged proteins on the NiChelate bead. (D) Schematic of AlphaLISA setup for measuring sdFab assembly. (E) Assembly AlphaLISA measurement of a panel of sdFabs. AlphaLISA signal is indicative of sdFab assembly, though the signal is subject to the hook effect⁵¹ resulting in a lower signal at higher concentrations. 190

Figure 8.4. Analysis of variability in AlphaLISA replicates. (A-D) From top to bottom: Histogram of raw AlphaLISA values (mean of n = 3 independent replicates) to visualize the spread of the data. A histogram of coefficient of variation (standard deviation divided by the mean) to visualize the typical error within a sample with the mean coefficient of variation displayed on the plot. Parity plots of the three replicates were fit to the line $y=x$ to visualize the consistency of replicates with the corresponding R^2 value is displayed on the chart. Only values found to be significantly different from the background are plotted ($p < 0.05$, two-sided t-test adjusted for multiple comparisons using FDR with a family-wise error rate of 5%) (A) sdFab assembly AlphaLISA. (B) SARS-CoV-2 S trimer binding AlphaLISA. (C) SARS-CoV-2 RBD binding AlphaLISA. (D) sdFab competition with ACE2 for the SARS-CoV-2 RBD AlphaLISA. 192

Figure 8.5. AlphaLISA profiling of 38 published antibodies. Data are a more quantitative depiction of the data for 38 of the antibodies in the heatmaps in Fig. 3.2a-b. (A) Schematic depicting AlphaLISA setup for measuring sdFab assembly.(B) AlphaLISA measurement of sdFab assembly. (C) Schematic depicting AlphaLISA setup for measuring S trimer binding.(D) AlphaLISA measurement of sdFab binding to the SARS-CoV-2 S trimer. (E) Schematic depicting AlphaLISA setup for measuring RBD binding. (F) AlphaLISA measurement of sdFab binding to the SARS-CoV-2 RBD. (G) Schematic depicting AlphaLISA setup for measuring sdFab competition with ACE2 for the RBD. (H) AlphaLISA measurement of sdFab competition

with ACE2 for the RBD. All AlphaLISA data are the mean of three independent replicates \pm the SEM. The dashed line indicates three standard deviations away from the background. Samples determined not to be significantly distinguished from the background ($p > 0.05$ two-sided t-test corrected using the FDR procedure) have bars that are filled dark grey. The samples are ranked within each category from worst (left) to best (right) neutralizers by their minimum neutralization IC_{50} value. 194

Figure 9.1. Site saturation mutagenesis of RBD to predict MON1 (LCB1v2.2) and MON2 (AHB2) escape mutants. (A-C) FACS gates for deep mutational scanning selections. Duplicate yeast display libraries expressing all mutations tolerated in the SARS-CoV-2 RBD were labeled with LCB1_v2.2 (A-B) or AHB2 (C). The population of cells indicated escaping minibinder binding was sorted and deep sequenced, enabling calculation of the “escape fraction” of each RBD mutation (the fraction of yeast expressing that mutant that were found in the escape bin). (D) Miniprotein escape profiles averaged across duplicate deep mutational scanning library selections. Line plots, left, show the total escape (sum of per-mutation escape) at all sites in the RBD. Sites marked in pink are shown in logoplots, right. Logoplots illustrate per-mutation escape by the height of letters reflecting individual amino acid mutations. Letters are colored according to their previously measured effects on ACE2-binding affinity³⁸⁹, according to the scale bar at right. (E-F) Correlation in escape between duplicate libraries, at the level of individual mutations (E) or total escape per site (F). 200

Figure 9.2. Cell-free DNA assembly and protein synthesis of multivalent minibinders. (A) A cell-free workflow for the expression and evaluation of multivalent minibinders. (B) Step-by-step workflow for cell-based DNA assembly and protein expression. (C) Standard step-by-step workflow for cell-free DNA assembly and protein expression. (D) Agarose gel of linear DNA templates for CFPS assembled via Gibson assembly and amplified via PCR. (E) SDS-PAGE of purified proteins expressed via CFPS. 202

Figure 9.3. Off-rate screen of multivalent minibinders. Dissociation of minibinder constructs from RBD or S6P over time was measured via competition AlphaLISA (mean \pm SEM, $n=3$ technical replicates from a single independent experiment). (A) Dissociation of FUS proteins with G10 linkers from RBD measured via competition with MON1. (B) Dissociation of FUS proteins with G10 linkers from S6P measured via competition with MON1. (C) Dissociation of TRI1 proteins from S6P measured via competition with TRI2-1. (D) Dissociation of TRI3 proteins from S6P measured via competition with TRI2-1. (E) Dissociation of

TRI2 proteins from S6P measured via competition with TRI2-1. (F) Dissociation of FUS proteins from S6P measured via competition with FUS231-P24.203

Figure 9.4. Kinetics of MON1, MON2, and MON3 binding to the SARS-CoV-2 RBD. Kinetic curves of Minibinders MON1 at 2-fold dilutions from 20 – 0.625 nM, MON2 at 2-fold dilutions from 10 – 0.625 nM, and MON3 at 3-fold dilutions from 20 – 0.25 nM were analyzed against monovalent soluble RBD covalently linked to AR2G biosensors. (A) Sensorgrams of minibinders binding to RBD ($n = 1$ technical replicate from a single experiment). Red curves are fits obtained from 1:1 Langmuir binding model. Sensorgram images were copied directly from export files from ForteBio Data Analysis v11.0 software without modification. (B) Summary of determined kinetic and equilibrium binding parameters for each minibinder. The < symbol indicates a cut-off of less than 5% dissociation observed during the dissociation phase, indicating insufficient time to accurately quantify the dissociation rate constant. The reported rate constants and equilibrium binding constants are estimates due to the very tight binding of the minibinders studied.204

Figure 9.5. Heat-based purification of mono- and multivalent minibinders. Proteins were expressed in *E. coli* using standard protein overexpression procedures. Harvested cells were heat treated at 70 °C for 5 min to lyse cells and precipitate cellular proteins. Bands observed in the Heat insoluble lane are indicative of protein precipitated during the heat treatment step. Bands observed in the Heat soluble step are indicative of proteins that did not precipitate during the heat treatment step. Data are from a single experiment. (A) SDS PAGE of MON1. (B) SDS PAGE of FUS231-P12. (C) SDS PAGE of TRI2-2.....205

Figure 9.6. Negative stain EM analysis of minibinder-mediated crosslinking/aggregation of S trimers. Representative electron micrograph and 2D class averages of SARS-CoV-2 S in complex with TRI2-2 (A), FUS231-G10 (B), FUS231-P24 (C), FUS31-G8 (D), TRI1-5-G2 (E), TRI1-5-G4 (F), TRI1-5-G6 (G) or FUS31-G10 (H). Scale bar: 100 nm. After two-dimensional classification, the number of particles assigned to classes with well-defined single (blue) or two neighboring cross-linked (orange) S trimers are presented as pie charts. The fraction of total cross-linked S trimers is underestimated since higher-order cross-linked trimers did not yield well-defined 2D averages.....206

Figure 9.7. Cryo-EM data collection and processing of the S6P/TRI2-2, S6P/FUS31-G10, and S6P/FUS231-P24 datasets. (A-D) Representative electron micrograph and 2D class averages of SARS-CoV-2 S in complex with TRI2-2 embedded in vitreous ice. Scale bar: 100 nm. (C) Gold-standard Fourier

shell correlation curves for TRI2-2-bound SARS-CoV-2 S trimer (black line) and locally refined RBD/TRI2-2 region (gray line). The 0.143 cutoff is indicated by a horizontal dashed line. (D) Local resolution maps calculated using cryoSPARC for the whole reconstruction as well as for the locally refined RBD/MON2 region. (E) Ribbon diagram of the RBD/MON2 designed model (pink) superimposed with the MON2 cryoEM structure (cyan). (F) MON2 designed model (pink) superimposed with the S6P/TRI2-2 cryoEM structure (cyan) with side chains displayed as sticks. (G-J) Representative electron micrograph (G), 2D class averages (H), gold-standard Fourier shell correlation curve (I) and local resolution map (J) of SARS-CoV-2 S in complex with FUS231-P24 embedded in vitreous ice. (K-N) Representative electron micrograph (K), 2D class averages (L), gold-standard Fourier shell correlation curve (M), and local resolution map (N) of SARS-CoV-2 S in complex with FUS31-G10 embedded in vitreous ice.207

Figure 9.8. Cryo-EM processing scheme of SARS-CoV-2 S/TRI2-2 complex.209

Figure 9.9. Cryo-EM processing scheme of SARS-CoV-2 S/FUS31-G10 complex.210

Figure 9.10. Cryo-EM processing scheme of SARS-CoV-2 S/FUS231-P24 complex.211

Figure 9.11. FUS231-P12 BRET sensor does not detect monomeric RBD. (A) Luminescence emission spectra and image of the BRET sensor (100 pM) in the presence (yellow trace, 100 pM) and absence (blue trace) of RBD. (B) Titration of RBD at 100 pM sensor (Mean \pm SEM, n=3 technical replicates, error bars that aren't visible are obscured by the marker).212

Figure 9.12. Kinetic analysis of interactions between multivalent minibinders and SARS-CoV-2 S protein variants using surface plasmon resonance. Multivalent minibinders were injected at concentrations ranging from 20 nM to 0.08 nM in three-fold serial dilutions against S protein variants covalently coupled to the chip via amine coupling. The double-referenced data were fit globally to a simple 1:1 Langmuir binding model in Carterra's Kinetic tool. (A) Sensograms of FUS231 proteins binding to the RBD, S6P, and S6P mutants (n = 1 technical replicate from a single independent experiment). (B) Summary of kinetic and equilibrium parameters of the measured interactions. *All listed K_D values are apparent K_D s due to the potential for multivalency. The < symbol indicates a cut-off of less than 5% dissociation observed during the dissociation phase, indicating insufficient time to accurately quantify the dissociation rate constant.213

Figure 9.13. Competition of ACE2 and mini binder constructs for S6P. (A) Competition ELISA curves (mean, n=2 technical replicates representative of two independent experiments). (B) Summary of minibinder

construct competition IC₅₀ values. An * denotes an IC₅₀ value less than 2-fold lower than the limiting concentration of S6P variant present in the well. N/A indicates an IC₅₀ value higher than the tested concentration range and greater than 20,000 pM. (C) Summary of EC₅₀ values of S6P variants binding to ACE2 used as the concentration for minibinder construct competition. (D) Ratio of mutant to wild type (WT) IC₅₀ values of minibinder constructs. The ratio in cells containing a slash was determined using the highest tested minibinder construct concentration which is indicative of a measurement at the upper limit of detection.....215

Figure 9.14. Neutralization of additional SARS-CoV-2 variants. (A) Neutralization of SARS-CoV-2 pseudovirus variants by minibinder constructs (mean, n=2 technical replicates representative of two independent experiments). (B) Neutralization of SARS-CoV-2 variants by minibinder constructs (mean ± SEM, n=4 technical replicates from two independent experiments). (C) Table summarizing neutralization potencies of multivalent minibinder constructs against SARS-CoV-2 pseudovirus variants. N/A indicates an IC₅₀ value above the tested concentration range and an IC₅₀ greater than 50,000 pM. (D) Table summarizing neutralization potencies of multivalent minibinder constructs against SARS-CoV-2 variants.217

Figure 9.15. Representative confocal images of human kidney organoids. Organoids were derived from H9 human embryonic stem cells (LTL, *Lotus tetragonolobus* lectin, proximal tubule marker in magenta; PODXL, podocalyxin, podocyte marker in red).218

Figure 9.16. Replicate plates for VSV-SARS-CoV-2 escape studies. Plaque assays were performed to isolate VSV-SARS-CoV-2 chimera virus escape mutants against a control neutralizing antibody (2B04) and the FUS231-P12 and TRI2-2 multivalent minibinders. Large plaques are indicative of escape. FUS231-P12 and TRI2-2 replicates were performed in three separate experiments consisting of two plates each. (A) No inhibitor in overlay. (B) 2B04 neutralizing mAb in overlay. (C) FUS231-P12 in overlay. (D) TRI2-2 in overlay.219

Figure 9.17. Measurement of SARS-CoV-2 viral titers in the lung of K18-hACE2-transgenic mice post intranasal delivery of TRI2-2. At 6 dpi (B.1.351) or 7 dpi (B.1.617.2), animals (n = 6 from two independent experiments) were sacrificed and analyzed for SARS-CoV-2 viral titers in the lung by plaque assays (line at median, two-tailed Mann-Whitney test: ** P < 0.01).....220

| | |
|---|-----|
| Figure 9.18. Pharmacokinetics of TRI2-2 delivered intranasally in C57BL/6J mice. (A) Serum and tissues were collected at 1, 8, 24, and 48 hours post-treatment with 200 μ g of TRI2-2 and quantitative competition ELISAs were used to determine concentrations of TRI2-2 present in the samples (n = 3 from a single experiment). The drawn line connects the mean values at each time point. (B) Raw data from competition ELISAs. Each curve represents a biological replicate with each data point consisting of n = 2 technical replicates. | 221 |
| Figure 9.19. Sequence alignment of AHB2 and ACE2. The main interacting helix of AHB2 and ACE2 were aligned using the Clustal Omega tool ⁴⁹² . Conserved residues between the two sequences are highlighted in cyan. | 222 |
| Figure 10.1. SEM micrograph of a microcantilever device. | 232 |
| Figure 10.2. Histogram of measured error for all cantilever deflection measurements in this manuscript. Values are binned \pm 0.1 nm of the center value. Dashed line indicates a gaussian fit (mean = 0.65, standard deviation = 0.36). | 233 |

List of Tables

| | |
|--|-----|
| Table 1.1. Strategies for protecting linear DNA in CFPS. | 40 |
| Table 1.2. Examples of cell free synthesized and functionally active, complex proteins. Abbreviations: scFv: antibody single-chain variable fragment, vtPA: variant of human tissue-type plasminogen activator, GM-CSF: granulocyte macrophage colony stimulating factor, IGF-I: insulin-like growth factor I, cIFN- α : consensus human interferon-alpha, rhGM-CSF: human granulocyte macrophage colony-stimulating factor, PD: Haemophilus influenzae protein D, IgG: Immunoglobulin G, β 2AR: β 2 adrenergic receptor, GPCR: G protein coupled receptor. *Indicates a non-batch CFPS reaction yield. An ~ indicates that the exact yield was not reported..... | 43 |
| Table 7.1. Design Sequences. Amino acid sequences of all designs tested. Two versions of DHD 15 heterodimers were used here, with slight differences in the loop sequence. For simplicity they are all labeled 2 and 2', differentiated with an additional asterisk (*). See attached Supplementary Table 1.xlsx for design sequences and Vectors.zip for plasmid maps. These materials have been deposited to zenodo and can be accessed at the following link: https://doi.org/10.5281/zenodo.6377529 | 167 |
| Table 7.2. SAXS analysis statistics. Rg is the radius of gyration, Rc is the cross-sectional radius of gyration determined from Guinier fitting, and Px is the Porod exponent. The Rc value for most designs cluster around 12 Å, in a close agreement with design models. | 168 |
| Table 7.3. CD data parameter fitting. Fitted parameters for equilibrium chemical denaturation. Errors represent fitting errors. | 169 |
| Table 7.4. Native MS spectra of individual designs. 5 μ L protein (10-100 pmol) were injected and detected by online buffer exchange coupled to an Exactive Plus EMR Orbitrap instrument. Spectra were deconvoluted by UniDec. Species and determined molecular weights are depicted. Signals corresponding to monomers are indicated by an orange circle, dimers are indicated by a red square. | 170 |
| Table 7.5. Protein concentrations native MS titration experiments. | 174 |
| Table 7.6 Native MS spectra of complexes. Samples were prepared by mixing equimolar amounts of 1, 2, and the corresponding dimerizer(s) (i.e. 1'-2' for induced dimerization, 1'-3' + 3-2' for 2-in AND, linker 1'-4+ 4-3'+ 3-2'+ for 3-in AND). After denaturation, refolding and dialysis against 200 mM AmAc, the mixtures were diluted 12.5 fold with 200 mM AmAc and 5 μ L were flow-injected into an Exactive Plus EMR Orbitrap | |

instrument. Spectra were deconvoluted by UniDec. Species and determined molecular weights are depicted. Signals corresponding to the 1-(dimerizer)_n-2 complex are indicated by a red circle, free dimerizer is indicated by a yellow star..... 175

Table 7.7. Detailed native MS titration spectra of AND logic gates. Samples were prepared by mixing 10 μM 1 and 2 with 0 – 60 μM of (A) 1'-2' with 0-residue linker, (B) 1'-2' with 2-residue linker, (C) 1'-2' with 6-residue linker, (D) 1'-2' with 12-residue linker, (E) 1'-2' with 24-residue linker, (F) 1'-3' + 3-2', and (G) 1'-4+ 4-3'+ 3-2'. After denaturation, refolding and dialysis against 200 mM AmAc, the mixtures were diluted 12.5 fold with 200 mM AmAc and 5 μL were flow-injected into an Exactive Plus EMR Orbitrap instrument. Signals corresponding to the full trimer complexes in (A) - (E), full tetramer complex in (F) and full pentamer complex in (G) are indicated by a red circle. Free dimerizer/ dimerizer complexes are indicated by a blue triangle. The spectra for each titration were normalized and stacked to visualize the change in complex abundance in dependency of added dimerizer(s). As indicated by a red arrow and dashed line, in (G) relative intensities for signals with $m/z > 4300$ are magnified 4x, and in (E) relative intensities for signals with $m/z > 4200$ are magnified 5x for $m/z > 4200$ to help visualize formation of full complexes in the presence of the easily ionizable dimerizers. Lists of all identified and quantified species in the titration has been deposited. ... 177

Table 7.8. Native MS spectra of serial dilution experiments. Samples at 60 μM monomer concentration were prepared by denaturation, refolding and dialysis against 200 mM AmAc. Proteins were diluted to the indicated concentrations (4.8 μM to 0.0375 μM). 5 μL (24 pmole to 0.1875 pmole) were flow-injected into an Exactive Plus EMR Orbitrap instrument. Spectra were deconvoluted by UniDec. Species and determined molecular weights are depicted. Signals corresponding to monomers are indicated by an orange circle, dimers are indicated by a red square. Total ion chromatograms (1000-8000 m/z) are depicted for flow-injection profiles..... 181

Table 8.1. Summary of antibody screening studies designed to identify SARS-CoV-2 neutralizing antibodies. Studies were evaluated for their efficiency at identifying potent neutralizing antibodies. For each study, the total number of antibodies evaluated as well as the number of neutralizing antibodies with a neutralization IC_{50} less than 0.01 $\mu\text{g}/\text{mL}$ and 0.25 $\mu\text{g}/\text{mL}$ were summarized. Either authentic- or pseudovirus neutralization IC_{50} was considered based on the breadth of the reported data. The 0.01 $\mu\text{g}/\text{mL}$ cutoff was chosen because this value is approximately within an order of magnitude of the most potent reported

neutralizing antibodies³³⁸. 0.25 µg/mL was chosen as a practical cutoff for moderately potent neutralizing antibodies³³⁵. For Hansen *et al.*³⁴⁴ the neutralization potencies were reported in M and were converted to µg/mL assuming an antibody molecular weight of 150 kDa. For this analysis, only the studies containing antibodies expressed in this manuscript were used. Furthermore, only studies whose purpose was to identify neutralizing antibodies from a large set of candidates were considered. Studies that didn't describe the results of their antibody discovery process in sufficient detail to collect the desired information were omitted. NR indicates not reported in sufficient detail to determine. 195

Table 8.2. Reported parameters and expected behaviors for tested antibodies. Information about antibody target epitope, pseudo- or authentic virus neutralization IC₅₀, and equilibrium dissociation constant from literature about the antibodies used in this study are summarized. Data are presented in two tabs, one for antibodies from diverse sources and one for the Brouwer *et al.* data set, in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>. 196

Table 8.3. Variable heavy and light chain sequences of tested antibodies. Sequences are classified by their heavy (VH) or light chain (VL) as well as the light chain class (kappa or lambda). The variable domain protein sequence, the variable domain *E. coli* codon-optimized DNA sequence, and the ordered DNA sequence containing all additional (Gibson assembly homology, n-terminal expression tag, etc.) sequences are listed. Note that the antibodies COVA2-15 and COVA1-18 were ordered and evaluated twice in this data set, and there are thus two separate entries for the two sets of sequences. Sequences are listed in a single tab in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>. 197

Table 8.4. Summary of observed antibody binding significance statistics and comparison to literature. Results from the AlphaLISA screen are compared against literature data and a qualitative match with the reported values are summarized. Comments on the possible origin of observed inconsistent results are listed for all data with reported neutralization IC₅₀ values less than 10 µg/mL. Data are presented in two tabs, one for antibodies from diverse sources and one for the Brouwer *et al.* data set, in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>. 198

Table 8.5. Raw and processed data for AlphaLISA measurements for the 120 antibodies evaluated in this manuscript. The individual replicates, average, standard deviation, coefficient of variation, p-value from a two-sided t-test against the background, and the FDC corrected p-value are reported. Note that the antibodies COVA2-15 and COVA1-18 were ordered and evaluated twice in this data set, and there are thus two separate entries for the two sets of sequences. Data are presented in four tabs, one for each AlphaLISA measurement modality, in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>. 199

Table 9.1. List of abbreviations used to describe multivalent minibinders in this manuscript.223

Table 9.2. Oligomerization domains tested in this work.225

Table 9.3. CryoEM data collection and refinement statistics.....227

Table 9.4. List of DNA and protein sequences for multivalent minibinders used in this manuscript (Separate csv file). DNA sequences are the open reading frame coding for the expressed protein. Protein sequences are annotated as follows. Minibinder and homotrimer sequences are denoted by square brackets []. Secondary sequences (e.g., expression tag, purification tag, etc.) are annotated by parenthesis (). Non-minibinder or non-homotrimer sequences are annotated by curly brackets {} (captures linkers and other secondary sequences). These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>. 229

Table 9.5. Estimates of the diversity of mutants in the VSV-SARS-CoV-2 chimera virus pool used in the multivalent minibinder escape selections. Estimates were calculated using the Pedel-AA tool ⁴⁵² for analyzing diversity in randomly mutated protein libraries. Upper and lower estimates for the VSV RNA polymerase L error rate ⁴⁰⁸ were used to calculate the range of plausible library compositions.230

Table 9.6. Comparison of multivalent minibinder and FDA authorized neutralizing antibody potencies. For mAb data, pseudovirus neutralization IC₅₀ values (columns 3, 4, and 5) were taken from ³⁷⁰ and authentic virus neutralization IC₅₀ values (columns 6, 7, and 8) were taken from ³⁷¹. *Value for WA1/2020 strain as opposed to WA1/2020 D614G.231

Table 10.1. Summary of the measured RT-qPCR Ct values, lineage designations, and GISAID identifiers for all reported patient samples (as available).....234

Table 10.2. Summary of the measured RT-qPCR Ct values, their corresponding concentration of RNA (copies/mL), and their measured deflection (\pm standard deviation) for all reported patient samples.235

Chapter 1 - Introduction

1.1 Motivation

At the outset of my PhD, I was fascinated by the amazing properties of biological systems, including but not limited to their ability to self-replicate, to sense and respond to their environment, to efficiently produce complex molecules from simple molecules, and much more. In particular, I became fascinated by proteins, the sequence-defined polypeptide polymers which act as macromolecular machines carrying out nearly every biological process¹. The phenomenal features of proteins have led to enormous efforts by scientists and engineers not only to understand their function, but also to repurpose proteins to solve pressing societal problems. A few recent compelling examples of this include treatments for COVID-19²⁻⁵, multiplexed pathogen diagnostics⁶, enzymes that degrade non-natural pollutants⁷, low-cost water quality diagnostics⁸⁻¹⁰, materials proteins as replacements for hydrocarbon-based synthetic materials^{11,12}, synthetic carbon fixation pathways^{13,14}, and carbon negative manufacturing of industrially relevant chemicals¹⁵. The possibilities for utilizing proteins, and biological engineering more broadly, to tackle the major challenges of the 21st century is enormous¹⁶⁻¹⁹.

Despite this potential, efforts to identify, engineer²⁰, or design²¹ proteins with the desired function often are slow and are limited by the methodologies available to researchers to express and functionally characterize proteins. In the first year of my PhD, I worked on a project that relied on the production of a protein in living *Escherichia coli* and found my research stifled by the inability to actualize experiments. I spent most of my time cloning plasmids, culturing cells, and purifying proteins instead of engineering protein function. Something that I learned about myself during my PhD is that I would rather spend my time trying to improve or circumvent slow or cumbersome processes rather than trying to ask the questions I am interested in with tools that only allow for incremental progress. As a result, I found the problem of high throughput protein expression and evaluation to be highly motivating, and it became a through line early on in my PhD.

When trying to identify proteins for a given function — whether it be through identifying natural homologs or engineering or designing a protein — there are two general approaches taken by researchers. The first is pooled library-based approaches, where all the variants are pooled together, and variant function

must be decoupled using various types of a selection or sorting step. The second is individual screening of the desired variants, where each variant is processed and analyzed separately. These two approaches are highly complementary and are frequently combined within a given workflow. Due to the complex nature of proteins, both approaches typically requires the hijacking of cellular machinery to manufacture the protein of interest, though in very select cases they can be chemically synthesized²². For decades recombinant protein production in living hosts — typically in *Escherichia Coli* (*E. coli*)²³, yeasts²⁴, insect cells²⁵, or mammalian cells²⁶ — has been the gold standard technique for protein manufacture. Historically, cell-based expression has been a powerful approach for pooled library experiments²⁷ but has been of more limited utility for the high-throughput screening of individual proteins. Techniques for the high-throughput production of individual proteins using cells in 24 or 96 well plates do exist^{28,29}, but they rely on poorly scalable and time-consuming steps including plasmid purification, transformation of the host organism with the cloned plasmid, and cell culture to manufacture the protein of interest²⁷. An excellent example of the discrepancy between the level of effort required to perform the pooled library and individual screening approaches is in modern antibody discovery workflows where the pooled library step to select active binding proteins from billions of variants takes 2-3 days and the subsequent individual evaluation of the hundreds of identified candidates takes a week or more²⁹. Furthermore, it is usually the case that only a fraction of possible hits are evaluated due to the bottleneck of individual screening³⁰. It is also important to note that many protein functions, particularly enzymatic activity²⁷, are often difficult to evaluate in pooled library approaches, and thus require individual screening²⁷. To me it was clear that better technologies to accelerate the individual screening approach are needed.

Cell-free protein synthesis (CFPS), the manufacture of proteins with non-living cellular components, is a promising alternative to cell-based protein expression because it can overcome the limitations of cell-based approaches for individual screening. This is primarily because CFPS converts a process that requires cultivating living cells into one that only requires the combination of various non-living reagents, which enables protein production in hours, the ability to use linear DNA as a template for protein synthesis, and straightforward integration with liquid handling robotics among other benefits. These advantages allow researchers to go from synthetic DNA to a functionally evaluated protein without ever going into cells and in a matter of hours. I felt that these crucial differences could enable a hundred fold or

even greater improvement in throughput for individual functional screens over the typical hundreds or thousands of proteins screened with current methods²⁷, which has precedent in the small molecule drug discovery world high-throughput screening is regularly used to individually evaluate compound libraries containing millions of distinct compounds^{31,32}. This area, high throughput functional protein screening, is where I decided to focus my PhD efforts.

I was initially unsure what to apply a high throughput protein screening platform to but quickly became excited about de novo protein design, or the design of proteins from scratch, which has the potential to make custom proteins for any application where proteins are used²¹. After consideration and discussion, Mike Jewett and I decided to reach out to David Baker, a leader in the field of protein design, about trying to meld high throughput screening in CFPS and protein design. We initially collaborated on the design of protein heterodimer-based logic gates that integrate signals in both cell and cell-free environments³³. However, the coronavirus disease 2019 (COVID-19) pandemic forced me to reevaluate the possible uses of the high-throughput screening methods I'd developed, and I next turned to the design and discovery of protein therapeutics. The COVID-19 pandemic highlighted the limitations of modern antibody screening workflows and motivated me to apply the high-throughput CFPS to the antibody discovery pipeline³⁴. Furthermore, in a continued collaboration with the Baker lab, we designed de novo protein therapeutics for the treatment of COVID-19 that prevent the SARS-CoV-2 virus from engaging its host receptor³⁵. As a byproduct of this work, I also worked with the Dravid lab at Northwestern University to manufacture nanomechanical sensors for COVID-19 detection based on the same proteins we'd developed as COVID-19 therapeutics.

1.2 Dissertation Roadmap

My dissertation focuses largely on the development of methods that leverage cell-free protein synthesis and liquid handling robotics to accelerate the ability of researchers to build and test individual proteins and to discover and engineer protein-protein interactions. Each chapter is an individual work that has either been published in or is intended for submission to a peer reviewed journal. What follows is a brief roadmap of the chapters of my dissertation. **Chapter 2** details the design and implementation of protein-protein interaction-based logic gates built from de novo designed proteins. This work was led by

Zibo Chen and I collaborated with him and others to both evaluate the computationally designed protein heterodimers and to prototype logic gates. Our results paved the way for alternate mechanisms for information processing in engineered biological systems and laid the groundwork for the development of a high-throughput protein-protein interaction screening workflow. In **Chapter 3** I develop an end-to-end automatable system for manufacturing and screening antibodies with cell-free protein synthesis. This chapter details the developed high-throughput protein screening workflow in its fullest embodiment and illustrates that the workflow can be used to go from DNA to functionally evaluated protein in less than 24 hours using cell-free protein synthesis and an Echo acoustic liquid handler for hundreds of proteins. **Chapter 4** comprises a collaborative effort with Brett Case, Young-Jun Park, Longxing Cao, Kejia Wu, Alexandra Walls, and many others to design and validate multivalent minibinder therapeutics for the treatment of COVID-19. Here I describe our effort to design proteins that inhibit a key host pathogen interaction that enables the SARS-CoV-2 viruses' ability to infect human cells. I leveraged many of the developments in Chapter 3 to accelerate this effort and guide the selection of the lead candidates. We found that our lead candidates protected from SARS-CoV-2 challenge in mice indicating that this type of molecule represents a promising therapeutic modality against SARS-CoV-2 and potentially other pathogens. In **Chapter 5** I describe a collaborative effort with Dilip Agarwal to adapt the protein designed in Chapter 4 to build a diagnostic for SARS-CoV-2 infection. We built a nanomechanical sensor that enables rapid detection SARS-CoV-2 antigen at levels equivalent to the limit of detection of gold standard RT-qPCR diagnostics. In **Chapter 6** I conclude my dissertation by discussing my perspectives on possible future directions and the accomplishments of my work.

1.3 Background

1.3.1 Cell free protein synthesis

Cell free proteins synthesis (CFPS, also referred to as in vitro transcription translation (IVTT or TXTL)^{36,37} or cell free gene expression (CFE)³⁸) is the activation of transcription and translation from non-living cellular components. The transcription and translation machinery required to drive CFPS typically comes from either crude cellular extract from various organisms³⁹ (e.g., *E. coli*, CHO, etc.) or from purified components as in the PURE system⁴⁰. CFPS reactions are run by combining these components with a

template and a cocktail of other molecules required to activate and sustain transcription and translation⁴¹. There are two typical categories of CFPS reaction format, batch reactions where all components are added initially and semi-continuous reactions where a semi-permeable membrane is used to enable exchange of small molecules during the reaction³⁹. These reaction formats lead to different reaction durations and yields, with batch reactions reaching up to 4 mg/mL³⁷ and >20 hours⁴² of protein synthesis and semi-continuous reactions reaching yields of 8 mg/mL³⁷ and 14 days⁴³ of sustained protein synthesis. CFPS reactions are highly scalable, with batch reaction volumes ranging from femtoliters in microfluidic systems^{44,45} to >100 L in traditional large-scale bioreactors⁴⁶. In this background section, I will review the elements of CFPS systems relevant to this dissertation.

1.3.2 Templates for CFPS

The templates used to drive CFPS take a variety of forms (Fig. 1.1). The unifying feature of these templates is that they contain the appropriate elements to activate transcription (e.g., promoter and terminator) and or translation (e.g., ribosome binding site (RBS) and open reading frame (ORF)) in the CFPS system of interest (reviewed elsewhere⁴⁷). In vitro transcribed (IVT) mRNA is commonly used to drive uncoupled, translation only CFPS reactions. However, a separate IVT step is not necessary for many CFPS systems which can perform transcription and translation from either circular or linear DNA in a one pot reaction with high yields. Furthermore, coupled transcription and translation reactions are most suitable for high throughput experimentation due to their one pot nature. Genomic DNA and RNA, for example phage genomes⁴⁸, have also been used as templates for CFPS.

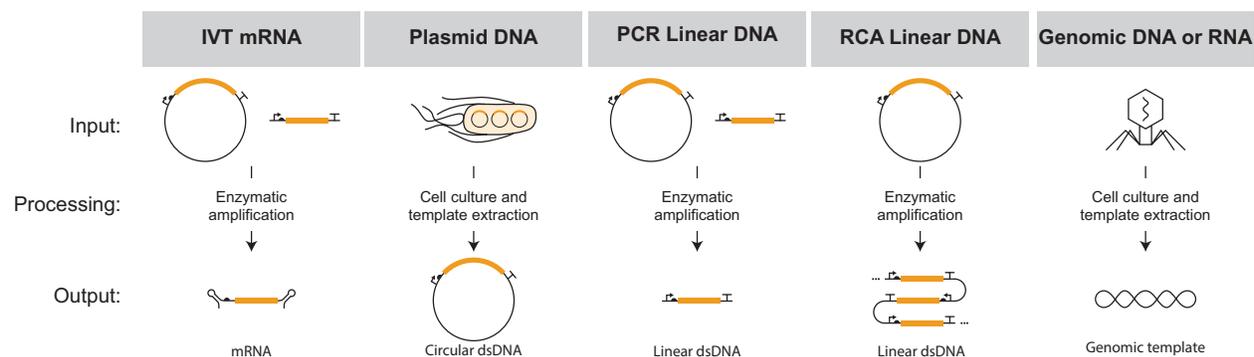


Figure 1.1. Templates used to drive transcription and or translation in CFPS systems. IVT: in vitro transcription; DNA: deoxyribonucleic acid; mRNA: messenger ribonucleic acid; PCR: polymerase chain reaction; RCA: rolling circle amplification.

Linear DNA templates, frequently referred to as linear expression templates (LETs) (reviewed elsewhere^{49–51}), are particularly attractive because they circumvent the time consuming and low-throughput steps associated with plasmid assembly and isolation from an organism, saving days of effort and enabling throughputs much greater than cell-based workflows^{34,52–54}. Generally, these methods work by coupling an assembly step building an LET capable of supporting transcription and translation to an amplification step manufacturing the template in sufficient quantity to drive a CFPS reaction. The assembly step typically involves assembly of the gene with other regulatory elements (e.g., promoter, ribosome binding site, terminator, etc.) and potentially with other components of the ORF. To date, assembly has been accomplished with Gibson assembly^{34,35,53,55}, golden gate assembly⁵³, gateway assembly⁵⁴, or polymerase chain reaction (PCR) based assembly^{36,43,56,57}. After assembly, LETs are subsequently amplified to generate sufficient quantities to drive a CFPS reaction either by the polymerase chain reaction (PCR)^{36,53,54,57–60} or rolling circle amplification (RCA)^{52,54,61,62}. PCR generates linear dsDNA containing only a single copy of the template and requires thermal cycling. RCA leverages random primers and a strand displacing polymerase to generate large fragments of dsDNA containing many copies of the template in a single molecule. Due to the nature of the assembly and amplification reactions being in vitro enzymatic reactions, they can be easily carried out in well plates, making them particularly valuable for high-throughput screening efforts.

A major consideration in the use of linear DNA in CFPS is that crude extract-based systems contain nucleases which degrade LETs which can hinder protein expression. Interestingly, cell culture and extract processing steps play a major role in the ability of a given CFPS system to utilize LETs, for example *E. coli* extracts that are not processed with runoff reaction or dialysis steps yield protein expression within 50% of that of plasmid-based expression³⁴. However, frequently methods to protect LETs from nuclease degradation are required and there are a variety of methods that have been developed for this purpose (Table 1.1). The RecBCD exonuclease complex is known as a major contributor to the stability of linear DNA in *E. coli*⁶³ and has been the major target for inhibition in *E. coli* CFPS systems. Strategies tested to inhibit RecBCD include genomic modification^{58,64,65}, competitive substrate inhibitors⁶⁶, protein inhibitors⁵³, and small molecule inhibitors⁶⁷ which have all yielded varying degrees of success and are summarized in Table 1.1. The success inhibiting the RecBCD exonuclease in *E. coli* has led to other strategies to inhibit

exonucleases more broadly that may be applicable in CFPS systems derived from other organisms. Several different DNA modifications with potential to increase nuclease resistance, including methylation⁶⁸ and chemical modification^{53,65}, have been evaluated but yielded little to no improved protein expression. A more successful general strategy has been to utilize DNA binding proteins that block the terminal ends of the LET^{68–70}, which in one case resulted in LET-based expression on par with that of plasmid-based expression⁶⁹. Circularization of the linear template DNA has also been a successful strategy for protecting against degradation, yielding similar results between linear and plasmid based templates⁷¹.

Table 1.1. Strategies for protecting linear DNA in CFPS.

| Strategy | Implementation | Mechanism | Result | Organism | Reference |
|----------------------|--|---|------------------------------------|--|---------------------------------------|
| Genomic modification | A19ΔrecCBD::Plac-red-kan-ΔendA | Replacement of RecBCD with bacteriophage λ red recombination system | 45-63% of plasmid yield | <i>E. coli</i> | Michel-Reydellet et al. ⁵⁸ |
| Genomic modification | SBP tagged RecD and PNPase | removal of RecD and PNPase from extract | 2-4x improved LET yield | <i>E. coli</i> | Seki et al. ⁶⁴ |
| Genomic modification | ΔrecCBD | Knockout of RecBCD | 48–78% of plasmid yield | <i>E. coli</i> | Batista et al. ⁶⁵ |
| Nuclease inhibition | Chi6 oligos | Competitive inhibition of RecBCD | 23% of plasmid yield | <i>E. coli</i> | Marshall et al. ⁶⁶ |
| Nuclease inhibition | gamS | Protein inhibitor of RecBCD | 38% of plasmid yield | <i>E. coli</i> | Sun et al. ⁶³ |
| Nuclease inhibition | CID 697851, CID 1517823 | Small molecule inhibitor of RecBCD | 2.5-3x improved LET yield | <i>E. coli</i> | Shrestha et al. ⁶⁷ |
| Nuclease inhibition | Tus-Ter | Protein binds linear dsDNA termini | 100-146% of plasmid yield | <i>E. coli</i> , <i>V. natriegens</i> | Norouzi et al. ⁶⁹ |
| Nuclease inhibition | Ku | Protein binds linear dsDNA termini | improvement in transcription yield | <i>E. coli</i> and 4 non-model organisms | Yim et al. ⁷⁰ |
| Nuclease inhibition | ssCro | Protein binds linear dsDNA termini | 2-28% of plasmid yield | <i>E. coli</i> , <i>V. natriegens</i> | Zhu et al. ⁶⁸ |
| DNA modification | Terminal 5' phosphorothioate linkages (x2) | 5' DNA more resistant to nuclease degradation | 36% improved LET yield | <i>E. coli</i> | Sun et al. ⁶³ |
| DNA modification | Methylation | Methylation of linear DNA potentially inhibits degradation | No significant improvement | <i>E. coli</i> | Zhu et al. ⁶⁸ |
| Template design | Flanking DNA | 5-500 bp of flanking DNA as a buffer for exonucleases | 2.4-6x improved LET yield | <i>E. coli</i> | Sun et al. ⁶³ |
| Template design | Circularization | PCR-based circularization of LET | same as plasmid yield | <i>E. coli</i> | Wu et al. ⁷¹ |

There are several other factors to consider when choosing a template for CFPS. Template “quality” — both the purity of the DNA and damage to the DNA during purification — can have a large influence on CFPS productivity and repeatability⁷². Towards improving template preparation outcomes, best practices for template preparation have recently been suggested⁷³. Fortuitously, in applications utilizing linear DNA,

the buffers associated with both PCR and RCA have been found in some cases not to be inhibitory to CFPS, enabling the use of unpurified template^{34,54,55,61}. It should also be noted that linear DNA may not be suitable for all applications. In efforts to evaluate *E. coli* σ^{70} promoters in CFPS, multiple researchers have observed that the strength of a given promoter correlated better with in vivo results for plasmid based CFPS than for LET based CFPS^{53,74}. It was hypothesized that this was related to DNA supercoiling and its impact on transcription rate⁷⁴, and thus applications sensitive to transcription rate and DNA supercoiling more generally may require plasmid templates. While this may be a limitation for some applications for the time being, recent advances in cell-free plasmid replication systems may enable plasmid preparation with the same ease as linear DNA⁷⁵.

1.3.3 Quantifying the results of CFPS

CFPS systems are complex mixtures, and this complexity can convolute the quantification of the products of a CFPS reaction. Evaluating the functional output of the expressed molecule (e.g., fluorescence, binding, catalytic activity, etc.) or system of molecules (e.g., complex assembly, metabolic pathway productivity, etc.) is frequently the goal, however, quantifying the results of transcription and translation in the system is often crucial to understanding the functional result. Several methods to track RNA or protein production in CFPS systems have been developed that do not require modification of the expressed molecule. Stochastic incorporation of nucleotide radiolabels into RNA⁷⁶ and amino acid radiolabels into proteins⁷⁷ is widely utilized to quantify the molecules produced during the CFPS reaction and does not require modification of the expressed sequences. Transcription in the PURE CFPS system has been monitored without modifying the transcript using Förster resonance energy transfer (FRET) based binary DNA primer probes⁷⁸. As an alternative to radiolabels, tRNAs pre-charged with a non-canonical amino acid containing a trackable modification (e.g., a fluorophore or biotin) are also widely used to stochastically label the native protein sequences produced in CFPS⁷⁹. Furthermore, standard protein visualization techniques like Coomassie staining of a protein gel can visualize protein expression from an unpurified CFPS reaction provided enough protein is expressed⁸⁰. Several strategies for tracking transcription and translation that require modification of the mRNA or protein sequence have also been developed. Modification of the sequence can enable higher throughput experimentation but can interfere with the native function in some cases. To quantify transcription, it is common to fuse a fluorescent aptamer

to the RNA of interest^{8,42,81–83}. Protein fusions — including fluorescent proteins^{33,54}, high affinity luciferase complementation reporters⁸⁴, tetracysteine tags^{85,86}, and purification or antibody epitope tags for western blotting^{84,87–90} — have also been used to quantify the products of translation. Other fusions have also been suggested⁹¹, including a fluorescent protein complementation reporter⁹², protease cleavable fluorescent proteins^{54,93}, and nanoluciferase⁹⁴.

1.3.4 Expression of complex proteins using CFPS

Many proteins require specialized conditions or modifications to fold and be active many of which may not be recapitulated by basic CFPS systems. However, the flexibility of CFPS enables supplementation of the system with exogenous factors that enable proper folding or modification of the proteins. To date, CFPS systems have been designed that enable the expression of disulfide bonded proteins, post-translationally modified proteins, membrane proteins, multi-subunit proteins, as well as proteins that require chaperones to properly fold (Table 1.2).

Table 1.2. Examples of cell free synthesized and functionally active, complex proteins. Abbreviations: scFv: antibody single-chain variable fragment, vtPA: variant of human tissue-type plasminogen activator, GM-CSF: granulocyte macrophage colony stimulating factor, IGF-I: insulin-like growth factor I, cIFN- α : consensus human interferon-alpha, rhGM-CSF: human granulocyte macrophage colony-stimulating factor, PD: Haemophilus influenzae protein D, IgG: Immunoglobulin G, β 2AR: β 2 adrenergic receptor, GPCR: G protein coupled receptor. *Indicates a non-batch CFPS reaction yield. An ~ indicates that the exact yield was not reported.

| Year | Protein | Protein Class | CFPS System | Yield (μ g/mL) | Complex Elements | Reference |
|------|---|-----------------------------|----------------|---------------------|---|--------------------------------|
| 1997 | α -Hemagglutinin scFv | Antibody | <i>E. coli</i> | 8.3 | folding chaperone required, disulfide bonded | Ryabova et al. ⁹⁵ |
| 2003 | α -Salmonella O-antigen scFv | Antibody | Wheat germ | 13 | folding chaperone required, disulfide bonded | Kawasaki et al. ⁹⁶ |
| 2004 | urokinase protease | Protease | <i>E. coli</i> | 40 | disulfide bonded | Kim et al. ⁹⁷ |
| 2004 | vtPA | Protease | <i>E. coli</i> | 60 | folding chaperone required, disulfide bonded | Yin et al. ⁹⁸ |
| 2005 | scFv-GM-CSF fusion protein | Antibody Cytokine Fusion | <i>E. coli</i> | 43 | folding chaperone required, disulfide bonded | Yang et al. ⁹⁹ |
| 2005 | β 2AR | GPCR | <i>E. coli</i> | ~1,000* | membrane bound | Ishihara et al. ¹⁰⁰ |
| 2006 | IGF-I | Hormone | <i>E. coli</i> | 400 | disulfide bonded | Swartz ¹⁰¹ |
| 2008 | HydA1 | [Fe Fe] Hydrogenase | <i>E. coli</i> | 22 | cofactors, oxygen sensitive | Boyer et al. ¹⁰² |
| 2011 | rhGM-CSF | Cytokine | <i>E. coli</i> | 700 | folding chaperone required, disulfide bonded | Zawada et al. ⁴⁶ |
| 2011 | cIFN- α | Cytokine | <i>E. coli</i> | 400 | solubility issues | El-Baky et al. ¹⁰³ |
| 2011 | ATP Synthase | Molecular Machine | <i>E. coli</i> | NA | folding chaperone required, membrane bound, multi-subunit | Matthies et al. ¹⁰⁴ |
| 2014 | VH3-7/Vk3-20 IgG | Antibody | <i>E. coli</i> | ~1,500 | folding chaperone required, disulfide bonded, multi-subunit | Groff et al. ¹⁰⁵ |
| 2014 | ERBB2 (HER2) | Receptor Kinase | <i>E. coli</i> | 2 | membrane bound | He et al. ¹⁰⁶ |
| 2017 | α -NIP IgG | Antibody | CHO | 114* | folding chaperone required, disulfide bonded, multi-subunit, glycosylated | Martin et al. ¹⁰⁷ |
| 2017 | α -SMAD2 IgG | Antibody | CHO | 9* | folding chaperone required, disulfide bonded, multi-subunit, glycosylated | Stech et al. ¹⁰⁸ |
| 2019 | α -HER2 IgG | Antibody | PURE | 124 | folding chaperone required, disulfide bonded, multi-subunit | Murakami et al. ¹⁰⁹ |
| 2021 | PD glycosylated with <i>F. tularensis</i> O antigen | Glycoprotein | <i>E. coli</i> | 20 | glycosylated | Stark et al. ⁸⁷ |
| 2021 | PD glycosylated with <i>C. jejuni</i> glycan | Glycoprotein | <i>E. coli</i> | 43 | glycosylated | Hershewe et al. ⁹⁰ |

1.3.4.1 Folding chaperones

Many proteins, particularly large, multi-domain proteins, require the assistance of chaperones^{110,111} and other regulatory mechanisms like modulated translation rates¹¹² to properly fold. Interestingly, *E. coli* CFPS systems typically have a translation rate more than an order of magnitude lower than living *E. coli*¹¹³

and closer to that of eukaryotic systems¹¹⁴, which may be beneficial for proteins that require more time for co-translational folding¹¹². However, this alone is frequently not sufficient to enable proper folding of complex proteins in CFPS systems. Furthermore, purified CFPS systems by design contain no chaperones and extract based systems may or may not have sufficient levels of endogenous chaperones to adequately support proper folding, and thus CFPS systems are frequently supplemented with additional chaperones^{95,105}. Two different chaperones systems derived from *E. coli* — DnaK/DnaJ/GrpE and GroEL/GroES — have been used extensively in CFPS^{95,109,115}. In *E. coli*, these two systems work together to assist in proper protein folding and collectively provide additional time (DnaK system) and an isolated environment (GroEL system) for the protein of interest to fold^{110,111}. A ribosome targeted fusion of the HSP 70 chaperone BiP, which has a function analogous to that of DnaK, showed an improvement in soluble protein yields of eukaryotic proteins expressed in an *E. coli* CFPS system¹¹⁶. To improve antibody folding in CFPS, Groff et al.¹⁰⁵ supplemented chaperones known facilitate antibody folding in mammalian cells and found that the FkpA and SylD prolyl isomerases as well as the SkpA deaggregase improved the yields of soluble antibody¹⁰⁵. Artificial chaperone systems consisting of polysaccharide nanogels have also been shown to improve the soluble expression of aggregation proteins¹¹⁷, highlighting the highly tunable CFPS reaction environment can be engineered beyond what is naturally present inside the cell. CFPS has also been used as a strategy to produce chaperones and subsequently assess their impact on folding of a target protein, a strategy that could accelerate researcher's ability to identify the chaperones required for a given protein¹¹⁸. In addition to improving the ability to synthesize complex proteins, chaperone function and specificity can also be studied in cell-free systems as has been done for the DnaK/DnaJ/GrpE chaperone system¹¹⁹. In addition to traditional chaperones, fusion partners (e.g., maltose binding protein (MBP), thioredoxin (TRX), small ubiquitin-related modifier (SUMO), glutathione-S-transferase (GST), etc.) which are thought to improve the folding and solubility of proteins, have also been utilized in CFPS⁸⁴. Interestingly, proteases like the tobacco etch virus (TEV) protease are active in some CFPS systems, enabling production of protein fusions and subsequent cleavage of the fusion partner in a one-pot format⁸⁴. Collectively, these advances enable the manufacture of proteins with complex folding requirements in environments customized for the protein of interest.

1.3.4.2 Disulfide bond formation

Disulfide bonds are critical structural elements of many proteins that stabilize their 3D structure¹. Disulfide bonds are formed in specialized compartments or organelles — the endoplasmic reticulum in eukaryotes and the periplasm in prokaryotes — which maintain an oxidizing environment to enable spontaneous disulfide bond formation¹²⁰. These environments are enabled by systems of enzymes that maintain the redox environment and assist with the formation of the correct disulfide linkages¹²⁰, and both features of these compartments can be recapitulated in a cell free system. The non-living nature of the CFPS reaction allows for the oxidizing environment to be maintained through the addition of small molecules (typically oxidized (GSSG) and reduced (GSH) glutathione)¹²¹. Extract-based CFPS systems are also frequently treated with iodoacetamide (IAM), an alkylating reagent that inactivates the endogenous redox enzymes of the lysate, to stabilize the redox environment⁹⁷. Furthermore, the enzymes responsible for disulfide bond isomerization — protein disulfide isomerase (PDI) in eukaryotes¹²⁰ and DsbC in prokaryotes¹²⁰ — can be supplemented to enable the proper formation of disulfide bonds¹⁰⁵.

Many different proteins containing disulfide bonds have been successfully manufactured using CFPS including human hormones¹⁰¹, cytokines^{46,103}, other enzymes^{97,98}, and antibodies^{95,96,105,107–109,121} (Table 1.2). To quantify the progress of CFPS systems capable of manufacturing disulfide bonded proteins, it is instructive to track the manufacture of antibodies¹²¹. The first efforts to manufacture antibodies started with simpler antibody single chain variable fragments (scFvs) and yielded 8.3 µg/mL of protein⁹⁵. Modern CFPS systems enable the manufacture of full length heterotetrameric IgG antibodies with yields from 114 µg/mL (Chinese hamster ovary cell (CHO) CFPS, glycosylated)¹⁰⁷ to approximately 1,500 µg/mL (*E. coli* CFPS, aglycosylated)¹⁰⁵. These advances enable CFPS not only to be a research tool for studying and prototyping disulfide bonded proteins, but also to be a viable manufacturing platform^{46,122}.

1.3.5 Enabling technologies for CFPS applications

The simplicity of setting up CFPS reactions have led multiple groups to develop automated systems for the manufacture and purification of proteins. Early embodiments input template and CFPS reagents into an automated system and output up to milligram quantities of purified protein^{123–126}. A recent iteration of this idea, biologically derived medicines on demand (Bio-MOD), leverages this concept to produce good-manufacturing-practice (GMP) biologics at small scale at the point of care¹²⁷. Systems have also been

developed that leverage onboard DNA synthesis and amplification combined with CFPS to manufacture purified protein directly from a digital DNA sequence within an integrated instrument^{36,57}. These automated protein production systems point to a possible future where protein production is an entirely automated process with the only input from the user being the desired protein sequence, analogous to the advancements made in the DNA synthesis industry¹²⁸.

Beyond the expression of a single gene, there are many technologies that can be combined with CFPS to increase the throughput and complexity of the experiments performed. These technologies generally focus on methods to connect genotype (information) to phenotype (function) in a CFPS reaction, which has been done through compartmentalization, spatial organization, or direct physical linkage. When combined with CFPS, these technologies can be used to evaluate hundreds to millions of reaction conditions and hundreds to trillions of protein variants. Here, these technologies are briefly summarized to provide context for their use throughout different applications of CFPS.

Microwell plates are an extension of single tube-based reaction formats but enable fast and simple setup of hundreds to thousands of distinct conditions when combined with liquid handling robotics. Traditional tip-based liquid handling robots have been interfaced with CFPS systems largely to optimize reaction conditions and enabled the setup of hundreds of distinct reaction conditions^{129–131}. More recently, Echo acoustic liquid handlers have enabled flexible 384 well reaction setup in minutes with reaction volumes as low as 0.5 μL ¹³². Echo liquid handlers have thus far been used to optimize reaction conditions¹³³, improve experimental reproducibility^{72,134}, generate data for model parameterization^{82,133,135}, and enable high-throughput genetic part^{82,136,137} and protein^{34,84} prototyping. Custom microwell systems also highlight the potential for smaller reaction volumes and higher density microwells¹³⁸. While lower throughput than the other technologies discussed in this section, microwell plate and liquid handler-based screening has the significant advantage of requiring less specialized knowledge and being highly adaptable to the application of interest.

Spatial and compartmental organization on a surface is another common strategy for extending the capabilities of CFPS systems. Protein microarrays are systems where proteins are spatially organized and immobilized on a solid surface and several microarray formats have been developed to enable in situ production of the proteins of interest using CFPS (frequently referred to as self-assembling protein

microarrays)¹³⁹. The developed versions of self-assembling microarrays include the protein in situ array (PISA)^{140,141}, the multiple spotting technique (MST)¹⁴², the nucleic acid programmable protein array (NAPPA)^{143–145}, the DNA array to protein array DAPA¹⁴⁶, and microintaglio printing (μ IP)^{147,148}. These strategies vary in implementation, but generally rely on spatial segregation of the template DNAs in a DNA array and non-covalent capture of the synthesized protein to the chip to manufacture the protein array. This allows proteins to be produced in situ immediately prior to experimentation and without having to express and purify each individual protein, two major benefits over traditional protein microarrays. Self-assembling protein microarrays have been used to generate arrays of thousands of proteins¹³⁹ for high-throughput screening. Analogous to protein microarrays, several microfluidic systems containing distinct compartments for both CFPS-based protein expression and analysis have been developed. These microfluidic systems frequently have hundreds or thousands of chambers and can be used to measure gene expression and genetic circuits^{44,149–151}, molecular interactions^{152–155}, or enzymatic activities¹⁵⁶. A major benefit of this approach is that the information in each chamber is linked to the known genotype. Some iterations of this technology enable millions of defined compartments¹⁵⁷, although the genotype is not known until individual wells are recovered. Microfluidic chambers have also been used to study the self-assembly of materials proteins¹⁵⁸ and ribosomal subunits¹⁵⁹.

Non-surface associated compartments are also frequently used to encapsulate CFPS systems. The methods vary widely in implementation and function and include emulsion-based droplets, liposomes, polymersomes, and more¹⁶⁰. These techniques can generate millions^{161–164} of compartments via vortex mixing, microfluidics, or other methods depending on the compartment type. Compartments are exciting prospects for CFPS-based high throughput screening and directed evolution^{161–167}, though unlike the surface-based microfluidic chambers above, they typically do not provide coupled information about the identity and activity of each condition screened. However, this is not the case for all applications, droplet-based compartment combined with optical barcoding, the incorporation of unique combinations and concentrations of fluorophore barcodes, has been used to enable the optimization of large parameter spaces for genetic circuits¹⁶⁸. Compartments have also been used to study the impacts of compartmentalization on gene expression^{160,169–172} and build towards synthetic cells^{160,173,174}.

Several display technologies that utilize CFPS have also been developed³⁰. While cell-based display methods typically create the genotype to phenotype linkage through compartmentalization, CFPS based selection methods, like ribosome¹⁷⁵ and mRNA¹⁷⁶ display, frequently leverage a direct physical link between the displayed protein transcript and its polypeptide chain to maintain the genotype and phenotype linkage. These methods circumvent the traditional reformation limitation with cell-based techniques and in principle enable library sized of up to 10^{14} variants^{177,178}. Generally, these systems have been applied to the selection of proteins that bind a desired target (Section 1.3.6.3), although they have also been expanded to map protein-protein interaction networks^{179–182}, evaluate protein stability^{183–187}, and evolve certain classes of enzymes¹⁸⁸.

1.3.6 CFPS for protein-protein interaction screening and engineering

Protein-protein interactions (PPIs) are critical biological processes. The characterization and engineering of these processes are important for both the understanding of natural protein biology and for diverse applications in biotechnology and synthetic biology. Many of the properties of CFPS systems make them an ideal tool to study and engineer PPIs. They have been utilized for a variety of applications in this area, from evaluating interactomes to engineering some of the highest-affinity antibodies reported to date. These applications constitute three main categories based on the methodologies used: plate-based screening, microarrays and microfluidic systems, and in vitro display.

1.3.6.1 Individual Screening

PPIs are frequently characterized by individually expressing the proteins of interest and evaluating their interactions by a variety of measurement modalities. Here, CFPS, coupled with an appropriate measurement modality, can greatly accelerate the speed and throughput of this process. The open reaction environment allows for direct interrogation of the produced proteins without purification, which is a bottleneck in typical PPI workflows. The scalability of CFPS reactions enables proteins to be expressed at small (1-10 μL) volumes in 96-, 384-, or 1536-well microwell plates, allowing protein expression to match the throughput of the plate-based assay used to analyze the expressed proteins. Furthermore, there is a gap in throughput between high-throughput interaction screens and methods to validate the observed interactions which individual screening with CFPS can help to fill¹⁸⁹.

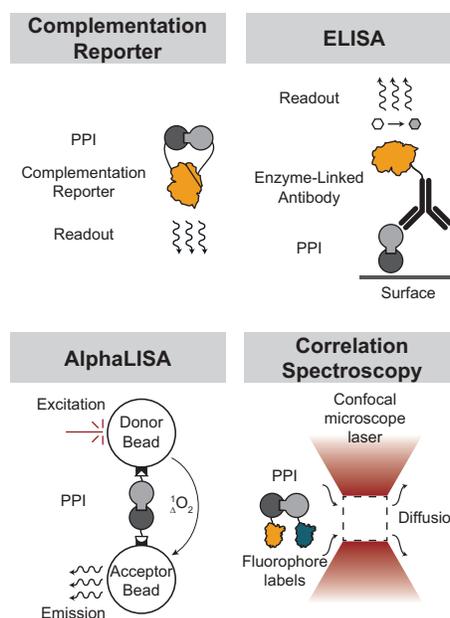


Figure 1.2. Methods for evaluating PPIs in CFPS that don't require purification of the interaction partners.

There are several standard PPI quantification methodologies that enable the quantification of PPIs in the complex CFPS environment without purification (Fig. 1.2). A key feature of these methodologies is that they must be able to specifically probe the interaction of interest in the complex CFPS environment. Protein complementation assays¹⁹⁰, where reporter proteins are fragmented and genetically fused to the proteins of interest, have been coupled to CFPS for probing natural PPIs^{191,192}, identifying inhibitors of PPIs^{193,194}, and prototyping PPI-based logic gates³³. Fluorescence correlation spectroscopy has been applied directly to CFPS systems and has been used to evaluate the interactions and oligomeric state of the expressed proteins or complexes^{189,195–199}. The technique uses confocal microscopy to observe fluorescently labeled (either by genetic fusion with a fluorescent protein¹⁹⁶ or conjugation with small molecule fluorophores¹⁹⁵) single molecules or complexes that transit through a small observation volume (~1 fL) with the intensity and wavelength of the signal corresponding to the oligomeric state of the transiting protein(s) (Fig. 1.2). Enzyme-linked immunosorbent assays (ELISAs)²⁰⁰, a gold standard method in PPI screening relying on immobilization of antigens on a surface and detection with an enzyme-linked antibody, have been utilized to screen antibodies produced in CFPS^{55,201,202}. The amplified luminescent proximity homogeneous linked immunosorbent assays (AlphaLISA)²⁰³, a high-throughput and homogeneous ELISA-

like alternative²⁰³, has also found utility when coupled to CFPS systems. AlphaLISA is particularly well suited for high-throughput microwell plate screening as it has been shown to tolerate a number of crude CFPS systems^{34,189,204} and it can be performed in 384-well plates without wash steps. AlphaLISA has been utilized to measure the interaction specificities between several natural proteins^{189,197–199} as well as to profile 204 mouse genes for autoantigenicity²⁰⁴. When coupled with state-of-the-art liquid handling robotics, AlphaLISA also enabled the rapid evaluation of hundreds of individual antibodies³⁴ and computationally designed binding proteins³⁵. An in vitro two-hybrid (IVT2H) assay, where PPIs are quantified by transcriptional activation of a reporter gene, has also been developed CFPS²⁰⁵.

1.3.6.2 Microarrays and Microfluidic Systems

There is great interest in screening of large numbers of PPIs to characterize interaction networks¹⁸², to profile GPCR binding¹⁸², to aid in antibody engineering efforts¹⁸², and more. CFPS is an attractive technology to enable the high-throughput evaluation of PPIs due to its ability to synthesize libraries of proteins and spatially pattern the resulting proteins to gain individual insight into each interaction. These benefits have been leveraged by a myriad of technologies to evaluate thousands of interactions on chips or in microfluidic devices.

Self-assembling protein microarrays and microfluidic systems are one tool used to interrogate protein-protein interactions. Self-assembling microarrays assay for PPIs by exposing the microarray to a labeled target protein, and interactions are measured by the intensity of the label output localized at each microarray spot. Self-assembling microarrays have been used to map the interaction network of 841 interactions between 29 human replication proteins¹⁴³, profile antibody targets^{206–208}, investigate protein small molecule interactions²⁰⁹, and more^{210,211}. In addition to the standard fluorescence-based readouts^{139,212}, self-assembling microarrays have also been analyzed using fluorescent single-walled carbon nanotubes¹⁵⁴ and surface plasmon resonance (SPR) sensors²¹³.

Microfluidic systems have also been developed to measure PPIs. The protein interaction network generator (PING) is a microfluidic system that combines CFPS with mechanical trapping of molecular interactions (MITOMI) to evaluate the interactions of the expressed proteins¹⁵⁵. The system was used to evaluate 1,849 interactions between 43 *Streptococcus pneumoniae* proteins in quadruplicate, mapping a rich interaction network. Impressively, PING also enabled the evaluation of membrane proteins, with recent

iterations of this technology enabled the simultaneous profiling of 2,100 membrane proteins to characterize host-pathogen interactions²¹⁴. A similar platform has also been used to study and engineer protein-DNA interactions^{153,154}. Droplet-based microfluidic systems have also been used to evolve peptide binders using the IVT2H assay²¹⁵.

Several techniques have been developed that leverage the spatial segregation of Illumina DNA sequencing flow cells and in vitro display technologies (see section 1.3.4.3) to assay PPIs^{182,216,217}. In single molecule interaction sequencing (SMIseq)¹⁸², proteins are displayed using ribosome display, immobilized in a polyacrylamide gel, and panned against a ribosome displayed library of other variants. The mRNA barcodes are then amplified, sequenced, and interactions are quantified by co-localization of the barcodes. This approach enables 'library versus library' scale experiments and was used to profile the specificity of a library of 200 scFvs against 55 different human proteins constituting 11,000 possible interactions¹⁸². In protein display on a massively parallel array (Prot-Map)²¹⁶, a library of DNA templates coding for protein variants are sequenced on an Illumina MiSeq flow cell to determine variant locations. The flow cell is then incubated with a CFPS system and ribosome display is used to display the protein variants at the known location on the flow cell. Protein are probes using fluorescently labeled target molecules and imaged on a TIRF microscope to quantitatively assay binding. The system was used to probe the sequence specificity (12,739 variants) of the M2 anti-FLAG antibody and to study the mutational landscape (156,140 variants) of the SNAP-tag self-labeling enzyme²¹⁶. These techniques are promising for generating quantitative data sets on large numbers of protein variants.

1.3.6.3 In Vitro Display Technologies

There are several methods that have been developed to leverage CFPS for directed evolution to select for protein and peptide binders against new targets that afford several benefits over their cell-based counterparts³⁰. The in vitro nature of these technologies means that library size is limited by number of ribosomes and mRNA present as opposed to the transformation efficiency, which leads to library sizes up to 10^{14} variants^{177,178}, as opposed the maximum library size of 10^9 in cells^{176,218}. Library diversification steps can also be performed without the requirement for transformation and cell growth, leading to faster selection cycles. The lack of a cell viability constraint means that single selection steps can be carried out over the course of multiple days, which is valuable for working with toxic proteins²¹⁹ or performing off rate selection-

based affinity maturation for binders with equilibrium dissociation constants in the pM range^{220,221}. Furthermore, the lack of a viability constraint also enables widespread genetic recoding for the incorporation of non-canonical amino acids²²².

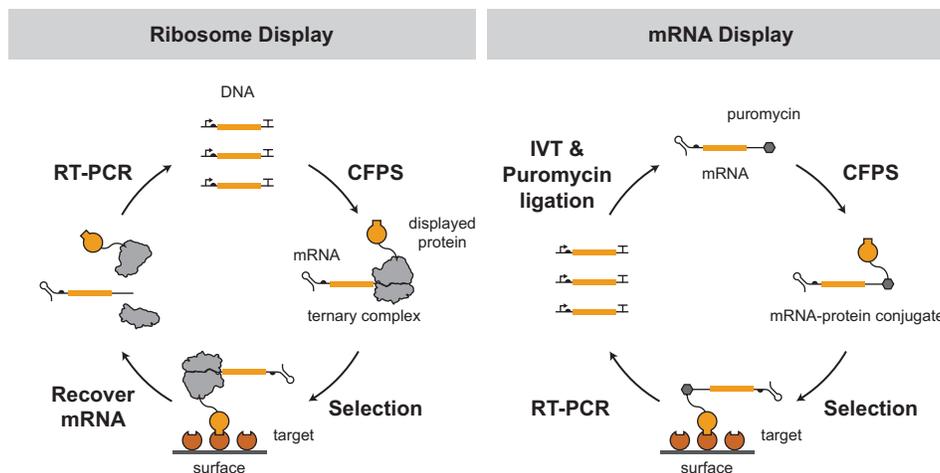


Figure 1.3. Schemes depicting ribosome and mRNA display, the two most widely used display technologies leveraging CFPS.

Ribosome display is designed to maintain the genotype to phenotype linkage non-covalently by stalling the ribosome on its transcript without releasing the polypeptide chain^{175,223–225}. The ribosome is stalled on the transcript through the omission of a stop codon, which creates a ternary complex that can be selected for binding using the displayed protein followed by recovery of the winning sequences by reverse transcription of the transcript¹⁷⁵ (Fig. 1.3). Ribosome display has been successfully applied to a variety of different protein binders, from single chain variable fragments (scFvs)^{220,221}, to antigen binding fragments (Fabs)²²⁶, designed ankyrin repeat proteins (DARPINs)²²⁷, and single domain antibodies^{228–231}. Of note, one of the strongest reported monovalent antibody affinities, an scFv targeting a bovine prion protein (KD = 1 pM), was engineered using ribosome display using extended off rate selections for affinity maturation^{220,232,233}. Ribosome display has also been used to select for improved protein stability using chemical denaturants¹⁸³ and protease treatment¹⁸⁴. A limitation of ribosome display is that selections must be carried out under conditions under which the ternary complex is stable, although ternary complex has been reported to be quite stable even under denaturing stresses²²³.

mRNA display makes the genotype to phenotype linkage by covalently conjugating the nascent peptide chain to the mRNA using an mRNA-puromycin conjugate^{176,178,188,234,235} (Fig. 1.3). Puromycin is an antibiotic that interferes with translation by binding to the ribosomal A site and accepting the nascent peptide chain¹⁷⁶. By conjugating puromycin to the mRNA transcript, a covalently linked peptide-mRNA fusion can be generated. mRNA display has been used widely for the selection of binding proteins from random linear peptides²³⁶, antibody mimics²³⁷, single domain antibodies²³⁸, and antibody antigen binding fragments²³⁹. mRNA display has also been used to probe the functional areas of random sequence space by selecting functional sequences from random libraries²⁴⁰. It also features prominently in the random non-standard peptide integrated discovery (RAPID) system^{222,241}. The RAPID system has been used to select linear and cyclic peptides containing more than ten distinct non-canonical amino acids with properties akin to natural products^{222,241}. mRNA display can also be used to perform selections from natural proteome libraries by displaying natural proteins and panning for proteins that bind to a given target^{179–181}. The covalent nature of the mRNA display genotype/phenotype linkage has also enabled the selection for protein stability using chemical denaturants^{185,186} and temperature¹⁸⁷. Major limitations of traditional mRNA display include the stability of mRNA as well as the effort required to prepare mRNA-puromycin conjugates. Several other systems have been developed to address these limitations. cDNA display²⁴² is a modified version of mRNA display developed to overcome the mRNA stability issue of mRNA display by linking puromycin directly to the cDNA. Transcription-translation coupled with association of puromycin linker (TRAP) display²⁴³ and more recently cDNA TRAP display²⁴⁴ enable one pot transcription, translation, and puromycin conjugation, which simplifies and accelerates the selection process.

Many other in vitro display techniques have also been developed that diverge from ribosome and mRNA display. CIS display leverages the ability of the RepA protein to bind the DNA template from which it was expressed, which has been used to select protein binders²⁴⁵. Similarly, covalent antibody display (CAD) leverages the P2A enzyme to covalently link the expressed protein to its DNA template, enabling selection of scFvs²⁴⁶. SNAP display²⁴⁷ is a technique that leverages in vitro compartmentalization in emulsions and a genetic SNAP tag fusion, an enzyme and small molecule covalent inhibitor pair, to link the expressed protein to its coding DNA and has been used to select a variety of binding proteins including DARPINS²⁴⁸. Several different microbead display systems have been developed^{249–251}, all of which utilize

in vitro compartmentalization in emulsions to bind both the binding protein and its coding DNA template to a microbead that can be used to perform selections. Liposome display is a technique developed to couple genotype and phenotype via compartmentalization by an emulsion to enable the engineering of membrane proteins, toxic pore-forming proteins, transporters, and receptors²¹⁹.

Chapter 2 - De Novo Design of Protein Logic Gates

2.1 Abstract

The design of modular protein logic for regulating protein function at the post-transcriptional level is a challenge for synthetic biology. Here we describe the design of 2-input AND, OR, NAND, NOR, XNOR, and NOT gates built from *de novo* designed proteins. These gates regulate the association of arbitrary protein units ranging from split enzymes to transcriptional machinery in vitro, in yeast and in primary human T cells, where they control the expression of the TIM3 gene related to T cell exhaustion. Designed binding interaction cooperativity, confirmed by native mass spectrometry, makes the gates largely insensitive to stoichiometric imbalances in the inputs, and the modularity of the approach enables ready extension to 3-input OR, AND, and disjunctive normal form gates. The modularity and cooperativity of the control elements, coupled with the ability to *de novo* design an essentially unlimited number of protein components, should enable design of sophisticated post-translational control logic over a wide range of biological functions.

The following text appeared in:

Chen, Z.; Kibler*, R.D.; Hunt*, A.C.; Busch*, F.; Pearl*, J.; *et al.* De novo design of protein logic gates. *Science* (2020). (* Indicates equal contribution)

<https://doi.org/10.1126/science.aay2790>

2.2 Introduction

Protein-protein interactions are ubiquitous in cellular decision making and controlling them will be increasingly important in synthetic biology^{252–255}. Although protein interactions are central to natural biological circuits, efforts to create new logic circuits have focused on control at the level of DNA^{256,257}, transcription^{258–269}, or RNA^{264,270–273}. Recently, protein-based circuits have been generated by rewiring native signaling pathways^{274–279}, bringing proteins together with coiled coils²⁸⁰, or creating protease cascades^{281,282}; these approaches rely on a small number of components which limits their modularity and scalability. The ability to *de novo* design protein-based logic gates modulating arbitrary protein-protein interactions could open the door to new protein-based control systems in and out of cells.

In principle, it should be possible to design a wide range of logic gates *de novo* using a set of heterodimeric molecules. For example, given hypothetical heterodimer pairs **A:A'**, **B:B'**, and **C:C'**, an AND gate modulating the association of **A** with **C'** can be constructed by genetically fusing **A'** and **B**, and **B'** and **C**: association occurs only in the presence of both **A'-B**, and **B'-C** (here and below ":" denotes noncovalent interaction, and "-" a genetic fusion through flexible linkers). Several building block properties are desirable for constructing such associative logic gates. First, there should be many mutually orthogonal heterodimeric pairs, so that gate complexity is not limited by the number of individual elements. Second, the building blocks should be modular and similar in structure so that differences in building block shape and other properties do not have to be considered when constructing the gates. Third, single building blocks should be able to bind to multiple partners with different and tunable affinities, allowing inputs to perform negation operations by disrupting pre-existing lower affinity interactions. Fourth, the interactions should be cooperative so gate activation is not sensitive to stoichiometric imbalances in the inputs. In the above AND gate, for example, if the interactions are not cooperative, a large excess of **A'-B** will pull the equilibrium towards partially assembled complexes (**A'-B** with either **A** or **B'-C** but not both), which will limit gate activation.

Here, we explored the possibility of designing logic gates satisfying all four of the above criteria using *de novo* designed protein heterodimers with hydrogen bond network-mediated specificity²⁸³. Sets of mutually orthogonal designed heterodimers (DHDs, hereafter referred to by numbers, e.g. **1** and **1'** form one cognate pair. Table 7.1.) with hydrogen bond network mediated specificity (see Fig. 2.1A inset for

example) are available for logic gate construction, satisfying condition 1 (orthogonality). The heterodimeric interfaces all share the same four helix bundle topology (Fig. 2.1A), satisfying condition 2 (modularity). The shared interaction interface allows a limited amount of cross talk between pairs, leading to a hierarchy of binding affinities, satisfying condition 3 (multiple binding specificities). Inspired by cooperative systems in nature^{284,285}, we sought to achieve condition 4 (cooperativity) by constructing the monomer fusions (**A'-B** and **B'-C** in the above example) in such a way that the interaction surfaces (with **A** and **C'**) are buried within the fusions. The free energy required to expose these buried interfaces would oppose gate activation, and we reasoned that the system could be tuned so that the sum of the binding energies of the two partners, but not either one alone, would be sufficient to overcome this barrier, ensuring cooperative gate activation. If condition 2 (modularity) holds, then a single scheme for ensuring cooperativity could in principle work for a wide range of gate configurations.

2.3 Results

2.3.1 Design and testing of an induced dimerization system

To explore the design of cooperative building blocks, we focused on the simple system **A** + **A'-B** + **B'** (we refer to this as induced dimerization below, **A** and **B'** as the monomers, and **A'-B** as the dimerizer). If binding is not cooperative, the amount of the trimeric complex decreases when **A'-B** is in stoichiometric excess relative to **A** and **B'**: the formation of intermediate dimeric species of the dimerizer binding to either of the monomers competes with formation of trimeric complexes. On the contrary, if binding is cooperative such that no binding to either monomer occurs in the absence of the other, the amount of trimeric complex formed becomes insensitive to an excess of the dimerizer. A simple thermodynamic model of the effect of binding cooperativity on the stoichiometry dependence of such induced dimerization systems (Fig. 2.1B, supplemental materials modeling section) shows that as the binding cooperativity decreases, there is a corresponding decrease in the population of full trimeric complexes at high dimerizer concentrations (Fig. 2.1C).

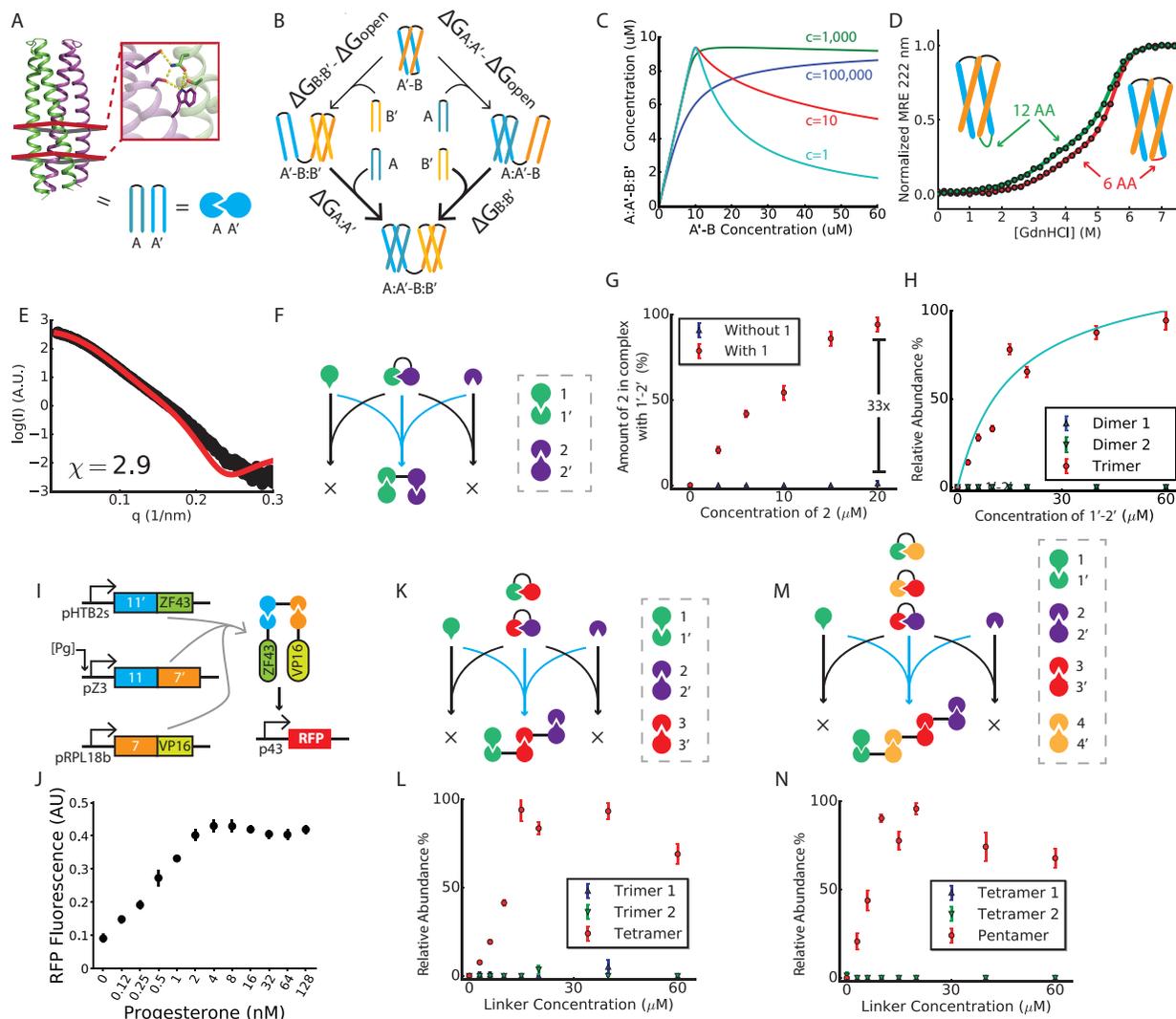


Figure 2.1. Cooperativity of CIPHR logic gates. (A) Left: Backbone structure of A:A' heterodimer building block, with its hydrogen bond network in inset. Bottom: Shorthand representations used throughout figures. (B) Thermodynamic cycle describing the induced dimerization system. (C) Simulation of the induced dimerization system under thermodynamic equilibrium. A and B' monomers were held constant at 10 μM each while titrating in various initial amounts of the A'-B dimerizer proteins. If binding is not cooperative (small c), the final amount of trimeric complexes decreases when the dimerizer protein is in excess. (D) Equilibrium denaturation experiments monitored by CD for designs with 6- and 12- amino acid (AA) linkers. Circles represent experimental data, and lines are fits to the 3-state unimolecular unfolding model. (E) Experimental SAXS profile of 1'-2' with a 6-residue linker (in black), fitted to the calculated profile of 1:1' heterodimer. (F) An induced dimerization system using a 6-residue linker. (G) Native MS titration of 2 against 1'-2' in the presence (red) or absence (blue) of 1. (H) Native MS titration of 1'-2' against 1 and 2. Dimer 1 and 2 refer to partial dimeric complexes consisting of the dimerizer binding to either of the monomers. For comparison, the thermodynamic model result with $c = 991,000$ is shown in cyan. (I) Schematic of testing of the induced dimerization system in yeast, with *in vivo* results in (J). Pg, progesterone. (K) A two-input AND gate schematic, with native MS titration results in (L). Trimer 1 and 2 refer to partial trimeric complexes of the two dimerizer proteins binding to either one of the monomers. (M) A three-input AND gate, with native MS titration results in (N). Tetramer 1 and 2 refer to partial tetrameric complexes of the three dimerizer proteins binding to either one of the monomers. All error bars are reported as standard deviations of $n=3$ independent replicates.

We hypothesized that a folded four helix bundle like state of the **A'-B** dimerizer could oppose binding to either **A** or **B'**, as the relatively hydrophobic interacting surfaces would likely be sequestered within the folded structure (Fig. 7.1A). We tested different flexible linker lengths connecting **A'** with **B** using heterodimers **1:1'** and **2:2'** as a model system. At all linker lengths tested (between 0 and 24 residues), constructs were folded and stable in circular dichroism (CD) guanidine hydrochloride (GdnHCl) denaturation experiments, with unfolding free energies greater than 13 kcal/mol (Fig. 2.1D, Fig. 7.2, Table 7.3). Although **1'-2'** dimerizer constructs with short linkers of 0 and 2 residues, or with a very long 24 residue linker could be purified as monomers (Fig. 7.1B), they were prone to aggregation, perhaps due to domain swapping. In contrast, designs with 6 and 12 residue linkers remained largely monomeric (Table 7.4). Small angle x-ray scattering (SAXS) experiments²⁸⁶ indicate their hydrodynamic radii are close to those of folded four-helix bundle DHDs (Fig. 2.1E, Table 7.2). Linkers in this length range likely allow the two monomers (**1'** and **2'**) to fold back on each other such that the largely hydrophobic interaction surfaces are buried against each other; such a structure would have to partially unfold for **1'-2'** to interact with either **1** or **2**. The magnitude of the unfolding energy (ΔG_{open} in Fig. 1B), determines the extent of cooperativity for the gate. We selected linker lengths of 6-, 10- or 12- residues for all of the following experiments.

We studied the cooperativity of the induced dimerizer system *in vitro* using native mass spectrometry (nMS,^{287,288}) which can directly measure the populations of different oligomeric species in a sample (Tables 7.5-7.8, for calibration curve see Fig. 7.3). We first measured the extent to which **1** activates the binding of **2** to **1'-2'** (Fig. 2.1F). **1**, **2** and **1'-2'** were separately expressed in *E.coli* and purified. At 10 μM each of **2** and **1'-2'**, the fraction of **2** in complex with **1'-2'** increased from 3% to 100% upon addition of 20 μM **1** (Fig. 2.1G); a fold increase comparable with naturally occurring allosteric systems²⁸⁴. To assess the sensitivity of binding to stoichiometric imbalance, 10 μM **1** and **2** were titrated with increasing concentrations of **1'-2'** (Fig. 2.1H), and the species formed determined by nMS. The heterotrimeric **1:1'-2':2** complex was observed over a wide range of **1'-2'** concentrations (Fig. 2.1H). Even in the presence of a 6 fold excess of **1'-2'**, there was no decrease in the amount of **1:1'-2':2** formed, and neither **1:1'-2'** nor **1'-2':2** were detected (Fig. 2.1H). We define a cooperativity parameter *c* as the ratio of the affinities in the presence and absence of the other monomer, which in our model directly relates to the free energy of opening of the dimerizer ($c = e^{\Delta G_{\text{open}}/RT}$, see supplementary materials). The estimated *c* value from fits of

the thermodynamic model to nMS data (Fig. 2.1H, cyan line) is $991,000 \pm 21$ (for reference, the c value of the naturally occurring N-Wasp system is 350, but system differences complicate quantitative comparisons). This value of c corresponds to ΔG_{open} of 8.2 kcal/mol, which is about half the measured unfolding free energy of **1'-2'** (Table 7.3), suggesting that binding may not require complete unfolding of the four helix bundle state of the dimerizer.

To investigate the cooperativity of the induced dimerizer system in living cells, we used a two-hybrid like assay in yeast. **11'** was fused to the DNA binding domain ZF43²⁶⁵, **7** to the transactivation domain VP16, and the dimerizer **11-7'** was placed under the control of a progesterone responsive element. Association of the DNA binding and activation domains results in transcription of red fluorescent protein (RFP) (Fig. 2.1I). Treating cells with increasing amount of progesterone resulted in up to a 4.5-fold increase in RFP signal, with only a small drop at saturating progesterone concentrations (Fig. 2.1J). Based on calibration curves, under these conditions **11-7'** is expected to be in greater than 5-fold molar excess over **11'** and **7** (Fig. 7.4), suggesting that **11-7'** binds cooperatively to **11'** and **7** in cells. Thus the cooperativity of the dimerizer system makes it robust to fluctuating component stoichiometries in cells.

With dimerizers displaying cooperative binding, we reasoned that the lack of dependence on stoichiometric excesses of one of the components should extend to more complex gates. Using nMS, we investigated the cooperativity of a 2-input AND gate constructed with the two dimerizers **1'-3'** and **3-2'** as inputs, and monomers **1** and **2** brought together by the two inputs (Fig. 2.1K). As the concentration of the 2 inputs increased, the amount of heterotetrameric complex plateaued at a stoichiometry of 2:1, and then largely remained constant with a small drop at molar ratio of 6:1. Only very small amounts of partial complexes (heterotrimers and heterodimers) were observed, further indicating high cooperativity (Fig. 2.1L). We constructed a 3-input AND gate from **1'-4'**, **4-3'**, and **3-2'**, which together control the association of **1** and **2** (Fig. 2.1M). Similar to the 2-input AND gate, the amount of full, pentameric complexes only decreased slightly at greater than stoichiometric concentrations of inputs with no detectable competing tetrameric complexes (Fig. 2.1N).

2.3.2 Design and testing of multi-input logic gates

We explored the modular combination of DHDs (Table 7.1) to generate a range of 2-input Cooperatively Inducible Protein HeterodimeR (CIPHR) logic gates. Monomers from individual DHDs were

linked to effector proteins of interest via genetic fusion such that the inputs (linked heterodimer subunits) control colocalization or dissociation of the effector proteins. Taking advantage of previously measured all-by-all specificity matrices for the DHDs²⁸³, we explored constructing gates from two interaction modalities: cognate binding between designed protein pairs, or competitive binding involving multispecific interactions (Fig. 2.2A).

We began by constructing AND and OR gates, reading out gate function using a yeast-two-hybrid (Y2H) setup similar to previously described yeast-four-hybrid systems^{289,290}. To construct an AND gate, we fused **2** to the Gal4 activation domain (AD), and **1** to the Gal4 DNA binding domain (DBD). In this scheme, the colocalization of AD and DBD, and resulting transcriptional activation of the *His3* gene, should require the expression of both input proteins (**1'-5**, **5'-2'**). Indeed, growth in media lacking histidine required expression of both inputs (Fig. 2.2B). An OR gate was similarly constructed by linking the **1-6** fusion to the AD and **7'** to the DBD. Expression of either of the inputs **1'-7** or **6'-7** results in growth by driving association of AD with DBD (Fig. 2.2C).

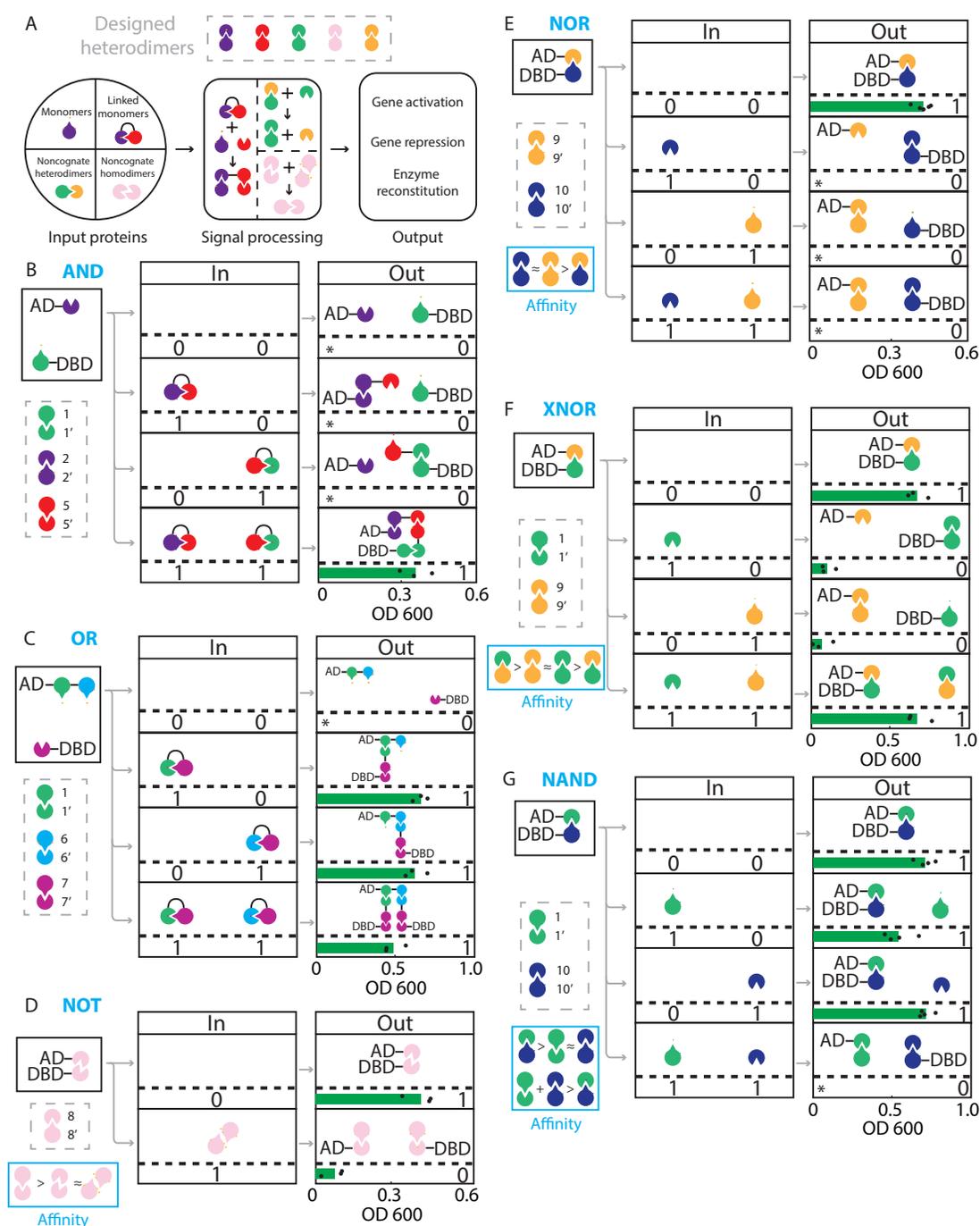


Figure 2.2. CIPHR two input logic gates. (A) CIPHR gates are built from DHDs (top) with monomers or covalently connected monomers as inputs (left); some gates utilize only the designed cognate interactions (left side of middle panel), while others take advantage of observed binding affinity hierarchies (right side of middle panel). (B-C) Two-input AND (B) and OR (C) CIPHR logic gates based on orthogonal DHD interactions. (D-G) NOT (D), NOR (E), XNOR (F), and NAND (G) CIPHR logic gates made from multispecific and competitive protein binding. For each gate, black dots represent individual Y2H growth measurement corrected over background growth, with their average values shown in green bars. * indicates no yeast growth over background. 0s and 1s in the middle and right blocks represent different input states and expected outputs, respectively.

We explored the construction of additional boolean logic gates by exploiting binding affinity hierarchies identified in all by all Y2H experiments²⁸³. **8** alone forms a homodimer, but in the presence of **8'** it dissociates to form the **8:8'** heterodimer (Fig. 7.5A). We constructed a NOT gate by fusing **8** to both AD or DBD; the **8:8** homodimer supports yeast growth but in the presence of co-expressed **8'** input protein, the interaction is broken and growth is slowed (Fig. 2.2D). Based on the affinity hierarchy $9:9' \approx 10:10' > 9:10'$ (Fig. 7.5B), we constructed a NOR gate in which **9** was fused to the AD and **10'** to the DBD, with **9'** and **10** the two inputs. Either or both of the inputs outcompete the **9:10'** interaction and hinder yeast growth (Fig. 2.2E). Based on the affinity hierarchy $9':1' > 9:9' \approx 1:1' > 9:1$ (Fig. 7.5B), an XNOR gate was constructed by fusing **9** to AD, **1** to DBD, and using **9'** and **1'** as the two inputs: the presence of either outcompetes the **9:1** binding and blocks growth, but when both are expressed they instead interact with each other and growth is observed (Fig. 2.2F). Similarly, a NAND gate was designed based on the interaction hierarchy $1':10' > 1:1' \approx 10:10' > 1:10$ (Fig. 7.5B). Neither **1** nor **10** alone can outcompete the **1':10'** interaction and hence growth occurs, but when both are expressed, the free energy of formation of both **1:1'** and **10:10'** outweighs that of **1':10'** and growth is blocked (Fig. 2.2G).

We next investigated 3-input CIPHR logic gates. We first used native MS to characterize a 3-input AND gate (Fig. 2.1M) in which monomers **1** and **2** are brought into proximity by the three inputs **1'-4'**, **4-3'**, and **3-2'**. We experimentally tested all eight possible input combinations (Fig. 2.3A) with both **1** and **2** present, quantifying all complexes using nMS. Consistent with 3-input AND gate function, **1** and **2** only showed significant co-assembly when all three inputs are present (Fig. 2.3B).

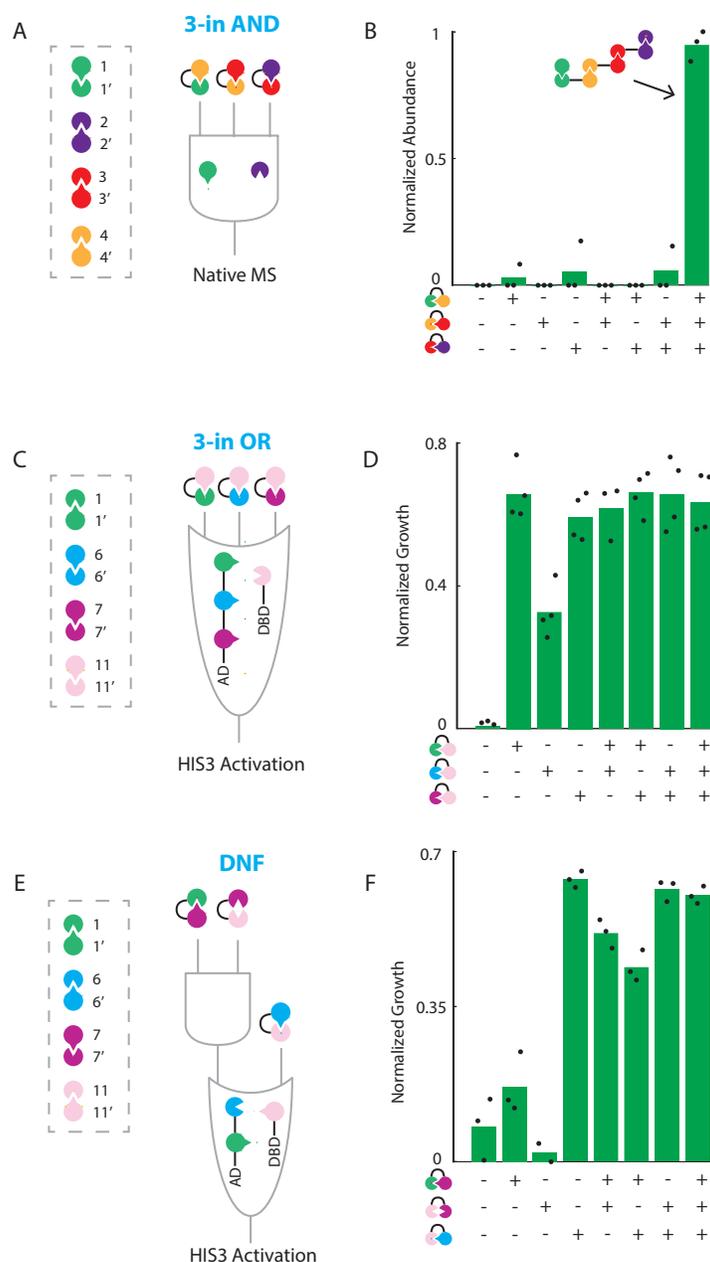


Figure 2.3. Three-input CIPHR logic gates. (A) Schematic of a three-input AND gate. (B) Native MS results indicate proper activation of the 3-input AND gate only in the presence of all three inputs. (C) Schematic of a three-input OR gate. (D) Y2H results confirmed activation of the 3-input OR gate with either of the inputs. (E) Schematic of a DNF gate. (F) Y2H results confirmed proper activation of the gate. For each gate, black dots represent individual measurements with their average values shown in green bars. For Y2H-based measurements (D,F), the growth measurements are corrected over background growth.

To test 3-input CIPHR gate function in cells, we designed two additional gates using the same 4 pairs of DHDs and tested them by Y2H. To make a 3-input OR gate, **1'-6-7** was fused to AD, and **11'** to

DBD. Any one of the 3 inputs (**11-1**, **11-6'**, **11-7'**) connects the AD to the DBD through 1', 6 or 7 respectively (Fig. 2.3C). Y2H results confirmed the expected behavior of this logic gate in cells: any of the input proteins induced cell growth (Fig. 2.3D). We constructed a CIPHR disjunctive normal form (DNF, [**A AND B**] OR **C**) gate by fusing **1'-6** to AD, **11'** to DBD with inputs **11-7'**, **7-1**, or **11-6'** (Fig. 2.3E). In Y2H experiments, the DNF gate functioned as designed, with low yeast growth levels when no input or only one of the **11-7'** and **7-1** input proteins are present, and high yeast growth levels otherwise (Fig. 2.3F).

2.3.3 Evaluation of CIPHR logic gates in different environments

To test the transferability of CIPHR logic gates, we explored the ability of CIPHR logic gates to reconstitute split enzyme activity by controlling the association of the two halves of the NanoBiT split luciferase system^{191,291,292}. Monomers from **1:1'**, **2:2'**, **4:4'** and **9:9'** (Fig. 2.4A) were fused in pairs to the two split domains (smBiT and IgBiT), and produced by in vitro transcription and translation, which facilitated a rapid testing cycle enabling the full 4x4 interaction affinity hierarchy to be determined by monitoring luciferase activity following mixing (Fig. 7.6A). Based on this hierarchy, we constructed and experimentally verified an induced dimerization circuit with **4-smBiT**, **1-IgBiT**, and **1'-4'** as the input (Fig. 2.4B, Fig. 7.6C-D); characterization of the time dependence of the response revealed a 7-fold increase in signal 5 minutes after adding inputs (Fig. 7.6D). We also constructed an AND gate with **4-smBiT**, **1-IgBiT**, and **1'-2** and **2'-4'** as the inputs (Fig. 2.4C) and a NOR gate with **1'-smBiT**, **2'-IgBiT**, and **1** and **2** as the inputs (Fig. 2.4D), both of which had the designed dependence of gate function (*i.e.*, luciferase activity) on the inputs. We investigated the response of the NOR gate to varying concentrations of the inputs against the NanoBiT components held at 5nM, and found a sharp drop in signal around 5 nM for both inputs consistent with NOR logic (Fig. 2.4E, Fig. 7.6E).

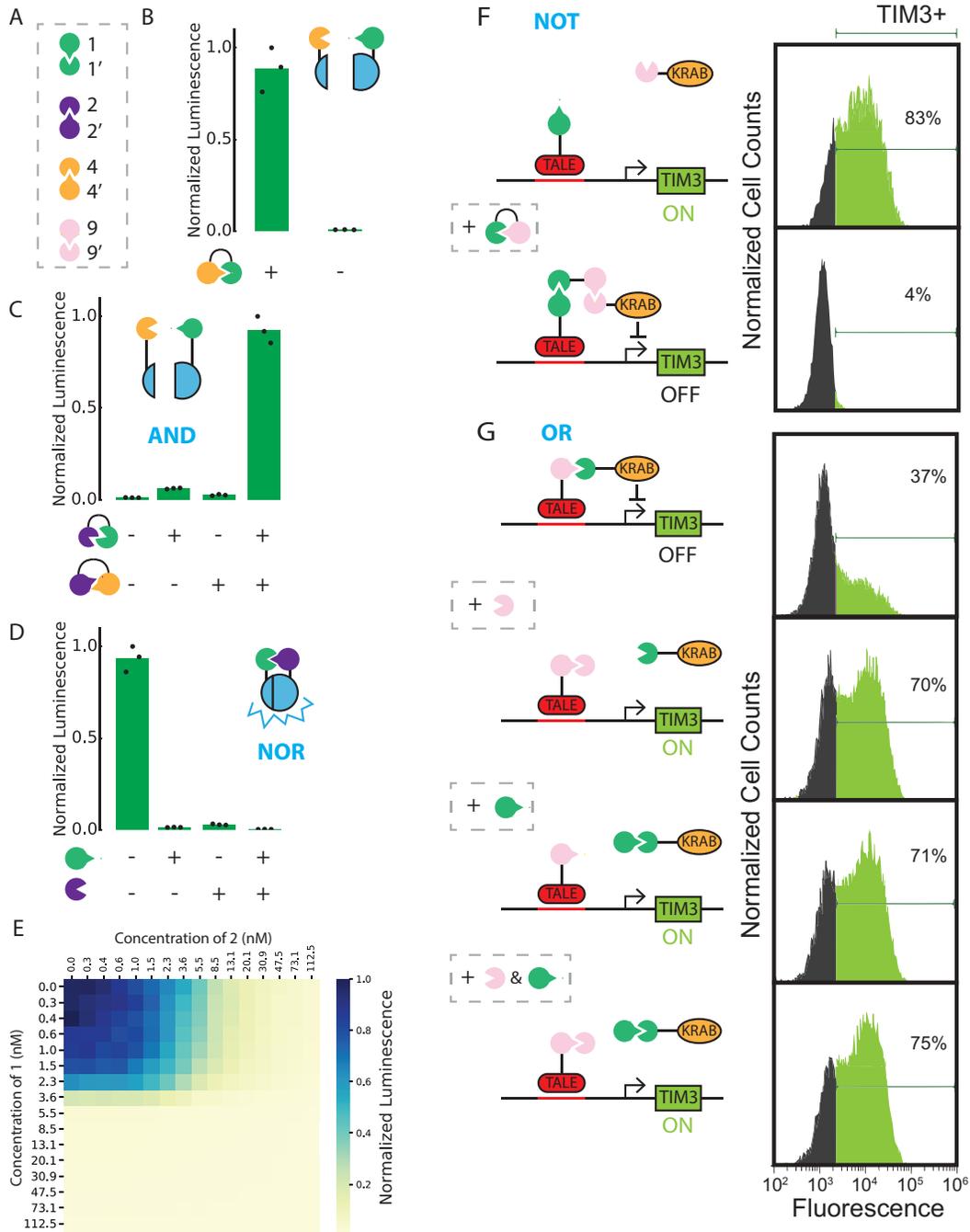


Figure 2.4. Transferability of CIPHR logic gates. (A) Four pairs of DHDs were modularly combined to construct CIPHR logic gates that can be used to control different functions: (B-E) catalytic activity of split luciferase, and (F-G) gene expression in primary human T cells. (B) Induced dimerization system, (C) AND gate, and (D) NOR gate coupled to NanoBiT split luciferase system, tested by *in vitro* translation and monitoring luminescence. (E) *In vitro* titration of the two inputs of the NOR gate in D while keeping 1'-smBiT and 2'-IgBiT fixed at 5 nM. (F) NOT gate and (G) OR gate using a split TALE-KRAB repression system to control expression of TIM3 proteins in primary human T cells, tested by flow cytometry.

Engineered T cell therapies are promising therapeutic modalities^{293–295}, but their efficacy for treating solid tumors is limited at least in part by T cell exhaustion^{296,297}. Immune checkpoint genes including *TIM3* are believed to play critical roles in modulating T cell exhaustion^{298–300}. To put the transcription of such proteins under the control of the CIPHR logic gates, we took advantage of potent and selective transcriptional repressors of immune checkpoint genes in primary T cells that combine sequence-specific transcription activator-like effector (TALE) DNA binding domains with the Krüppel-associated box (KRAB) repressor domain³⁰¹. Repression activity is preserved in split systems pairing a DNA recognition domain fused to a DHD monomer with a repressor domain fused to the complementary DHD monomer³⁰¹. We reasoned that this system could be exploited to engineer programmable therapeutic devices by making the joining of the DNA recognition and transcriptional repression functionalities dependent on CIPHR gates. Use of a repressive domain effectively reverses the logic of CIPHR gates when expression level of the target gene is measured as the output.

To test the feasibility of this concept, we used a TALE-KRAB fusion engineered to repress the immune checkpoint gene *TIM3*³⁰¹. We designed a NOT gate, with **1** fused to the TALE DNA recognition domain, **9'** fused to KRAB, and the **1'-9** dimerizer protein as the input (see Fig. 7.7A for T cell DHD specificity matrix). In this scheme, **1'-9** brings KRAB to the promoter region bound by the TALE, therefore triggering repression of *TIM3* (Fig. 2.4F). Taking advantage of the interaction between **9** and **1'**, we built an OR gate with **9**-TALE and **1'**-KRAB fusions; *TIM3* is repressed in the absence of inputs, but upon addition of either **9'** or **1**, the weaker **9:1'** interaction is outcompeted in favor of the stronger **9:9'** and **1:1'** interactions, restoring *TIM3* expression (Fig. 2.4G). These results suggest that the combination of CIPHR and TALE-KRAB systems could be directly applied to add signal processing capabilities to adoptive T cell therapy.

2.4 Discussion

The systematic design of logic gates described in this paper takes advantage of the strengths of *de novo* protein design. Since the building block heterodimers are designed *de novo*, many more components can be generated for gate construction with nearly identical overall topology than are available by repurposing biological motifs. The encoding of specificity using designed hydrogen bond networks enables a wide range of binding affinities between monomers with similar structures, which in turn allows

the construction of more complex gates based on competitive binding. From the protein biophysics perspective, our results highlight the strong synergy between *de novo* design of protein complexes and native MS, and more generally, the ability of *de novo* protein design to generate complex cooperative assemblies. For example, detecting and quantifying the 33-fold trans-activation of binding in Fig. 2.1G depended critically on the ability to resolve all species formed in solution by native MS. Analysis of the three input logic gates in Fig. 2.3B required distinguishing the designed heteropentameric assemblies — composed from five distinct protein chains — from the very large number of alternative possible heterotetrameric, trimeric, and dimeric complexes. The ability to generate highly cooperative and well-defined assemblies composed of five distinct polypeptide chains demonstrates that *de novo* protein design is starting to approach the complexity of naturally occurring protein assemblies, which are responsible for much of biological function.

Unlike nucleic acid based logic gates, CIPHR gates can be directly coupled to arbitrary protein actuation domains, offering greater diversity in the types of functional outputs. We illustrate here the coupling to transcriptional activation and repression, and split enzyme reconstitution; in principle any function that can be modulated by protein-protein association can be put under the control of the CIPHR gates. Since the designed components are hyperstable proteins and no additional cellular machinery is required, the gates should function in a wide range of conditions inside and out of cells (here we demonstrate function with purified components, in cell free extracts, yeast cells, and T-cells). The small size of DHDs and hence their genetic payload make them attractive for mammalian cell engineering. The sophistication of the circuits could be further increased by proteolytic activation as in elegant recent protease based protein circuits^{281,282}; our purely protein interaction based circuits have advantages in bioorthogonality, demonstrated scalability to three inputs, composability (the output, like the input and the computing machinery, are interactions between building blocks with common design features), and extensibility as an essentially unlimited repertoire of heterodimeric building blocks can be created using *de novo* design.

2.5 Materials and Methods

2.5.1 Buffer and media recipe for protein expression and yeast-two-hybrid experiments

TBM-5052

1.2% [wt/vol] tryptone, 2.4% [wt/vol] yeast extract, 0.5% [wt/vol] glycerol, 0.05% [wt/vol] D-glucose, 0.2% [wt/vol] D-lactose, 25 mM Na₂HPO₄, 25 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, 10 μM FeCl₃, 4 μM CaCl₂, 2 μM MnCl₂, 2 μM ZnSO₄, 400 nM CoCl₂, 400 nM NiCl₂, 400 nM CuCl₂, 400 nM Na₂MoO₄, 400 nM Na₂SeO₃, 400 nM H₃BO₃

Lysis buffer

20 mM Tris, 300 mM NaCl, 20 mM Imidazole, pH 8.0 at room temperature

Wash buffer

20 mM Tris, 300mM NaCl, 30 mM Imidazole, pH 8.0 at room temperature

Elution buffer

20 mM Tris, 300 mM NaCl, 250 mM Imidazole, pH 8.0 at room temperature

TBS buffer

20 mM Tris pH 8.0, 100 mM NaCl

YPAD buffer

Peptone 20 g/L, yeast extract 10 g/L, Adenine hemisulfate 10 μg/L, dextrose (20 g/L)

C-Trp-Ura-Leu-His+Adenine hemisulfate+Glucose

Yeast nitrogen base w/o amino acids (6.7 g/L), synthetic DO media (-Leu/-His/-Trp/-Ura) (1.4 g/L), dextrose (20 g/L), adenine hemisulfate (10 μg/L)

2.5.2 Construction of synthetic genes

For the expression of proteins in *E.coli*, synthetic genes were ordered from Genscript Inc. (Piscataway, N.J., USA) and delivered in pET21-NESG *E. coli* expression vector, inserted between the NdeI and XhoI sites. For each expression construct, a hexahistidine tag followed by a tobacco etch virus (TEV) protease cleavage site (GSSH₆HSSGENLYFQGS) were added in frame at the N-terminus of the protein. A stop codon was introduced at the 3' end of the protein coding sequence to prevent expression of the C-terminal hexahistidine tag in the vector.

Genes for yeast-two-hybrid (Y2H) studies were cloned into plasmids bearing the GAL4 DNA-binding domain (poDBD) and the GAL4 transcription activation domain (poAD)³⁰². Input proteins were cloned into plasmids V510 (uracil auxotrophic selection marker) and MX1 (bleomycin selection marker). Genes were expressed under the control of ADH1 promoters.

Genes for NanoBiT assay were ordered from Genscript Inc. (Piscataway, N.J., USA) and were synthesized and cloned into one or more of the following vectors: pJL1, pJL1-IgBiT, pJL1-smBiT, pJL1-IgBiT-sfGFP, pJL1-smBiT-sfGFP, or pJL1-sfGFP. NanoBiT constructs were designed with a c-terminal fusion of either the smBiT or IgBiT fusion partner (114 and 11s respectively²⁹¹, for sequences see Table 7.1) with a 15 amino acid linker (GGSGGGGSGSSSSGG) separating the inserted gene and the fusion partner. Where specified, genes were also fused to the c-terminus of sfGFP separated by a 30 amino acid linker (GGSGGSx5). Genes were expressed using a T7 promoter in cell-free protein synthesis.

The fluorescence-reporting yeast two-hybrid-like assay was constructed using hierarchical golden gate assembly³⁰³. Split transcription factor modules (²⁶⁵, ZF43_8 and p43_8(8x)) and DHDs were domesticated as parts using PCR amplification or gBlocks (IDT). The progesterone responsive transcription factor (Z3PM) and cognate promoter (pZ3) were domesticated using PCR (³⁰⁴, e.g. Addgene #87944 and #89195, respectively). All other promoters and terminators were taken from (3). Transcriptional units were assembled into cassette vectors and multi-transcriptional units were assembled into integration vectors for insertion into the yeast genome at the indicated loci (see “Progesterone-responsive yeast strain construction and growth media”).

2.5.3 Protein expression

Plasmids were transformed into chemically competent *E. coli* expression strain Lemo21(DE3) (New England Biolabs) for protein expression. Following transformation and overnight growth, single colonies were picked from agar plates into 5 ml Luria-Bertani (LB) medium containing 100 µg/mL carbenicillin (for pET21-NESG vectors) with shaking at 225 rpm for 18 hours at 37°C. Proteins were expressed using the autoinduction method³⁰⁵: starter cultures were further diluted into 500 ml TBM-5052 containing 100 µg/mL carbenicillin, and incubated with shaking at 225 rpm for 24 hours at 37°C.

2.5.4 Affinity purification

E. coli cells were harvested by centrifugation at 5000 rcf for 15 minutes at 4°C and the pellet resuspended in 18 ml lysis buffer. EDTA-free cocktail protease inhibitor (Roche), lysozyme, and DNase were added to the resuspended cell pellet, followed by cell lysis via sonication at 70% power for 5 minutes. Lysates were clarified by centrifugation at 4°C and 18,000 rpm for 45 minutes and applied to columns containing Ni-NTA (Qiagen) resin pre-equilibrated with lysis buffer. The column was washed two times with 5 column volumes (CV) of wash buffer, followed by 5 CV of elution buffer for protein elution.

2.5.5 Size-exclusion chromatography (SEC)

Eluted proteins were buffer exchanged into lysis buffer. N-terminal hexahistidine tags were removed with TEV protease cleavage overnight at room temperature, at a ratio of 1 mg TEV for 100 mg of protein. After TEV cleavage, sample was passed over a fresh Ni-NTA column and washed with 1.5 CV of lysis buffer, collecting flow through. The resulting proteins were purified by SEC using a Superdex 75 10/300 increase column (GE Healthcare) in TBS buffer.

2.5.6 Circular dichroism (CD) measurements

Circular dichroism (CD) wavelength scans (260 - 195 nm) and temperature melts (25 - 95 °C) were performed using an AVIV model 420 CD spectrometer, with protein samples diluted to 0.25 mg/ml in PBS pH 7.4 in a 0.1-cm cuvette. Temperature melts were carried out at a heating rate of 4 °C/min and monitored by the change in ellipticity at 222 nm.

GdmCl titrations were performed on a JASCO model J-1500 with automated titration apparatus in PBS pH 7.4 at 25 °C, with protein concentrations between 0.08 mg/ml to 0.025 mg/ml in a 1-cm cuvette with stir bar. Each titration consisted of at least 34 evenly distributed GdmCl concentration points up to 7.4 M with 30 seconds of mixing time for each step. Titrant solution consisted of the same concentration of protein in PBS and GdmCl.

2.5.7 CD data analysis and model fitting

Folding free energies were obtained by fitting equilibrium denaturation data. Fused hairpin constructs had biphasic unfolding transitions, indicating the existence of an intermediate on their respective energy landscapes. Since native MS showed that Linker 0, Linker 2, Linker 6, and Linker 12 were almost exclusively monomeric in buffer (Fig. 7.3), it was concluded that these intermediates were partially folded

monomeric species. Thus, the chemical denaturation data of these proteins was fitted to a unimolecular 3-state model:



where N represents the fully folded state, I a partially folded intermediate, and D the denatured state. The fraction of each species can be written as a function of $K_1 = [I]/[N]$ and $K_2 = [D]/[I]$, the equilibrium constants for the first and second transitions respectively:

$$f_N = (1 + K_1 + K_1 \cdot K_2)^{-1}$$

$$f_I = (1 + K_2 + \frac{1}{K_1})^{-1}$$

$$f_D = (1 + \frac{1}{K_2} + \frac{1}{K_1 \cdot K_2})^{-1}$$

In the context of equilibrium chemical denaturation experiments, the free energy of unfolding is a linear function of denaturant concentration:

$$\Delta G_{[den]} = \Delta G_{buffer} - m \cdot [den]$$

where $\Delta G_{[den]}$ represents the free energy of the system at a given concentration of denaturant, ΔG_{buffer} is the corresponding free energy change in the absence of denaturant, and m is a constant of proportionality that relates to the change in solvent-accessible surface area upon unfolding ($\Delta SASA$). Thus, the effect of denaturants on the equilibrium constant relating to each transition can be written as a function of its free energy difference in buffer, and a specific m -value:

$$K_1 = \exp\left(\frac{m_1 \cdot [den] - \Delta G_1}{R \cdot T}\right)$$

$$K_2 = \exp\left(\frac{m_2 \cdot [den] - \Delta G_2}{R \cdot T}\right)$$

By combining these expressions with the definitions for f_N, f_I, f_D , the fractional distribution of each species can be expressed as a function of denaturant concentration, and the free energy change corresponding to each transition (in buffer). Finally, for an ensemble spectroscopic technique such as CD, the observed signal (the dependent variable) as a function of denaturant concentration (the independent variable) can be expressed as a linear combination of the spectroscopic signals corresponding to each species, weighed by their fractional contribution to the ensemble:

$$MRE_{222nm} = f_N \cdot MRE_N + f_I \cdot MRE_I + f_D \cdot MRE_D$$

Where MRE_N , MRE_I , MRE_D represent the spectroscopic signatures (baselines) for the native, intermediate, and denatured states respectively. This equation was used to fit chemical denaturation data for the different linker proteins, and the fitted parameters are reported in Table 7.3. For Linker 24 in buffer, native MS revealed a significant proportion of dimer (Table 7.4). Therefore, this model is not entirely appropriate for describing the unfolding, and the fitted values for this construct should be interpreted with care. Nevertheless, denaturation performed at different concentrations of protein revealed that the position of the second transition was concentration-independent, and thus unimolecular. For this event, the model holds.

The total m -values for these linked hairpins were found to be around $3 \text{ kcal mol}^{-1} \text{ M}^{-1}$. It has been shown that m -values correlate with $\Delta SASA$ of unfolding³⁰⁶. For the folded state, $SASA$ was estimated from the structures of DHDs²⁸³ using PyMOL to be 8800 \AA^2 . For the unfolded state, $SASA$ was estimated using ProtSA^{307,308}, and is about $20,000 \text{ \AA}^2$. Thus, $\Delta SASA$ for the unimolecular unfolding of a fused hairpin should be around $11,000 \text{ \AA}^2$, which would have a predicted m -value of 3.3. This number is in close agreement with the fitted parameters reported here, in line with the notion that the folded state for these linker proteins has a four helix bundle topology.

2.5.8 Small Angle X-ray Scattering (SAXS)

Protein samples were purified by SEC in 25 mM Tris pH 8.0, 150 mM NaCl and 2% glycerol; elution fractions preceding the void volume of the column were used as blanks for buffer subtraction. Scattering measurements were performed at the SIBYLS 12.3.1 beamline at the Advanced Light Source. The sample-to-detector distance was 1.5 m, and the X-ray wavelength (λ) was 1.27 Å, corresponding to a scattering vector q ($q = 4\pi \sin \theta/\lambda$, where 2θ is the scattering angle) range of 0.01 to 0.3 \AA^{-1} . A series of exposures were taken of each well, in equal sub-second time slices: 0.3-s exposures for 10 s resulting in 32 frames per sample. For each sample, data were collected for two different concentrations to test for concentration-dependent effects; 'low' concentration samples ranged at 2.5 mg/ml and 'high' concentration samples at 5 mg/ml. Data were processed using the SAXS FrameSlice online serve and analysed using the ScÅtter software package^{286,309}. The FoXS online server^{310,311} was used to compare experimental scattering profiles to design models and calculate quality of fit (χ) values.

2.5.9 Yeast two-hybrid assay for logic gates

Chemically competent cells of yeast strain PJ69-4a (MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4(deleted) gal80(deleted) LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ) were transformed with the appropriate pair of plasmids containing DNA binding domains (DBD) or activation domains (AD), using the LiAc/SS carrier DNA/PEG method³¹². For two input CIPHR logic gates, genes encoding the input proteins (together with selection markers) were genetically integrated into either or both of the Ura3 locus (uracil auxotrophic selection marker) or the YCR043 locus (bleomycin selection marker). In the case of three input CIPHR logic gates, genes encoding two input proteins were genetically integrated as described, with the additional input cloned downstream of either the AD or DBD plasmid, separated by a p2a and nuclear localization sequence (GSGATNFSLLKQAGDVEENPGPGDKAELIPEPPKPKKRKVELGTA). The p2a sequence ensures translational cleavage to make the additional input protein a separate protein. The selection of transformed yeast cells was performed in synthetic dropout (SDO) medium lacking tryptophan and leucine for 48 h with shaking at 1,000 r.p.m. at 30 °C. The resulting culture was diluted 1:100 and grown for 16 h in fresh SDO medium lacking tryptophan and leucine, before being diluted 1:100 in fresh SDO medium lacking tryptophan, leucine and histidine. The culture was incubated with shaking at 1,000 r.p.m. at 30 °C. As it is necessary to bring the DBD and the transcription activation domain into proximity for the growth of yeast cells in medium lacking histidine, successful activation of logic gates was indicated by the growth of yeast cells^{313,314}. The optical density of yeast cells was recorded at 24 h, 48 h, and 72 h.

2.5.10 Native MS of individual proteins

Sample purity and integrity were determined by online buffer exchange MS using an UltiMate 3000 HPLC coupled to an Exactive Plus EMR Orbitrap instrument (Thermo Fisher Scientific) that was modified to incorporate a quadrupole mass filter and allow for surface-induced dissociation³¹⁵. Between 10 and 100 pmol protein was injected and online buffer exchanged to 200 mM ammonium acetate (AmAc) by a self-packed buffer exchange column (P6 polyacrylamide gel, BioRad) at a flow-rate of 100 μ L per min. Data were deconvoluted with UniDec (version 3.1.1)³¹⁶ and are shown in Table 7.4.

2.5.11 Native MS titration assay and data analysis

Titration assay: Protein concentrations were determined based on UV absorbance at 280 nm with a NanodropTM 2000c spectrophotometer (Thermo Fisher Scientific). Mixtures for titration were prepared as

outlined in Table 7.5. Other mixtures are mixed with a final monomer concentration at 10 μM for each component. Proteins were incubated at 75 $^{\circ}\text{C}$ for 30 min in the presence of 5 M guanidinium hydrochloride to ensure complete denaturation. The denatured proteins were subsequently refolded into 200 mM AmAc by dialysis using Pierce™ 96-well Microdialysis Plate, 3.5K MWCO (ThermoFisher Scientific). 50 μL of each sample was loaded into individual microdialysis devices with 1.6 mL 200 mM AmAc in each well. The buffer was changed 8 times over a total dialysis time of 36 hours at room temperature. 12.5, 25, and 50 fold dilutions in 200 mM AmAc were prepared. Dilutions were measured in triplicate by flow-injecting 5 μL into an Exactive Plus EMR Orbitrap instrument (Thermo Fisher Scientific) with 200 mM AmAc as flowing carrier stream. Mass spectra were recorded for 1,000 – 8,000 m/z at 17,500 resolution as defined at 200 m/z . The injection time was set to 200 ms for 1-in AND and 3-in AND titrations, and 150 ms for 2-in AND titrations. Voltages applied to the transfer optics were optimized to allow ion transmission while minimizing unintentional ion activation. Complex formations were observed and are shown in Table 7.6. Dilution series of individual proteins were also measured to ensure for a linear correlation between signal intensities and protein concentration (Table 7.7).

Data analysis: All mass spectra were deconvoluted and processed using Intact Mass software (Protein Metrics, Cupertino, CA). Software default deconvolution parameters were used with the exception of minimum charge state (3) and mass range (5,000-80,000 Da). A list of theoretical masses for complexes up to pentamers was generated and included in the Intact Mass software for species assignments. Mass areas (in the mass “zero-charge” domain) were made by setting the mass area integration width to 12 Da. Mass areas of all species were exported from Intact Mass software using the reports function and then combined in Excel to generate titration curves.

2.5.12 Cell-Free extract preparation

E. coli extracts were prepared following an established protocol in literature ²⁹², with minor modifications. Briefly, three 150 mL BL21 Star™ (DE3) starter cultures were inoculated in LB from a glycerol stock and cultured in a 250 mL baffled flask at 37 $^{\circ}\text{C}$ for 16 hours. The starter cultures were used to inoculate 10 L of 2xYTPG media (16 g/L tryptone, 10 g/L yeast extract, 5 g/L sodium chloride, 7 g/L potassium phosphate dibasic, 3 g/L potassium phosphate monobasic, 18 g/L glucose) in a Sartorius BIostat® Cplus fermenter at an initial OD600 of 0.08. The 2xYTP was prepared lacking glucose in 75% of the final

volume and sterilized using the fermenter sterilization cycle. A 4x glucose solution was prepared and autoclaved separately, then added to the fermenter. Cells were cultured at 37 °C with an air flow rate of 8 SLPM and a 600 RPM stir rate.

Cultures were grown until OD600 0.4-0.6, at which point the expression of T7 RNA polymerase was induced by the addition of IPTG to a final concentration of 0.5 mM. Cells were harvested at an OD600 of 3.0 via centrifugation at 5,000g for 10 minutes at 4 °C. Cell pellets were washed twice with 25 mL S30 buffer per 50 mL culture (10 mM Tris Acetate pH 8.2, 14 mM Magnesium Acetate, and 60 mM Potassium Acetate), centrifuging at 5,000g and 7,000g for the first and second washes respectively. Pellets were flash frozen in liquid nitrogen for storage. Pellets were thawed on ice for 1 hour and resuspended in 1 mL S30 buffer plus dithiothreitol per gram of cell mass (10 mM tris acetate pH 8.2, 14 mM magnesium acetate, and 60 mM potassium acetate, 2 mM dithiothreitol). Cell suspensions were lysed using an Avestin EmulsiFlex®-C3 Homogenizer at a lysis pressure of 23,000 PSI. Cell debris was separated via centrifugation at 10,000g for 10 minutes, and the clarified lysate was collected, flash frozen in liquid nitrogen, and stored at -80 °C.

2.5.13 Cell-Free Protein Synthesis reactions

CFPS reactions utilized the PANOx-SP system³¹⁷. CFPS reactions were composed of the following reagents: 12 mM magnesium glutamate, 10 mM ammonium glutamate, 130 mM potassium glutamate, 1.2 mM ATP, 0.5 mM of each CTP, GTP, and UTP. 0.03 mg/mL folinic acid, 0.17 mg/mL *E. coli* MRE600 tRNA (Roche 10109541001), 100 mM NAD, 50 mM CoA, 4 mM oxalic acid, 1 mM putrescine, 1 mM spermidine, 57 mM HEPES pH 7.2, 2 mM of each amino acid, 33.3 mM PEP, 30% v/v *E. coli* extract, linear DNA templates were added to a final concentration of 3.33 µM, and the remainder water. For reactions co-expressing multiple proteins, each template was added to a final concentration of 0.83 µM. The preparation of these reagents has been described in detail elsewhere³¹⁸. All reaction components were assembled on ice and 12 µL reactions were pipetted into 1.5 mL microtubes. Reactions were allowed to proceed at 30 °C for 20 hours. Protein expression yields were quantified by the addition of 10 uM of L-[¹⁴C(U)]-leucine (Perkin Elmer NEC279E250UC, 11.1GBq/mMole) to CFPS reactions, followed by scintillation counting of precipitated proteins⁷⁷. In cases where proteins were tagged with sfGFP, reaction yields were quantified by sfGFP fluorescence. A standard curve was prepared using radioactive leucine incorporation to measure sfGFP concentration and correlate with fluorescence. To quantify sfGFP fluorescence, 2 µL of a CFPS

reaction was diluted in 48 μ L of water in a Black Costar 96 Well Half Area Plate. Fluorescence was measured using a BioTek SynergyTM H1 plate reader with excitation and emission wavelengths of 485 and 528 respectively. In cases where proteins were tagged with sfGFP, reaction yields were quantified by sfGFP fluorescence. A standard curve was prepared using radioactive leucine incorporation to measure sfGFP concentration and correlate with fluorescence. To quantify sfGFP fluorescence, 2 μ L of a CFPS reaction was diluted in 48 μ L of water in a Black Costar 96 Well Half Area Plate. Fluorescence was measured using a BioTek SynergyTM H1 plate reader with excitation and emission wavelength of 485 and 528 nm respectively.

2.5.14 Linear templates for Cell-Free Protein Synthesis

Plasmid constructs were amplified via PCR using the Q5 High-Fidelity DNA Polymerase and primers ACH113 (ctgagatacctacagcgtgagc) and ACH114 (cgtcactcatggtgattctcacttg) to generate linear DNA templates for cell-free protein synthesis reactions. PCR products were purified using the Zymo Research DNA Clean & Concentrator-5 kits following the recommended protocol.

2.5.15 NanoBiT reactions

NanoBiT[®] reactions were setup using the Promega Nano-Glo[®] Live Cell Assay System following the Promega NanoBiT[®] Technical Manual with minor modifications. Working buffer was prepared by diluting the LCS Dilution Buffer 1:4 in 1x PBS pH 7.3. CFPS reactions were diluted to the desired final concentration in working buffer and allowed to equilibrate at room temperature for the desired amount of time. The Nano-Glo[®] Substrate was used at a 50x final dilution of the stock. Reactions were prepared at a final volume of 2 μ L, and were dispensed into 384 well plates (Nunc 267462) using the Echo[®] 525 liquid handler (Labcyte Inc.). Plates were immediately sealed (Bio-Rad MSB1001) and luminescence was monitored using a BioTek SynergyTM H1 plate reader.

Equilibrium logic gates were assayed by co-expressing all gate components in CFPS. Co-expressed logic gates were diluted 100-fold in working buffer and allowed to equilibrate for 72 hours before measurement. Kinetics of the induced dimerization system was measured by diluting individually expressed reaction components in working buffer to the desired level. The association reaction contained NanoBiT[®]

fusion proteins at a final concentration of 10 nM and input protein serially diluted across a range of concentrations. Reaction luminescence was monitored immediately after mixing components. The NOR gate input titration was set up by first diluting CFPS reactions for each component in working buffer. The mixture of NanoBiT® fusions was prepared at a concentration of 5 nM of each component. Input proteins were serially diluted in working buffer to reach the desired concentration range. Dilutions were equilibrated for 3 hours at room temperature to allow for association of the fusion proteins. Reactions were then set up by adding the appropriate volume of diluted input protein to the NanoBiT® fusion mixture. This was then allowed to equilibrate for 16 hours at room temperature before measurement. NOR gate kinetics were measured by adding 5 nM each of the NanoBiT® fusion proteins and monitoring association. After 20 minutes, 112.5 nM of **1** was added to the reaction and again luminescence was monitored.

2.5.16 T cell husbandry

CD3+ human T cells were ordered fresh from healthy donors from AllCells. Cells were activated using anti-CD3/CD28 Dynabeads from Invitrogen (Cat# 11132D). Cells were maintained in X-VIVO 15 media (Lonza, 04-418Q) with 10% FBS (HyClone Cat# SH30071), 2 mM Glutamax (Gibco Cat# 35050-061) and 1% penicillin-streptomycin (Cellgro, Cat# 30-001-CI) passed through a 0.22 um filter and supplemented with recombinant human IL-2 (Peprotech, 200-02) at a final concentration of 200U/mL. Cells were cultured in 37C w/ 5% CO₂ humidified HeraCell incubators (Thermo Scientific). Cell counts were performed by 0.4% trypan blue dye exclusion (Invitrogen, T10282).

2.5.17 Logic-gated control of TALE-KRAB repressors

We utilized the TALE DNA recognition domain and the KRAB repressor domain from a potent and specific synthetic transcription factor repressor targeting *TIM3*³⁰¹. Separated TALE DNA recognition domain and KRAB proteins each fused to obligate heterodimers using a 15 amino acid linker domain (GGGGGMDAKSLTAWs) retain activity comparable to the single chain TALE-KRAB protein when separate mRNAs encoding these molecules are electroporated into T cells³⁰¹. Logic gated control of TALE-KRAB repressors was engineered by pairing separated TALE and KRAB components with different orthogonal heterodimer pairs (e.g. – **1** of the **1:1'** pair). Design sequences were ordered as gBlocks from IDT and ligated into the TALE backbones.

2.5.18 In vitro transcription of repressors

mRNA was produced from the sequence-confirmed plasmids using the T7 mScript™ Standard mRNA Production System (CELLSCRIPT, C-MS100625). mRNA was purified using Agencourt RNAClean XP beads (Beckman Coulter Cat#A63947), and QC of mRNA was performed on a Fragment Analyzer Infinity (Agilent) using the standard RNA kit (DNF-471-1000).

2.5.19 T cell electroporation

mRNA encoding TALE-TFs or heterodimeric proteins and controls were electroporated into T Cells using the BTX ECM830 Square Wave electroporator from Harvard Apparatus. Cells were collected and spun down, washed twice in PBS, and resuspended at a concentration of 250,000 cells per well in 100 ul of BTXpress high performance electroporation solution (Harvard Apparatus, Cat# 45-0805). Cells were then mixed with 1 ug of mRNA unless otherwise indicated and multichannel pipetted into MOS 96-Multi-Well Electroporation Plate_2mm (BTX, Cat# 45-0450).

Electroporation Settings: Choose Mode: LV; Set Voltage: 250 V; Set Pulse Length: 5 ms; Set Number of Pulses: 1; Electrode Type: BTX Disposable Cuvette (2mm gap); Desired Field Strength: 3000 V/cm.

Following transfection, cells were gently pipetted into warm media in a 96-well 1 mL deep well block and placed in a humidified 37°C, 5 % CO₂ incubator. GFP mRNA was always included as a control for transfection efficiency, and observed on the Cytoflex 24 hour following transfection.

2.5.20 T Cell flow cytometry

Cells were counted and a minimum of 100,000 cells were collected for analysis using flow cytometry on a CytoFlex S (Beckman Coulter, B75442). Cells were spun down in a 96-well V-bottom plate (Corning, 3894) and washed once in 1x PBS before staining with fluorophore-conjugated antibody diluted with 1X PBS. Staining proceeded for 30 minutes in the dark at room temperature. Following staining, cells were washed once with 1X PBS, spun down, and resuspended in 100-200 ul FACS buffer (2% heat-inactivated Fetal Bovine Serum (HyClone, SH30071.03), 1 mM EDTA (OmniPur, 4050) in 1X PBS, passed through a 0.22 um filter) for flow cytometry analysis. Antibody Brilliant Violet 421™ anti-human CD366 (Tim-3) Antibody (BioLegend, 345008) was used at a concentration of 1:50.

2.5.21 Progesterone-responsive yeast strain construction and growth media

The base *S. cerevisiae* strain used in all progesterone experiments was BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). Successive transformations of this background strain with the multi-transcriptional units encoding the fluorescence-based yeast two hybrid system (“Construction of synthetic genes”) were performed using a standard LiAc/PEG protocol (12). All yeast cultures were grown in YPD media (10 g/L Bacto Yeast Extract, 20 g/L Bacto peptone, 20 g/L dextrose) or SDC media (6.7 g/L Bacto-yeast nitrogen base without amino acids, 2 g/L complete supplement amino acid mix, 20 g/L dextrose).

| | |
|------------------------------|--|
| yGD390 (Fig. 2.1J, Fig. 7.4) | <p>pHTB2s-11'-NLS-ZF43_8-tADH1 --</p> <p>pRPL18B-7-NLS-VP16-tSSA1 --</p> <p>pZ3-7'-11-NLS (SV40)-tENO1::<i>leu2</i>,</p> <p>pZ3-Venus-tENO2 --</p> <p>p43_8(8x)-mScarlet-tPGK1::<i>ura3</i>,</p> <p>pPAB1-Z3PM(fixed)-tTDH1 --</p> <p>pPOP6-GEM-tENO2 -- HIS3(<i>C. glabrata</i>)::<i>HO</i></p> |
|------------------------------|--|

2.5.22 Yeast cell culture

Three colonies from the transformation of the fluorescence-based yeast two-hybrid system were used to inoculate a 1 mL YPD culture in a 2 mL 96 well storage block (Corning) to grow to saturation overnight (about 16 hours).

2.5.23 Progesterone induction

Saturated, overnight culture was diluted 1:500 in SDC and 450ul were aliquoted into individual wells of a new 2 mL 96 well storage block for a two hour outgrowth at 30 °C and 900 r.p.m. in a Multitron shaker (Infors HT). A Progesterone (Fisher Scientific) gradient was prepared at a 10x concentration by making the appropriate dilutions into SDC from a 3.2 mM stock solution. After the two hour outgrowth, 50ul

of progesterone inducer was added to the corresponding wells of the 96 well block and the block was returned to the shaker for six hours of post-induction growth before measurement by flow cytometry.

2.5.24 Yeast cell flow cytometry

Measurement of fluorescent protein reporter expression was performed with a BD LSRFortessa flow cytometer (BD Biosciences) equipped with a high-throughput sampler. After the specified induction duration, cultures were diluted 1:1 in fresh SDC for running through the high-throughput sampler, such that 3,000–10,000 events per well were collected. YFP (Venus) fluorescence was measured using the FITC channel and RFP (mScarlet) was measured using the PE-CF594 channel. Values were calculated on the height measurement and normalized to cell size by dividing by side scatter height (SSC-H). All analysis was performed in Python 3.7 using the package FlowCytometryTools and custom scripts.

2.5.25 Data Deposition

Raw data from native MS experiments has been deposited to

http://files.ipd.uw.edu/pub/de_novo_logic_2019/190522_native_ms_raw.zip

Chapter 3 - A High-Throughput, Automatable, Cell-Free Expression and Screening Platform for Antibody Discovery

3.1 Abstract

Antibody discovery is bottlenecked by the individual expression and evaluation of antigen-specific hits. Here, we address this gap by developing an automated workflow combining cell-free DNA template generation, protein synthesis, and high-throughput binding measurements of antibody fragments in a process that takes hours rather than weeks. We apply this workflow to 119 published SARS-CoV-2 neutralizing antibodies and demonstrate rapid identification of the most potent antibody candidates.

The following text appeared in:

Hunt, A.C.; *et al.* A High-Throughput, Automatable, Cell-Free Expression and Screening Platform for Antibody Discovery. *BioRxiv* (2021).

<https://doi.org/10.1101/2021.11.04.467378>

3.2 Introduction

Antibodies are widely used as protein-based drugs and diagnostics. They are the critical component in immunoassays enabling rapid diagnostics³¹⁹ and constitute one of the fastest-growing classes of therapeutics with nearly 25% of new FDA-approved drugs in 2020 being antibodies^{320,321}. Antibodies have also recently garnered attention as potential countermeasures for emerging pathogens, and currently constitute the majority of emergency use authorized treatments for COVID-19 that inhibit the SARS-CoV-2 virus⁵.

Modern workflows for antibody discovery utilize either directed evolution or the isolation of single B-cell clones from convalescent patients or animals to go from $>10^8$ possible sequences to a pool of $\sim 10^3$ candidates targeting the desired antigen. However, once this pool of candidates has been generated, state-of-the-art workflows still rely on labor-intensive and poorly scalable procedures (e.g., plasmid-based cloning, transfection, cell-based protein expression, protein purification, binding assessment through enzyme-linked immunosorbent assays (ELISAs), etc.) to individually evaluate and identify the best antibody candidates^{29,322}. These labor-intensive procedures take weeks to months and represent a major bottleneck in antibody discovery. The effort to identify antibodies against emerging threats like the SARS-CoV-2 during the COVID-19 pandemic has highlighted (i) the importance of rapid and high-throughput antibody discovery platforms and (ii) the importance of identifying high-affinity antibodies targeting conserved epitopes^{323,324} or non-overlapping epitopes^{325,326} to resist viral escape and increase the ability to neutralize viral variants^{5,327}; both of which have required intensive screening campaigns. A further challenge is that existing antibody discovery processes frequently have low efficiency, with very few of the screened candidates being potent neutralizers in the case of SARS-CoV-2 (Table 8.1). Taken together, these limitations in existing antibody discovery processes suggest the urgent need for faster and higher throughput screens.

Cell-free protein synthesis (CFPS)^{38,328}, the manufacture of proteins without living cells using crude extracts or purified components, is an attractive tool to overcome these limitations. Towards this goal, a variety of CFPS systems for antibody expression have been developed^{55,107,109,121,122,201,329}. However, to our knowledge, an end-to-end (DNA to data) automatable antibody screening workflow combining CFPS with a high-throughput protein-protein interaction screen has yet to be developed.

Here we describe such an integrated pipeline for antibody expression and evaluation to address critical screening limitations in current antibody discovery pipelines. The workflow leverages four key developments (Fig. 3.1A): (i) DNA assembly and amplification methods that do not require living cells, (ii) CFPS systems that can work directly from linear DNA templates and can generate disulfide-bonded antibody molecules, (iii) an Amplified Luminescent Proximity Homogeneous Linked Immunosorbent Assay (AlphaLISA) that enables rapid protein-protein interaction (PPI) or binding characterization without protein purification³³⁰, and (iv) robotic and acoustic liquid handling that enables a highly parallel and miniaturized workflow. Our integrated workflow is end-to-end automatable and enables a single researcher to express and profile the antigen-specific binding of hundreds of antibodies in 24 hours. As a model, we applied our workflow to profile a diverse set of 120 previously published antibodies, 119 of which are antibodies targeting the SARS-CoV-2 spike glycoprotein (S trimer). These antibodies were selected based on the availability of sequence, structural, SARS-CoV-2 neutralization, and binding information, with 84 being drawn from Brouwer *et al.*³³¹ and the remainder from diverse sources^{323,332–346} (Table 8.2 and 8.3). The antibodies span four orders of magnitude in neutralization potency and target a variety of domains and epitopes.

3.3 Results

3.3.3 Development of a cell-free DNA assembly and amplification workflow

We first implemented a cell-free method for DNA assembly and amplification by adapting and optimizing recently reported protocols for high-throughput construction of DNA templates for CFPS^{53,55,201,347}. The method consists of a Gibson assembly step, followed by PCR amplification of the linear expression template (LET) using the unpurified Gibson assembly product as a template. The key idea was to create a versatile approach for rapid construction of DNA templates without the requirement of cell culture, allowing DNA assembly and amplification in less than 3 hours entirely in 384 well plates. To validate the method, we applied it to the assembly and amplification of a LET for sfGFP expression and only observed sfGFP expression in the presence of properly assembled DNA template (Fig. 8.1A-C). To assemble antibody DNA templates, we purchased synthetic, double-stranded linear DNA coding for the desired variable heavy (VH) and variable light (VL) chain sequences. These DNAs were assembled with

DNA coding for the appropriate heavy chain constant (CH1) or light chain constant (CL) antigen-binding fragment (Fab) domains in addition to a separate piece of DNA coding for the backbone of the pJL1 vector. These sequences were subsequently amplified by PCR to generate LETs (Fig. 8.1D-F). Previous works suggest that this workflow could be compatible with PCR products amplified from single B-cells from an immunized animal^{155,201,347}. In addition to being fast, this workflow also affords flexibility, allowing assembly of different antibody formats (e.g., full-length, Fab, sdFab) containing different purification or immobilization tags by using different antibody constant regions in the assembly reaction.

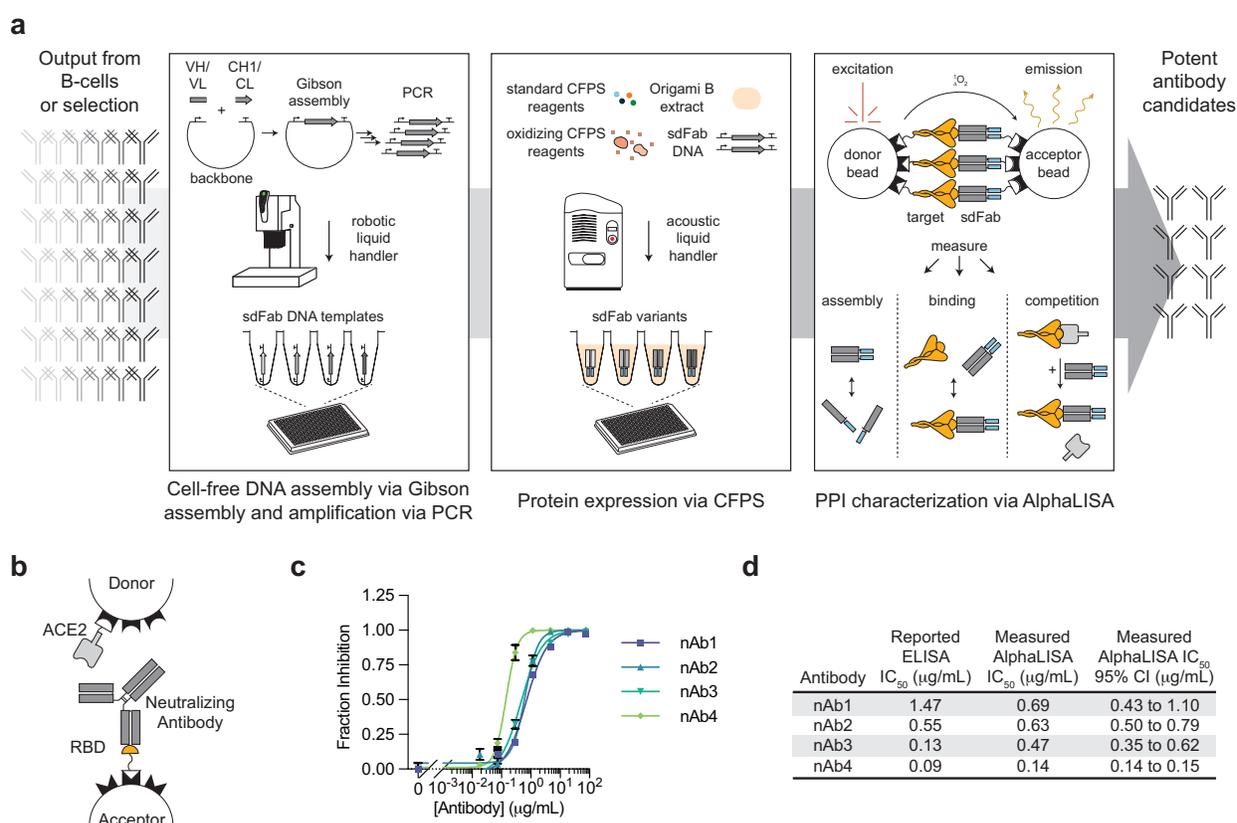


Figure 3.1. A high-throughput, cell-free antibody screening workflow. (A) Schematic of the steps involved in the cell-free antibody screening workflow. (B) Diagram of the AlphaLISA screen for neutralizing antibodies via competition with ACE2 for the SARS-CoV-2 RBD. (C) Evaluation of commercial neutralizing antibodies (nAbs) in the AlphaLISA ACE2 competition screen ($n=3$ independent replicates \pm SEM). (D) Comparison of the reported and measured potencies of commercial neutralizing antibodies.

3.3.3 Development of a crude extract based CFPS system for antibody expression

We next demonstrated rapid antibody expression in a crude *E. coli* based CFPS system. We developed a high-yielding ($1,390 \pm 32 \mu\text{g/mL}$ sfGFP, Fig. 8.1C) crude *E. coli* lysate-based CFPS system

from the Origami™ B(DE3) strain (Fig. 8.2), which contains mutations in the *E. coli* reductase genes *trxB* and *gor* to enable the formation of disulfide bonds in the cytoplasm³⁴⁸. By pretreating the extract with the reductase-inhibitor iodoacetamide (IAM) to further stabilize the redox environment^{97,98,349} and supplementing the reaction with purified *E. coli* disulfide bond isomerase DsbC and prolyl isomerase FkpA^{95,105,109}, we successfully expressed and assembled full-length trastuzumab, a model anti-HER2 antibody³⁵⁰, from linear DNA templates (Fig. 8.2A). However, like others^{55,107,109,201,329}, we found that the efficient assembly of full-length antibodies in CFPS can require further optimization (e.g., temperature, DNA template ratio, DNA template expression timing) which is not optimal for high-throughput screening. Like reports by Ojima-Kato *et al.*^{55,201,329}, we found that the assembly of synthetically dimerized antigen-binding fragments (sdFab, also called ecobodies^{55,201} or zipbodies²⁰¹,) were more consistent than their corresponding standard Fabs in CFPS for a small panel of antibodies and opted to utilize the sdFab format for expression (Supplementary Fig 2B-C). Using acoustic liquid handling we can assemble CFPS reactions to express each sdFab variant from cell-free assembled and amplified DNA in 384-well plates (Fig. 3.1A).

3.3.4 Integration of the AlphaLISA PPI screen into the workflow

Following DNA assembly and CFPS, antigen-specific binding was evaluated. To characterize the PPIs of the expressed sdFab antibody candidates, we developed an AlphaLISA method to characterize PPIs directly from CFPS reactions. AlphaLISA is an in-solution and wash-free assay that is designed for high-throughput screening and is compatible with crude cell-lysates³³⁰. In AlphaLISA, non-covalent capture chemistries are used to immobilize the proteins of interest on donor and acceptor beads, which generate a chemiluminescent signal when in proximity of one another and excited by a 680 nm laser. We developed AlphaLISA methods to enable the measurement of both direct binding to an antigen as well as competition for specific epitopes. We first sought to validate that AlphaLISA is tolerant of crude CFPS reactions. We observed that CFPS does not interfere with the measurement chemistry (Fig. 8.3A), but that certain reaction components can disrupt protein immobilization to the bead which can be circumvented with the appropriate choice of immobilization chemistry (Supplementary Fig 3b-c). We found that the Ni-Chelate beads were not tolerant of the high salt concentrations and high concentration of histidine present in CFPS, likely due to charge screening and Ni chelation respectively hindering immobilization of the hisx6 tagged protein. To validate the ability of AlphaLISA to profile neutralizing antibodies, we tested the ability of four different

commercial antibodies to compete with the SARS-CoV-2 target human receptor Angiotensin-Converting Enzyme 2 (ACE2) for binding of the SARS-CoV-2 Receptor Binding Domain (RBD) and found that our determined rank order of IC_{50} values aligns with the reported ELISA IC_{50} 's (Fig. 3.1B-D). Further, we utilized AlphaLISA to develop a sdFab assembly screen to monitor antibody expression and assembly in CFPS, a laborious step that traditionally requires SDS-PAGE. The measurement immobilizes the heavy and light chains of the sdFab to the AlphaLISA beads, resulting in signal when the two chains are assembled (Fig. 8.3d). The AlphaLISA assembly assay shows consistent prediction of antibody assembly with SDS-PAGE on a panel of sdFabs and can thus be used to identify when sdFab expression or assembly fails (Fig. 8.3E).

3.3.5 Evaluation of a large set of previously published antibodies

Using the developed workflow, we next evaluated a set of 120 unique antibodies using AlphaLISA to measure antibody binding to the SARS-CoV-2 S trimer, binding to the SARS-CoV-2 RBD, competition with ACE2 for RBD binding, and assembly of their heavy and light chains in CFPS (Fig. 3.1A and Fig. 3.2). Antibodies were expressed and evaluated in triplicate. AlphaLISA replicates were found to be consistent with one another, validating that the acoustic liquid handling workflow is robust (Fig. 8.4). Samples were evaluated for significant assembly, binding to, or competition with a given target using a two-sided student's t-test corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate procedure (FDR)³⁵¹. Within the diverse set of 36 antibodies, we observed assembly for 36 out of 36 tested antibodies, S trimer binding for 28 out of 35 antibodies reported to bind the S trimer, RBD binding for 23 out of the 34 antibodies reported to bind the RBD, and ACE2 competition for 16 out of 31 antibodies reported to compete with ACE2 (Fig. 3.2A, Fig. 8.5). For the set of 84 antibodies from Brouwer *et al.*, we observed assembly of 80 out 84 antibodies and binding to the S trimer and RBD for many of the antibodies that showed strong binding via ELISA (Fig. 3.2D-B). We compared ACE2 competition against neutralization since it has been reported that more than 90% of neutralizing antibodies block the RBD and ACE2 interaction^{334,352} and similar competition assays have been reported to correlate with neutralization potency^{334,353} (Fig. 3.2E). We observed ACE2 competition, as well as strong S trimer and RBD binding, for 4 out of 5 antibodies reported to compete with ACE2, which also represent the four most potent neutralizers in the Brouwer *et al.* data set.

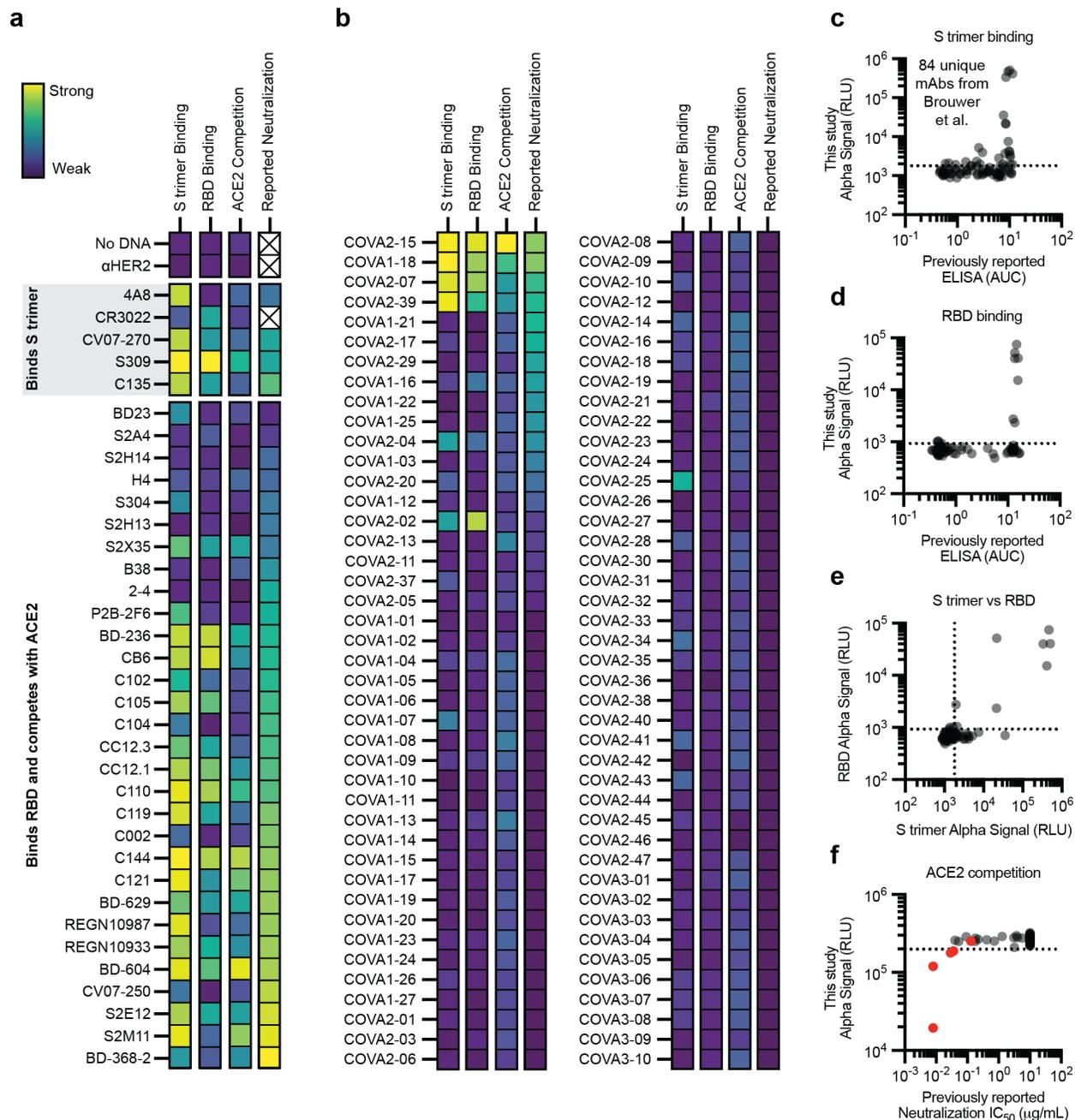


Figure 3.2. Performance of the cell-free antibody screening workflow evaluated on SARS-CoV-2 neutralizing antibodies. (A-F) AlphaLISA data are presented as the mean of 3 independent replicates. A dashed line indicates three standard deviations away from the background signal. (A-B) Heatmap of the binding of previously published antibodies measured using AlphaLISA to detect S trimer binding (log₁₀ scaled), RBD binding (log₁₀ scaled), and ACE2 competition (linearly scaled). AlphaLISA data are presented as the mean of 3 independent replicates. The lowest reported neutralization IC₅₀ value is also plotted for comparison (log₁₀ scaled) and an X indicates no relevant data available (Table 8.2). (A) Heatmap of the binding of 36 diverse antibodies. (B) Heatmap of the binding of all 84 antibodies in the Brouwer *et al.* data set. (C-D) Parity plots comparing the AlphaLISA the 84 antibodies in the Brouwer *et al.* data set vs the published ELISA data. A dashed line indicates three standard deviations away from the background. (C) S trimer binding. (D) RBD binding. (E) Comparison of the S trimer and RBD AlphaLISA binding data. (F) Parity plot comparing the AlphaLISA ACE2 competition data for the 84 antibodies in the Brouwer *et al.* data set

vs the published pseudovirus neutralization data. Antibodies that were reported to compete with ACE2 by Brouwer *et al.* are plotted in red.

3.4 Discussion

Notably, we observed ACE2 competition for 10 out of 13 antibodies in the overall data set whose neutralization IC₅₀ values are less than 0.01 µg/mL. While some less-effective neutralizers could not be completely characterized in our screen, we consistently identified potent neutralizing antibodies in our rapid cell-free screening workflow whose mechanism is ACE2 competition. Consistent with their binding specificities, we observed that 4A8, an n-terminal domain targeted antibody³⁴³, only showed strong interaction with the S trimer and that CR3022, whose target epitope is occluded in the S trimer^{340,354}, showed binding to the RBD, but weak binding to the S trimer. Surprisingly, the S309 antibody in the sdFab format exhibited competition with ACE2 although it has been previously reported not to compete with ACE2³²³, which will require further study. Taken together, the binding and competitive AlphaLISA data generated by our workflow are self-consistent and largely align with the literature (Table 8.4). Further improvements to the dynamic range of the PPI measurements could broaden the utility for performing antigenic mapping of the immune response to antigens. Inclusion of other binding targets could allow researchers to easily evaluate targeting to different domains (e.g., SARS-CoV-2 N-terminal domain) or look for antibodies targeting conserved epitopes by evaluating cross-reactivity with other related viruses (e.g., SARS-CoV, etc.).

In summary, we developed an integrated and automated workflow for antibody screening by combining methods for cell-free DNA assembly and amplification, cell-free protein expression, and highly parallel binding characterization via AlphaLISA. This workflow has two key features. First, it is fast. The entire workflow for all 120 antibodies evaluated in this study was completed in triplicate in less than 24 hours in two consecutive working days by a single researcher, highlighting the workflow's speed and throughput. Second, integration of the AlphaLISA assay in cell-free extracts without the need for protein purification facilitates direct evaluation of synthesized antibodies in high-throughput. This is important because this is frequently the limiting step in previously published methods. Looking forward, we anticipate that the increased speed and throughput afforded by our workflow will enable researchers to easily and

rapidly screen thousands of antibodies, facilitating down-selection to a few highly potent candidates that can be expressed at larger scales in cells or using CFPS and subjected to deeper developability testing. In this way, our method is poised to aid in the discovery of medical countermeasures in future pandemics, and more broadly, in the development of antibodies for therapeutic, diagnostic, and research applications.

3.5 Materials and Methods

3.5.1 Antibody sdFab sequence design

sdFabs were assembled based on a modified version of previously published protocols^{55,201,329}. Antibody sequences were collected from literature and their light chains were classified as either kappa or lambda via the terminal residue of the J-segment in the VL domain. The VH and VL domains were subsequently fused to their corresponding human constant heavy (Uniprot P0DOX5) or human constant light (kappa CL Uniprot P01834 or lambda 1 CL Uniprot P0CG04) chains. At the N-terminus of the VH and VL domains, we chose to include a modified expression tag based on the first 5-residues of the *E. coli* chloramphenicol acetyltransferase gene followed by a Tobacco Etch Virus (TEV) protease cleavage site (protein sequence: MEKKIENLYFQS, DNA sequence: atggagaaaaaatcgaaaacctgtacttccagagc)³⁵⁵ as opposed to the previously published SKIK tag³⁵⁶. The heavy chain was fused to the LZA heterodimer subunit (AQLEKELQALEKENAQLEWELQALEKELAQK) and a strep II tag. The light chain was fused to the LZB heterodimer subunit (AQLKKKLQALKKKNAQLKWKLQALKKKLAQK). Examples of the three types of antibody sequences are detailed below, with the important sequence features highlighted in square brackets [].

sdFab heavy chain constant strepII tagged:

[MEKKIENLYFQS][VH_Sequence][ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL
TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC]GGGGS[AQLEKEL
QALEKENAQLEWELQALEKELAQK]GSSA[WSHPQFEK]

sdFab light chain kappa:

[MEKKIENLYFQS][VL_Sequence][RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKSTYLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC]GGGGS[AQ
LKKKLQALKKKNAQLKWKLQALKKKLAQK]

sdFab light chain lambda 1:

[MEKKIENLYFQS][VL_Sequence][GQPKANPTVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSQRSYSCQVTHEGSTVEKTVAPTECS]GGGGS[AQLKKKLQALKKKNAQLKWKLQALKKKLAQK]

3.5.2 DNA assembly and linear expression template (LET) generation

Proteins to be manufactured via CFPS were codon-optimized using the IDT codon optimization tool and ordered as double-stranded linear DNA containing the desired Gibson assembly overhangs from IDT or GenScript. sfGFP was ordered containing the two pJL1 Gibson assembly overhangs. Antibody VH DNA was ordered with the pJL1 5' and the human IgG1 heavy chain constant 5' Gibson overhangs. Antibody VL DNA was ordered with the pJL1 5' and the human Ig light chain kappa or lambda 1 Gibson assembly overhangs. DNA was resuspended at a concentration of 50 ng/μL and used without amplification.

Additional linear DNA components for Gibson assembly (pJL1 backbone, sdFab heavy chain constant strepII tagged, sdFab light chain kappa constant, sdFab light chain lambda 1 constant) were ordered as gblocks from IDT. These components were amplified using PCR using Q5 Hot Start DNA polymerase (NEB, M0493L) following manufacturer instructions. Amplified DNA was purified using the DNA Clean and Concentrate Kit (Zymo Research, D4006) and diluted to a concentration of 50 ng/μL. Sequences of the utilized components are listed below, with Gibson assembly sequences being denoted by underlined lowercase text and primers for a given amplicon being listed below the DNA sequence.

Gibson assembly overhangs:

pJL1 5' Gibson: ttgtttaacttaagaaggagatatacat

pJL1 3' Gibson: gtcgaccggctgctaacaaagcccgaagg

Human IgG1 heavy chain constant 5' Gibson: gcgtcaacaaaaggtccttcagtttccattagcccct

Human Ig light chain kappa 5' Gibson: cgcacggtcgcgccgctgtcttttttctcct

Human Ig light chain lambda 5' Gibson: ggccaacccaaagcaaaccactgtcactttgttccc

Linear pJL1 plasmid backbone (Addgene plasmid # 69496):

gtcgaccggctgctaacaaagcccgaaggAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTGGCTGAAAGCCAATTCTGATTAGAAAACTCATCGAGCATCAAATGAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCG

TTTCTGTAATGAAGGAGAAAACCTCACCGAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCT
 GCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAG
 TGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATGCATTTCTTTCCAG
 ACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAATCACTCGCATCAACCAAACCGTTATTCAT
 TCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATC
 GAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTT
 CTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACG
 GATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCT
 GTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCAT
 ACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCATTTATACCCATATAAATCA
 GCATCCATGTTGGAATTTAATCGCGGCTTCGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACAC
 CCCTTGATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGATGATATATTTTTATCTTGCAAT
 GTAACATCAGAGATTTTGAGACACAACGTGAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGC
 GTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAGC
 TACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTG
 TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCC
 TGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTT
 ACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAA
 CGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGA
 GAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCA
 GGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTT
 TGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGATCCCGCGAAATTAAT
 ACGACTACTATAGGGAGACCACAACGGTTTTCCCTCTAGAAATAATtttgtttaactttaagaaggagatacat

pJL1_F gtcgaccggctgcta

pJL1_R atgtatatctccttctaaagftaaacaaaattatttcta

Linear sdFab heavy chain constant strepII tagged:

gcgtcaacaaaaggctcctcagttttccattagccctTCTTCTAAGTCAACTAGTGGCGGTAAGTGGCGCTCTTGGGTG
 TTTGGTTAAAGATTACTTCCAGAACCGGTTACGGTCTCGTGGAACCTCTGGTGCACTGACATCGGGC

GTACATACATTTCCCGCAGTTTTGCAGTCTTCGGGACTGTATTCTCTTTCATCGGTGGTTACAGTCCC
 TAGCTCTTCCCTGGGTACACAGACCTACATTTGTAATGTTAATCATAAGCCGAGTAATACTAAGGTGG
 AAAAAAGGTGGAACCGAAGTCTTGTGGTGGTGGCGGGTCAGCTCAACTGGAGAAGGAGTTACAGG
 CACTGGAAAAAGAGAATGCTCAACTTGAGTGGGAATTACAGGCGTTAGAAAAAGAACTGGCCCAGAA
 GGGTTCTAGCGCATGGTCACATCCCCAGTTTCGAAAAATAAAgtcgaccggctgctaacaagcccgaaagg

IgGC_F: GCGTCAACAAAAGGTCCTTCAGTTTTC

pJL1_3'Gib_R: CCTTTCGGGCTTTGTTAGCAGC

Linear sdFab light chain kappa constant:

cgcacggctcgccggcccgctctgtctttattttctctctTCTGATGAACAGCTTAAATCTGGGACAGCTTCTGTTGTATGTT
 TATTAACAACACTTTTACCCGCGTGAGGCAAAAGTTCAATGGAAGGTAGACAACGCACTGCAAAGCGG
 AAATTCGCAGGAGTCAGTTACCGAACAGGATTCCAAGGATAGTACCTACTCCTTAAGTTCAACATTAA
 CCCTGTCAAAGGCGGACTATGAAAAACATAAGGTATATGCCTGCGAAGTAACTCATCAGGGCTTATC
 ATCCCCAGTTACAAAATCTTTCAACCGTGGAGAATGCGGCGGCGGAGGTAGCGCGCAGCTTAAGAA
 AAAATTGCAAGCCCTTAAAAAATAAATGCCCAACTTAAATGGAAGCTGCAAGCCTTAAAAAAGAAAT
 TGGCGCAGAAGTAAAgtcgaccggctgctaacaagcccgaaagg

kLC_F: TCGCGGCGCCGTCTG

pJL1_3'Gib_R: CCTTTCGGGCTTTGTTAGCAGC

Linear sdFab light chain lambda 1 constant:

ggccaacccaaagcaaaccctgtcactttgttcccgCCCTCAAGCGAGGAACTTCAGGCTAATAAGGCCACGCTT
 GTTTGCCTGATCTCAGACTTTTATCCCGGTGCCGTAACAGTGGCTTGGAAGGCAGATGGTTCGCCG
 GTCAAAGCGGGCGTGGAAACTACAAAGCCATCGAAACAGTCAAACAATAAATATGCGGCATCAAGTT
 ACTTGAGCCTTACCCAGAACAGTGGAAAGTCACACCGCTCGTACAGTTGTCAAGTTACACACGAGGG
 AAGTACAGTTGAAAAGACCGTTGCCCAACTGAATGTTTCAGGCGGTGGTGGCTCAGCGCAGTTAAA
 GAAAAACTGCAGGCTTTGAAGAAAAAGAAATGCTCAATTAAGTGGAAATTGCAGGCGTTGAAGAAG
 AACTTGCGCAGAAGTAAAgtcgaccggctgctaacaagcccgaaagg

ILC_F: GGCCAACCCAAAGCAAACC

pJL1_3'Gib_R: CCTTTCGGGCTTTGTTAGCAGC

Gibson assembly was used to assemble protein open reading frame DNA with the pJL1 backbone following the published protocol with the addition of 3.125 µg/mL of ET SSB (NEB, product no. M2401S)^{357,358}. 20 ng of purified, linear pJL1 backbone, 20 ng of purified, linear sdFab VH or VL constant DNA, and 20 ng of the protein open reading frame insert were combined in 2 µL Gibson assembly reactions and incubated at 50°C for 30 minutes. The unpurified assembly reactions were diluted in 40 µL of nuclease-free water (Fisher Scientific, AM9937) and 1 µL of the diluted reaction was used as the template for a PCR to generate linear expression templates (LETs) for CFPS. Linear expression templates were amplified via PCR using the pJL1_LET_F (ctgagatacctacagcgtgagc) and pJL1_LET_R (cgtcactcatggtgattctcacttg) primers in a 50 µL PCR reaction using the Q5 Hot Start DNA polymerase (NEB, M0493L) following manufacturer instructions.

The DNA sequence of the *P. pyralis* luciferase containing a c-terminal strepII tag (fLuc, Uniprot Q27758) used as a negative control is below and was cloned into the pJL1 vector.

```
atggaagacgctaagaacattaagaagggacctgctccattctacccccctgaagacggcactgcaggtgagcagcttcataaagcgatgaagc
ggtatgcgtagtctctggcacgatgccttcactgacgcgcacatcgaagtcgaatcacctacgctgaatactttgagatgagtgctgctggcgga
agccatgaagcggtatggccttaacacgaaccaccgcatcgtgtttgtagcgagaattccttacaattctcatgccgctctggcgcgctgtttattgg
tgtggccggtgcaccagccaatgacatctataatgagcgcgagttgtgaactccatgaacatttctcaaccaacagtggtgttcggttcaaagaaagg
cttacagaaaatctaaacgttcaaaagaaactgccgattatccagaagatcatcattatggatagtaagactgactaccagggctccagtcattgt
atacattcgtgacgagtcacctgccccgggttttaacgagtagcactttgtcccagagagctttgatcgcgacaagaccatcgccctcattatgaata
gcagtggttcgacgggtagcccaaaggagtgccctgccccatcgtaccgctgctcgttctccatgcccgcgacccaatcttcggaatca
aatcatccccgacacggcaatctgtcggctgctccggttcaccatggctttggaatgtttacgacactcggttacctcatctcgggttccgctgctctg
atgtatcgcttcgaggaagagttgttctacgttcgcttcaggactacaagattcaatccgcccttctggtccccactttgttcagttctttgctaagagcac
cttaattgataagatgacctctcaacttacacgagattgcgagcgggtgctccccctcagcaagaggttgagagggcggttgctaagcggtttca
tctgcccggtatccgtaaggttacggcctcaccgaaaccacttctgccattctatcactccggaaggtgacgataagcctggggcagtggtgtaa
ggtgtacccttctcgaggctaaggtgtggattagatacggggaagaccttaggtgtgaaccagcgggtgaactgtgcttcggttcggttccgatgatta
tgtcgggttatgtaatgacccccgaggctacgaacgcgcttatcgataaggacggttgcttctccggcgacatcgcttactgggatgaggatgag
cacttctcatcgttgaccgctgaagagtctcatcaagtataagggatgcaagtcgctccggcagagttagagagcatcttactccagcacccta
atcttcgatgctgggggtgccgggctcccaggcgacgatgccggcgagctgccggcgcggtagttgttttagagcatggcaagaccatgaccgaa
aaggagattgtagactacgctcgcgagtcgaagtaaccacagcgaagaagctccgcggtggagtggtctttgtgacgaggtgcctaaaggcctgac
```

gggcaaaacttgacgcgcgtaagatccgtgagatcctcatcaaagcgaagaaggggtgggaagagtaagctggggagttcaggttggtcccacccg
caatttgagaagtga

3.5.3 Cell extract preparation for cell-free protein synthesis

E. coli Origami™ B(DE3) (Novagen, 70837) extracts were prepared using a modified version of established protocols^{33,292}. Briefly, a 150 mL Origami™ B(DE3) starter culture was inoculated in LB from a glycerol stock and cultured in a 250 mL baffled flask at 37 °C for 16 hours. The 2xYTP was prepared lacking glucose in 75% of the final volume and sterilized using an autoclave. A 4x glucose solution was prepared and autoclaved separately, then added to the medium immediately before use. The starter cultures were used to inoculate 1 L of 2xYTPG media (16 g/L tryptone, 10 g/L yeast extract, 5 g/L sodium chloride, 7 g/L potassium phosphate dibasic, 3 g/L potassium phosphate monobasic, 18 g/L glucose) in a 2.5 L Full-Baffle Tunair shake flask at an initial OD600 of 0.08. Cells were cultured at 37 °C at 220 RPM in a shaking incubator. Cultures were grown until OD600 0.4-0.6, at which point the expression of T7 RNA polymerase was induced by the addition of IPTG to a final concentration of 0.5 mM. Cells were harvested at an OD600 of 2.5 via centrifugation at 12,000g for 1 minute at 4 °C. Cell pellets were washed three times with 25 mL S30 buffer per 50 mL culture (10 mM Tris Acetate pH 8.2, 14 mM Magnesium Acetate, and 60 mM Potassium Acetate). Pellets were resuspended in 1 mL S30 buffer per gram of cell mass. Cell suspensions were lysed using a single pass on an Avestin EmulsiFlex-B15 Homogenizer at a lysis pressure of 24,000 PSI. Cell debris was separated via centrifugation at 18,000g for 20 minutes, and the clarified lysate was collected, flash-frozen in liquid nitrogen, and stored at -80 °C.

3.5.4 DsbC and FkpA expression and purification

Protein expression, purification, and his tag removal were performed similarly to previously reported³⁵. DsbC (Uniprot P0AEG6, residues 21-236) and FkpA (Uniprot P45523, residues 26-270) were ordered as gBlocks from IDT containing a c-terminal, TEV cleavable his tag (GSENLVYFQSGSHHHHHHHHH) and cloned into pET28a. Plasmids were transformed into BL21 Star™ DE3, plated on LB agar, and cultured overnight at 37°C. 1 L of Overnight Express TB (Fisher Scientific, 71491-4) was inoculated by scraping all colonies on a transformation plate and cultured at 37°C in 2.5 L tunair flasks (IBI Scientific, SS-8003) at 220 rpm overnight. Cells were harvested, resuspended at a ratio of 1 g cell mass to 4 mL resuspension buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 1X HALT protease

inhibitor without EDTA (Fisher Scientific, 78429), 1 mg/mL lysozyme, 62.5 U/mL cell suspension of benzonase (Sigma Aldrich, E1014-25KU) and lysed using an Avestin B15 homogenizer at 24,000 PSI. The lysate was spun down 14,000 x g for 10 min and the clarified supernatant was incubated with Ni-NTA Agarose (Qiagen, 30230) for 60 min on an end-over-end shaker. The resin was spun down 2,500 x g for 2 min, the supernatant removed, resuspended in wash buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 50 mM Imidazole), loaded on a gravity flow column, and subsequently washed with 20X resin volumes of wash buffer. Protein was eluted using elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM Imidazole) and exchanged into 50 mM HEPES pH 7.4, 150 mM NaCl using PD-10 desalting columns (Cytvia, 17-0851-01) according to manufacturer instructions.

His tags were removed via cleavage by ProTEV Plus (Promega, V6102). Before cleavage, 10% v/v glycerol was added to the protein. ProTEV Plus was added to a concentration of 0.5 U/ μ g purified protein and DTT was added to a concentration of 1 mM. Cleavage reactions were carried out at 30°C for 4 h. Free His tag and ProTEV Plus were removed by incubating with Ni-NTA Agarose for 1 hour at 4°C and collecting the supernatant. Proteins were subsequently concentrated to > 1mg/mL (Millipore, UFC800396). His tag removal was validated via SDS-PAGE and the AlphaScreen Histidine (Nickel Chelate) Detection Kit (Perkin Elmer, 6760619C).

3.5.5 Cell-free protein synthesis

CFPS reactions were composed of the following reagents: 8 mM magnesium glutamate, 10 mM ammonium glutamate, 130 mM potassium glutamate, 1.2 mM ATP, 0.5 mM of each CTP, GTP, and UTP. 0.03 mg/mL folinic acid, 0.17 mg/mL *E. coli* MRE600 tRNA (Roche 10109541001), 100 mM NAD, 50 mM CoA, 4 mM oxalic acid, 1 mM putrescine, 1 mM spermidine, 57 mM HEPES pH 7.2, 2 mM of each amino acid, 33.3 mM PEP, 20% v/v *E. coli* extract, varying concentrations of DNA template, and the remainder water. The preparation of these reagents has been described in detail elsewhere³¹⁸. For DNA templates, plasmids were used at a concentration of 8 nM, and unpurified linear PCR products were used at 6.66 %v/v. For the expression of antibodies, each template was added to a final concentration of 6.66 %v/v. For antibody and sdFab expression 4 mM oxidized glutathione, 1 mM reduced glutathione, 14 μ M of purified DsbC, and 50 μ M FkpA were also supplemented to the reactions. Additionally, for oxidizing CFPS reactions, cell-extracts were treated with 50 mM iodoacetamide (IAM) at room temperature for 30 minutes before use

in CFPS³⁵⁹. All reaction components were assembled on ice and were either run as 12 μL reactions in 1.5 mL microtubes or 2 μL reactions in 384 well plates (BioRad, HSP3801). For 2 μL reactions, components were transferred to the plate using an Echo 525 acoustic liquid handler. A mix containing all the CFPS components except for the DNA was dispensed from 384PP Plus plates (Labcyte, PPL-0200) using the BP setting. The DNA (unpurified PCR products) was dispensed from a 384LDV Plus plate (Labcyte, LPL-0200) using the GP setting. Reactions were allowed to proceed at 30 °C for 20 hours.

To quantify sfGFP fluorescence, a standard curve was prepared using previously reported methods³³. Radioactive leucine was added to CFPS at a final concentration of 10 μM of L-[14C(U)]-leucine (Perkin Elmer NEC279E250UC, 11.1GBq/mMole), followed by precipitation of the expressed proteins and scintillation counting⁷⁷. To quantify sfGFP fluorescence, 2 μL of a CFPS reaction was diluted in 48 μL of water in a Black Costar 96 Well Half Area Plate. Fluorescence was measured using a BioTek SynergyTM H1 plate reader with excitation and emission wavelengths of 485 and 528 respectively. Scintillation counts and fluorescence were fit to determine a standard curve for use with non-radioactive samples.

To visualize antibody assembly, proteins were labeled during CFPS with FluoroTectTM (Promega, L5001). FluoroTectTM was included in the CFPS reaction at 3.33 %v/v. After protein synthesis, RNaseA (Omega Bio-Tek, AC118) was added to 0.1 mg/mL and the sample was incubated at 37 °C for 10 minutes. Samples were subsequently denatured at 70 °C for 3 minutes, then separated via SDS-PAGE and imaged using a LI-COR Odyssey Fc imager on the 600 channel.

3.5.6 AlphaLISA reactions

AlphaLISA reactions were carried out in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mg/mL BSA, and 0.015% v/v TritonX-100 (hereafter referred to as Alpha buffer). All components were dispensed using an Echo 525 liquid handler from a 384-Well Polypropylene 2.0 Plus microplate (Labcyte, PPL-0200) using the 384PP_Plus_GPSA fluid type. All components of the AlphaLISA reactions were prepared as 4x stocks and added as 0.5 μL to the final 2 μL reaction to achieve the desired concentration. All AlphaLISA reactions were performed with CFPS reactions diluted to a final concentration of 2.5 %v/v. AlphaLISA beads were combined to prepare a 4X stock in Alpha buffer immediately before use and added to the proteins to yield a concentration of 0.08 mg/mL donor beads and 0.02 mg/mL acceptor beads in the final reaction. All reactions were incubated with AlphaLISA beads for 1 hour before measurement. AlphaLISA measurements

were taken on a Tecan Infinite M1000 Pro plate reader using the AlphaLISA filter with an excitation time of 100 ms, an integration time of 300 ms, and a settling time of 20 ms. Before measurement, plates were allowed to equilibrate inside the instrument for 10 min. For measurements involving sdFabs, protein A AlphaLISA beads were avoided due to the ability of protein A to bind human subgroup VH3 Fabs³⁶⁰.

The impact of CFPS reagents on AlphaLISA was determined by serially diluting the specified reagents in Alpha buffer and combining them with the specified AlphaLISA conditions. The TrueHits kit (Perkin Elmer, AL900) was used to assess the impact of the CFPS reagents on the Alpha detection chemistry. CFPS reagents were mixed with the donor and acceptor beads and incubated for 2 hours before measurement. His tagged RBD (Sino Biological, 40592-V08H) and human FC tagged human ACE2 (GenScript, Z03484) were used to evaluate the impact of CFPS reagents on capture chemistries. RBD and ACE2 were diluted in Alpha buffer, mixed at a final reaction concentration of 10 nM each, combined with the CFPS reagents, and allowed to incubate for 1 hour. Donor and acceptor beads were subsequently added and allowed to incubate for a further hour before measurement. Protein A Alpha donor beads (Perkin Elmer, AS102), Ni-Chelate AlphaLISA acceptor beads (Perkin Elmer, AL108), and anti-6xhis AlphaLISA acceptor beads (Perkin Elmer, AL178) were utilized for detection.

The commercial neutralizing antibody ACE2 competition experiment was performed with the following antibodies: nAb1 (Acro Biosystems, SAD-S35), nAb2 (Sino Biological, 40592-MM57), nAb3 (Sino Biological, 40591-MM43), nAb4 (Sino Biological, 40592-R001). ELISA IC₅₀ values were recorded from the product page at the time of purchase and converted to µg/mL assuming a MW of 150,000 Da if reported in M. Antibodies were serially diluted in Alpha buffer and mixed with SARS-CoV-2 RBD (Sino Biological, 40592-V02H) at a concentration of 10 nM in the final reaction and incubated for 1 hour. Mouse FC tagged human ACE2 (Sino Biological, 10108-H05H) was subsequently added and incubated for 1 hour, followed by simultaneous addition of the acceptor and donor beads. AlphaLISA detection was performed using Anti-Mouse IgG Alpha Donor beads (PerkinElmer, AS104) and Strep-Tactin AlphaLISA Acceptor beads (PerkinElmer, AL136). IC₅₀ values were calculated using Prism 9 by fitting the normalized data to [Inhibitor] vs. response -- Variable slope (four parameters) fit with the max constrained to a value of 1.

Assembly AlphaLISA reactions were composed of a final concentration of 10 nM of either Rabbit Anti-Human kappa light chain antibody (abcam, ab125919) or Rabbit Anti-Human lambda light chain

(abcam, ab124719). AlphaLISA detection was performed using Anti-Rabbit IgG Alpha Donor beads (PerkinElmer, AS105) and Strep-Tactin AlphaLISA Acceptor beads (PerkinElmer, AL136). CFPS reaction containing the expressed sdFab of interest was mixed with the appropriate anti-light chain antibody and allowed to equilibrate for two hours before the simultaneous addition of the acceptor and donor beads.

SARS-CoV-2 S trimer binding AlphaLISA reactions were composed of a final concentration of 5 nM His-tagged SARS-CoV-2 S trimer (Acro Biosystems, SPN-C52H9). AlphaLISA detection was performed using Strep-Tactin Alpha Donor beads (PerkinElmer, AS106) and Anti-6xHis AlphaLISA Acceptor beads (PerkinElmer, AL178). CFPS reaction containing the expressed sdFab of interest was mixed with the S trimer and allowed to equilibrate for two hours before the simultaneous addition of the acceptor and donor beads.

SARS-CoV-2 RBD binding AlphaLISA reactions were composed of a final concentration of 5 nM Human Fc-tagged SARS-CoV-2 RBD (Sino Biological, 40592-V02H). AlphaLISA detection was performed using Anti-Human IgG Alpha Donor beads (PerkinElmer, AS114) and Strep-Tactin AlphaLISA Acceptor beads (PerkinElmer, AL136). CFPS reaction containing the expressed sdFab of interest was mixed with the RBD and allowed to equilibrate for two hours before the simultaneous addition of the acceptor and donor beads.

ACE2 and RBD competition AlphaLISA reactions were composed of a final concentration of 2 nM Biotinylated SARS-CoV-2 S trimer (Acro Biosystems, SPN-C82E9) and 2 nM Human FC-tagged human ACE2 (GenScript, Z03484). AlphaLISA detection was performed using Anti-Human IgG Alpha Donor beads (PerkinElmer, AS114) and Anti-6xHis AlphaLISA Acceptor beads (PerkinElmer, AL178). CFPS reaction containing the expressed sdFab of interest was first mixed with S trimer and allowed to incubate for 1 hour. Subsequently, ACE2 was added and allowed to equilibrate for a further 1 hour before the simultaneous addition of the acceptor and donor beads.

3.5.7 Data analysis

Because of the nature of the intended use of the reported assay as an initial screen, we use the Benjamini and Hochberg False Discovery Rate procedure³⁵¹ to correct for multiple testing, which is less conservative than other standard methods. Statistical analyses were performed in python. Two-sided t-tests were performed using the scipy package and the FDR procedure was performed using the

statsmodels package with a family-wise error rate of 5%. For t-tests, the following samples were considered to be background, and the combined data were used in the t-test. Assembly: No DNA and Buffer controls. S trimer binding: No DNA and Buffer controls. RBD binding: No DNA and Buffer controls. ACE2 competition: No DNA and α HER2.

Chapter 4 - Multivalent Designed Proteins Protect Against SARS-CoV-2 Variants of Concern

4.1 Abstract

New variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) continue to arise and prolong the coronavirus disease 2019 (COVID-19) pandemic. Here we use a cell-free expression workflow to rapidly screen and optimize constructs containing multiple computationally designed miniprotein inhibitors of SARS-CoV-2. We find the broadest efficacy with a homo-trimeric version (TRI2) of the 75 residue ACE2 mimic AHB2 (TRI2-2) designed to geometrically match the trimeric spike architecture. In the cryoEM structure, TRI2 forms a tripod on top of the spike which engages all three RBDs simultaneously as in the design model. TRI2-2 neutralizes Omicron (B.1.1.529), Delta (B.1.617.2), and all other variants tested with greater potency than that of monoclonal antibodies used clinically for the treatment of COVID19 and provides prophylactic and therapeutic protection against SARS-CoV-2 challenge when administered intranasally in mice. Designed miniprotein receptor mimics geometrically arrayed to match pathogen receptor binding sites could be a widely applicable antiviral therapeutic strategy with advantages over antibodies in resistance to viral escape and antigenic drift by construction, and over native receptor domains in greatly reduced chance of autoimmune responses and higher avidity.

The following text appeared in:

Hunt*, A.C.; Case*, J.B.; Park*, Y.J.; Cao*, L.; Wu*, K.; Walls*, A.C.; *et al.* Multivalent Designed Proteins Protect Against SARS-CoV-2 Variants of Concern. *Science Translational Medicine* (2022). (* Indicates equal contribution)

<https://doi.org/10.1126/scitranslmed.abn1252>

4.2 Introduction

SARS-CoV-2 continues to cause a global pandemic with more than 300 million infections and 5.5 million deaths as of January 2022 (<https://covid19.who.int/>). Monoclonal antibodies (mAbs) targeting the SARS-CoV-2 spike (S) glycoprotein³⁶¹ have been an effective treatment for improving disease outcomes for patients with COVID-19²⁻⁵, but many are sensitive to viral escape via point mutations in their epitopes on the S trimer^{327,362}, and producing mAbs in sufficient quantities for population scale use during a global pandemic is technically and financially challenging³⁶³. Indeed, the continued emergence of variants of concern (VOC) jeopardizes the effectiveness of currently approved mAb treatments and vaccines³⁶⁴⁻³⁶⁹. In particular mutations in the rapidly spreading B.1.1.529 (Omicron) variant disrupt binding of most receptor binding motif targeted mAbs, and reduce neutralization potency more than 100-fold for five of the seven clinical mAbs used for the prophylactic or therapeutic treatment of COVID-19³⁷⁰⁻³⁷². Thus, there is an urgent need for interventions whose efficacy is not disrupted by the ongoing antigenic drift, as is the case for a few mAbs^{323,324,373-376}.

As an alternative to mAbs, we previously computationally designed two classes of minibinder proteins that block the SARS-CoV-2 receptor binding domain (RBD) interaction with its host receptor ACE2³⁷⁷. The first class, exemplified by AHB2, incorporates interactions from the RBD-interacting helix of ACE2 in a custom designed 3-helix bundle and has a similar overall binding mode to that of ACE2. The second class, exemplified by LCB1 and LCB3, contain an entirely new designed RBD binding interface. These minibinders neutralize the WA1/2020 SARS-CoV-2 virus with IC₅₀ values in the range of 23 pM (LCB1) to 15 nM (AHB2)³⁷⁷. The designs express at high levels in *Escherichia coli* and are highly thermostable³⁷⁷, which could considerably streamline manufacturing and decrease the cost of goods for clinical development. LCB1 has demonstrated protective activity as both pre-exposure prophylaxis and post-exposure therapy in human ACE2 (hACE2) transgenic mice, but mutations in the B.1.351 (Beta) and P.1 (Gamma) VOC reduce binding potency^{378,379}.

Here we sought to develop constructs containing three minibinder domains that could simultaneously engage all three RBDs on a single S-protein, and by virtue of this multivalent binding potentially neutralize all tested COVID-19 variants. Multivalency can increase the apparent affinity for target antigens³⁸⁰⁻³⁸², including against SARS-CoV-2³⁸³⁻³⁸⁸. We considered two classes of constructs. The first

contain multiple distinct minibinder domains linked together to maximize RBD binding avidity; these constructs have the advantages that LCB1 and LCB3 are very high affinity binders on their own, and the three domains contain different sets of contacts with the RBD making escape in principle more difficult^{325,384}. The second consists of self-assembling homotrimers of minibinders geometrically matched to the 3 RBDs on a single spike; while AHB2 is lower affinity than LCB1 and LCB3, and the sites targeted are less diverse than the first class, homotrimers of AHB2 have the advantage that the ACE2 binding site is inherently less mutable as the virus must bind ACE2 to infect cells^{376,389}. We describe the design, optimization, and escape resistance of both classes of constructs. We find that the top constructs have considerable promise as potential countermeasures in the ongoing COVID-19 pandemic.

4.3 Results

4.3.1 Impact of RBD mutations on minibinder binding

To determine the potential for mutations to arise that disrupt LCB1 and AHB2 binding to the RBD, we performed deep mutational scans using site saturation mutagenesis of the RBD³⁸⁹. We found that for LCB1, the widely observed K417N mutation results in a greater than 10-fold reduction in affinity and the E406W and Y453K/R mutations result in a greater than 100-fold reduction in affinity (Fig. 9.1). For AHB2, we similarly observed several mutations, including K417N, E406W, and Y453K/R that reduce the affinity of the minibinder for the RBD.

4.3.2 Multivalent minibinder design and experimental optimization

To improve the ability of the minibinders to neutralize circulating SARS-CoV-2 variants, we developed multivalent versions with geometries enabling simultaneous engagement of all 3 RBDs in a single S trimer³⁶¹ to increase binding avidity. Multivalent minibinders might be less sensitive to mutations that would escape binding of the monovalent minibinders (a 100x reduction in binding affinity of a sub-picomolar binder would still result in an affinity in a therapeutic range in a multivalent construct)³⁹⁰. We also hypothesized that constructs with binding domains containing different sets of contacts with the target epitope could prevent escape^{325,384}. To design multivalent constructs, we started from optimized versions of the previously described LCB1, AHB2, and LCB3 minibinders (hereafter referred to as monomers MON1, MON2, and MON3, respectively; Table 9.1)³⁷⁷.

To rapidly prototype multivalent minibinder designs, we developed a cell-free protein synthesis (CFPS) workflow which combines an *in vitro* DNA assembly step followed by PCR to generate linear expression templates that are used to drive CFPS and enable rapid prototyping of new minibinder designs (Fig. 9.2). The workflow enables assembly and translation of synthetic genes and generation of purified protein in as little as 6 hours, is compatible with high-throughput, automated experimentation using an acoustic liquid handler (Echo 525), and is easily scaled for the production of mg quantities of protein^{38,53}. To assess multivalent binding, we coupled the workflow to an AlphaLISA protein-protein interaction (PPI) competition assay to enable comparison of dissociation rates of the designed proteins against either the monomeric RBD or the trimeric HexaPro SARS-CoV-2-S-glycoprotein (S6P)³⁹¹. Because multivalency largely impacts dissociation rate constants of protein-protein interactions, we reasoned that an in-solution off-rate screen could distinguish differences between mono- and multivalent binding³⁹². Multivalent minibinders were allowed to fully associate with the target protein, then reactions were split in two and either 100-fold molar excess of untagged competitor (to prevent reassociation) or buffer was added. MON1, MON2, and MON3 target overlapping epitopes³⁷⁷, and thus mono- or multivalent versions of these minibinders were selected as competitors. The ratio of the competitor to no-competitor condition measurements were calculated to determine the fraction of the complex dissociated³⁹³.

Paralleling previous work where trimeric binders were targeted to the sialic acid-binding site on influenza hemagglutinin³⁸², we first designed self-assembling homotrimeric versions of the MON1, MON2, and MON3 miniproteins geometrically matched to the three RBDs in the S trimer (hereafter referred to as TRI; for example, TRI1-1 represents a homotrimer of MON1 with homotrimerization domain 1, Table 9.1). We designed and screened more than 100 different homotrimeric minibinders, with varied linker lengths and homotrimerization domains, using the CFPS workflow. We observed that many of the homotrimeric constructs exhibited slower dissociation rates than the corresponding monomers; much larger effects were observed with dissociation from the S trimer than monomeric RBD consistent with multivalent binding (Fig. 4.1 and Fig. 9.3). In total, we tested eleven different oligomerization domains and found that nine of these domains yielded at least one design with a linker length that improved dissociation rates on par with the top binders (Fig. 9.4). Designs with domains four and eleven exhibited slower dissociation rates compared to their monomeric counterpart, but faster than the top designs, likely indicative of an inability to

simultaneously engage all three target epitopes or dissociation of the oligomerization domains themselves. The top binders exhibited little to no dissociation from S trimer after 7 days of incubation with competitor, indicating a likely apparent dissociation rate constant of $1 \times 10^{-7} \text{ s}^{-1}$ or slower. This is a marked improvement, more than four orders of magnitude for the TRI2 proteins, over the dissociation rate constants of the corresponding monomeric minibinders (Fig. 9.4). We selected two trimeric scaffolds, the designed two ring helical bundle SB175 (domain 2) and the T4 foldon (domain 1)³⁹⁴ (Table 9.2), to proceed with based on the screening results and previous experience with these scaffolds.

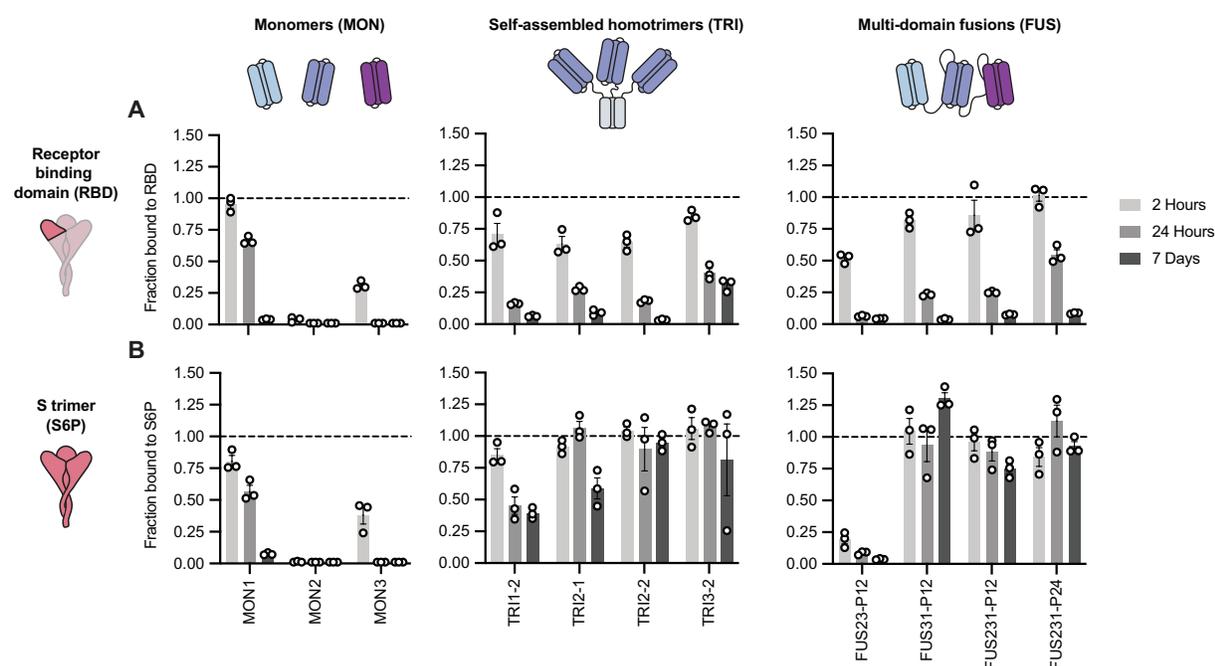


Figure 4.1. Multivalent minibinders exhibit very slow dissociation rates upon binding to the prefusion SARS-CoV-2-S glycoprotein. Dissociation of the minibinder construct was monitored via competition with 100-fold molar excess of untagged MON1 using AlphaLISA (Mean \pm SEM, $n = 3$ technical replicates from a single experiment). (A) Dissociation of the minibinder construct complexed with the receptor-binding domain (RBD). (B) Dissociation of the minibinder construct complexed with the S trimer (S6P).

Next, we generated two- and three-domain fusions of the MON1, MON2, and MON3 minibinders separated by flexible linkers (hereafter referred to as FUS; for example, FUS31-P12 represents a fusion of MON3 to MON1 separated by a 12 amino acid proline-alanine-serine (P12) linker, Table 9.1). We screened more than 100 different fusions using the CFPS workflow, evaluating different minibinder orderings and a range of linker compositions and lengths that span the distances between the termini of the domains when

bound to the “open” and “closed” states of the RBD (Fig. 4.1, and Fig. 9.3, A, B, and F)³⁶¹. We evaluated both glycine-serine (denoted as G) and proline-alanine-serine (denoted as P) linkers³⁹⁵ and observed similar binding characteristics (Fig. 4.1 and Fig. 9.3). We observed occasional truncation of the G linkers during expression and purification by *E. coli* proteases; however, this was less frequent for the P linkers than for the G linkers. FUS31 and FUS231 constructs showed slower dissociation against S6P than RBD, and exhibited slower dissociation than all monomeric minibinders tested, consistent with multivalent S6P engagement. The top binders exhibited little dissociation from S6P after 7 days, indicating a likely apparent dissociation rate constant of $1 \times 10^{-7} \text{ s}^{-1}$ or slower, representing one order of magnitude or greater improvement over the corresponding monomeric minibinder dissociation rate constant (Fig. 9.4).

To determine the potential for low-cost purification by heat treatment, we recombinantly expressed MON1, FUS231-P12, and TRI2-2 in *Escherichia coli*. The heat-treated soluble fraction was enriched with the expressed minibinder and contaminating background proteins were largely precipitated (Fig. 9.5).

4.3.3 Structural studies of minibinders in complex with SARS-CoV-2 S

We next determined how the designed multivalent proteins engage multiple RBDs on a single S trimer (multivalent engagement on a virion typically requires binding of a single S trimer due to the relatively sparse S distribution^{396–398}). For some designs, FUS31-G8 and TRI1-5 for example (Table 9.1), initial screening using negative stain EM revealed considerable cross-linking and aggregation of S trimers upon addition of the constructs (Fig. 9.6), consistent with binding to RBDs on different S trimers. In contrast, for constructs TRI2-2, FUS231-G10, FUS231-P24 and FUS31-G10, we observed less cross-linking, consistent with multivalent engagement of a single S trimer for each minibinder. To determine the binding modes of these compounds to the S trimer and characterize the structure of the MON2 and RBD interactions at high resolution, we carried out cryoEM characterization of these complexes (Fig. 4.2).

The cryoEM structures of the TRI2-2, FUS31-G10, and the FUS231-P24 constructs in complex with S6P were determined at resolutions of 2.8, 4.6, and 3.9 Å respectively (Fig. 4.2, Fig. 9.7-9.10, and Table 9.3), and a negative stain reconstruction was obtained with FUS231-G10 in complex with S6P (Fig. 2E). The TRI2-2/S6P cryoEM structure closely matched the TRI2-2 trimer design, with all three RBDs in the open state bound to MON2 (Fig. 4.2A and B, Fig. 9.7, and Fig. 9.8). In the FUS31-G10 and S6P complex, FUS31-G10 is bound to two RBDs adopting an open conformation (Fig. 4.2C, Fig. 9.7, and Fig.

9.9). The distance between the two RBDs in the open conformation is shorter in the FUS31-G10 than in the FUS231-P24 structure (Fig. 4.2C and D), suggesting that the bound minibinder holds the RBDs together, in agreement with the shorter linkers used in the former minibinder construct. In the structure, two molecules of FUS31-G10 are bound to a single S trimer with the third RBD being occupied by a second FUS31-G10 molecule. In the structure of FUS231-P24 bound to S6P, the three RBDs are participating in complex formation (Fig. 4.2D, Fig. 9.7, and Fig. 9.10). The limited resolution in the region comprising the minibinder-bound RBDs and linkers precludes definitive assignment of minibinder identity at each binding site and relative connectivity between each minibinder module. The distances between the termini of the minibinder domains, however, is compatible with the computational design models and suggestive of engagement of either 2 (FUS31-G10) or 3 of the RBDs (FUS231-P24) in a single S trimer by the multivalent minibinders.

The structure of MON2 in complex with the S trimer has not previously been determined. Starting from the TRI2-2/S6P cryoEM data, we improved the RBD/MON2 densities using focused classification and local refinement yielding a map at 2.9 Å resolution enabling visualization of the interactions formed by MON2 with the RBD. Superimposition of the design MON2 model to the corresponding cryoEM structure, using the RBD as reference, shows that the MON2 minibinder closely matched the design model with backbone C α RMSD of 1.3 Å (Fig. 9.7E and F). Together with previous structures of MON1 and MON3³⁷⁷, these data illustrate the accuracy with which both protein scaffolds and binding interfaces can now be computationally designed.

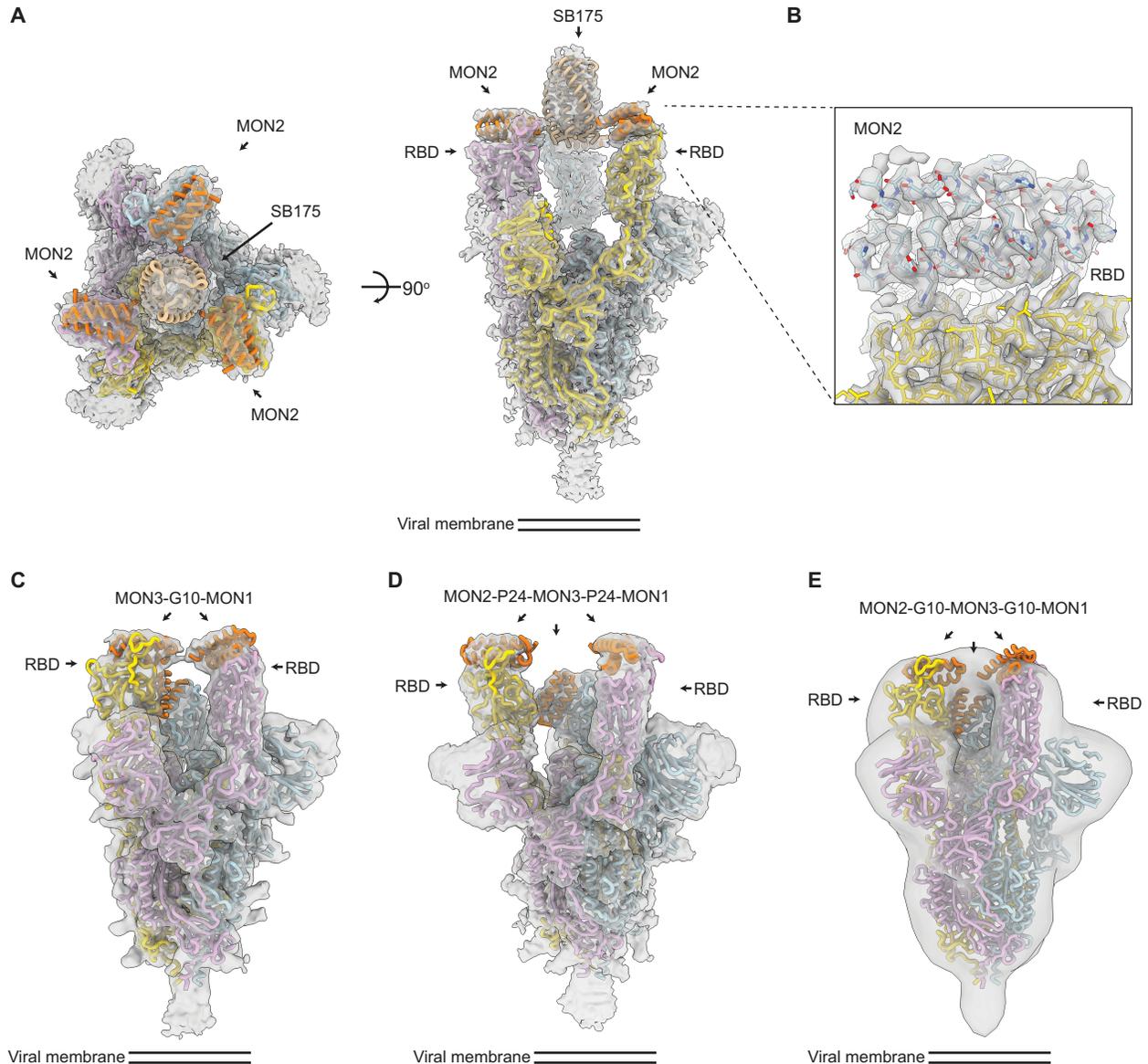


Figure 4.2. CryoEM structures of multivalent minibinders in complex with the SARS-CoV-2 S6P glycoprotein. (A) CryoEM map of TRI2-2 in complex with the S6P in two orthogonal orientations. (B) Zoomed-in view of the TRI2-2 and RBD complex obtained using focused 3D classification and local refinement. The RBD and MON2 built in the 2.9 Å resolution cryoEM map are shown in yellow and blue, respectively. (C) CryoEM map of FUS31-G10 bound to S6P. (D) CryoEM map of FUS231-P24 bound to S6P. (E) Negative-stain EM map of FUS231-G10 in complex with S6P. S and minibinder models were docked in the whole map by rigid body fitting for visualization. In all panels, the EM density is shown as a transparent gray surface, S protomers (PDB 7JZL) are rendered in yellow, cyan, and pink and minibinders (PDB 7JZU, 7JZM, and MON2 structure was determined in this study) are shown in orange.

4.3.4 Multivalent minibinder enables rapid detection of SARS-CoV-2 S

Having confirmed the binding mode of the FUS231 proteins via cryoEM, we designed an S trimer sensor, reasoning that the high affinity binding of the FUS231 proteins to the S trimer could make a useful

diagnostic³⁹⁹. We hypothesized that it would be possible to construct a bioluminescence resonance energy transfer (BRET) sensor for S trimer, where simultaneous engagement of all three minibinders in FUS231 with the S trimer would bring the N- and C-termini close enough together to enable efficient energy transfer. Towards this goal, we designed a BRET sensor based on FUS231-P12 with telLuc and mCyRFP3 fused to the N- and C-terminus of FUS231-P12 respectively (Fig. 4.3A)^{400,401}. Upon binding of the sensor protein to a stabilized S with 2 proline mutations (S2P)^{361,399}, we observed a 350% increase in the 590 nm/470 nm BRET ratio, which was not observed when bound to the RBD alone, and determined the limit of detection to be 11 pM S2P (Fig. 4.3B and C, and Fig. 9.11). Furthermore, these results support the proposed multivalent binding mode for the FUS231 proteins.

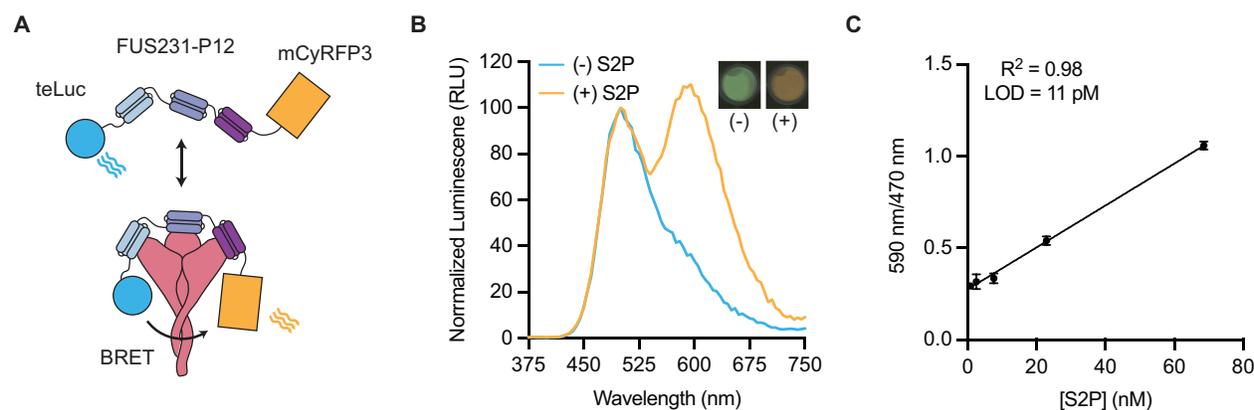


Figure 4.3. FUS231-P12 enables detection of SARS-CoV-2 S trimer via BRET. (A) Schematic representation of the BRET sensor, teluc-FUS231-P12-mCyRFP3, to detect S trimer. (B) Luminescence emission spectra and image of the BRET sensor (100 pM) in the presence (orange trace, 100 pM) and absence (blue trace) of S2P. Emission color change was observed using a mobile phone camera (inset top right). (C) Titration of S2P with 100 pM sensor protein (Mean \pm SEM, $n = 3$ technical replicates from a single experiment).

4.3.5 Multivalent minibinders bind tightly to SARS-CoV-2 variants

We next evaluated the resiliency of the binding of multivalent minibinders to the previously identified MON1 and MON2 escape mutants as well as mutations present in the B.1.1.7 (Alpha), B.1.351 (Beta), and P.1 (Gamma) SARS-CoV-2 VOC. We first measured the off-rate of the best multivalent minibinders using competition AlphaLISA with TRI2-1 against a panel of mutant S glycoproteins (Fig. 4.4A). Multivalent minibinders were allowed to fully associate with mutant S trimers and subsequently were competed with 100-fold molar excess of untagged TRI2-1 to measure dissociation of the complex. The two-domain fusions

(FUS23 and FUS31) showed little increased resilience to the tested point mutants. The three-domain fusions (FUS231) retained binding to the tested mutants, indicating they are more resistant to mutations than their monomeric counterparts, although E406W, Y453R, and the combination of K417N, E484K, and N501Y mutations (present in the B.1.351 S trimer) increased the dissociation rate more than 100-fold. Consistent with these results, we also observed increased dissociation rates for the FUS231 proteins against the B.1.351 and P.1 spikes via surface plasmon resonance (SPR) (Fig. 9.12). The TRI1 and TRI3 homotrimers showed similar mutational tolerance in the competition experiment, with the same E406W, Y453R, and B.1.351 mutations causing increased dissociation rates. Strikingly, the TRI2 designs showed little dissociation after 24 hours against any of the tested S trimer mutants.

We subsequently screened the top multivalent minibinders for binding to mutant S trimers by ACE2 competition ELISA, which correlates with neutralization potency³³⁴. The minibinders were pre-incubated with the S6P variants before binding to immobilized ACE2 (Fig. 4.4B and Fig. 9.13). In line with deep mutational scanning data, we observed impaired binding to the E406W, K417N, and Y453R mutants in addition to several other mutants. Two mutations, Y453F and E484K, improved MON2 binding, consistent with MON2 mimicry of the ACE2 interaction interface³⁸⁹. Compared to the monovalent minibinders, we observed reduced effects of mutations in the competition IC₅₀ values of the FUS231 and TRI2 minibinders and to a lesser extent of the TRI1 and TRI3 minibinders against the tested S6P variants except for E406W (Fig. 4.4B and Fig. 9.13D).

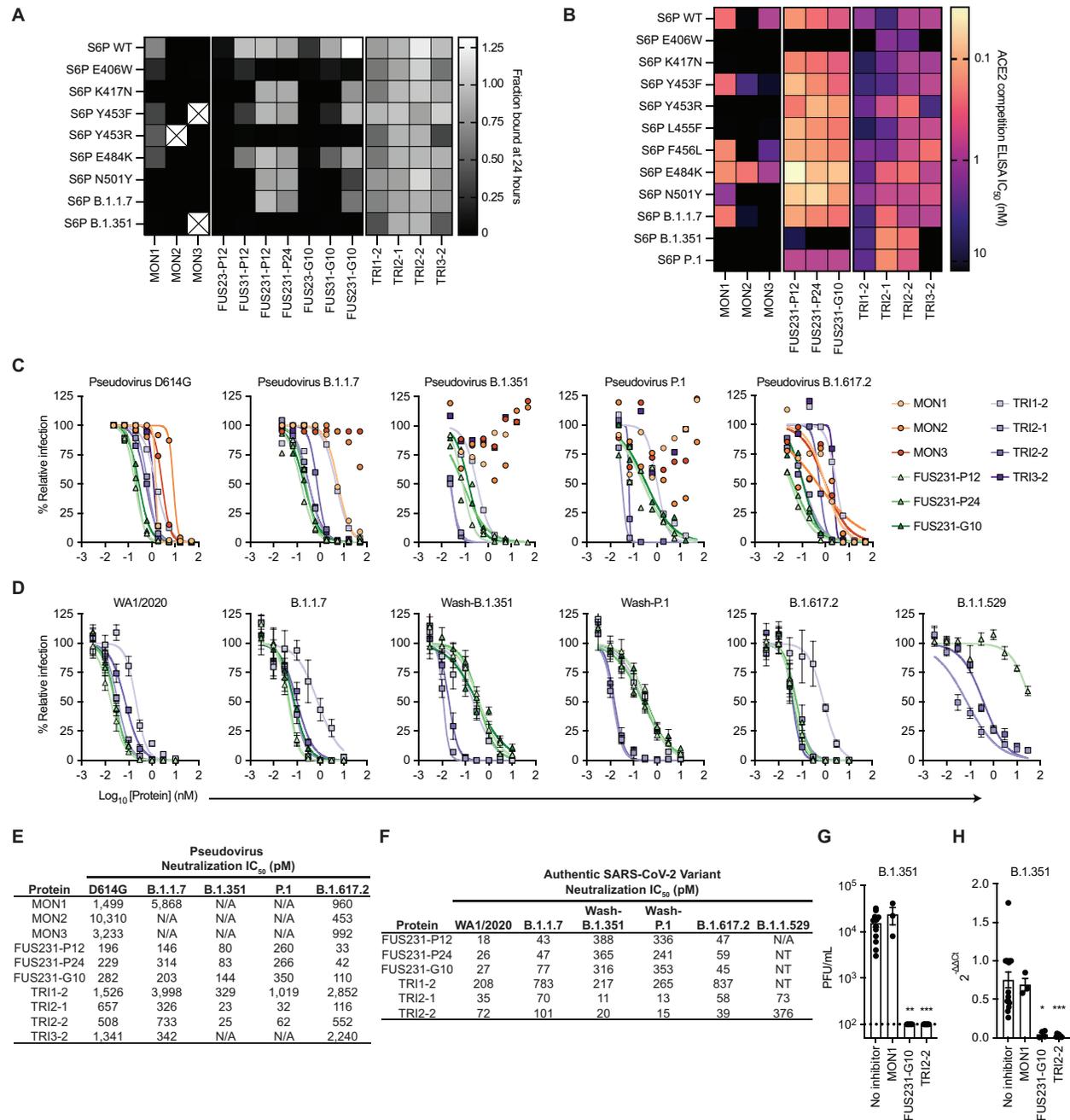


Figure 4.4. Multivalency enhances both the breadth and potency of neutralization against SARS-CoV-2 variants by minibinders. (A) Dissociation of minibinder constructs from S6P variants after 24 hours was measured via competition with untagged TRI2-1 using AlphaLISA (mean, $n = 3$ technical replicates from a single experiment). Cells containing an X indicate insufficient signal in the absence of a competitor to quantify the fraction of protein bound. (B) Competition of minibinder constructs with ACE2 for binding S6P measured via ELISA (mean, $n = 2$ technical replicates representative of two independent experiments). (C) Neutralization of SARS-CoV-2 pseudovirus variants by minibinder constructs (mean, $n = 2$ technical replicates representative of two independent experiments). (D) Neutralization of authentic SARS-CoV-2 by minibinder constructs (mean \pm SEM, $n = 4$ technical replicates from two independent experiments for all but B.1.1.529 where $n = 8$ technical replicates from four independent experiments). (E) Summary of neutralization potencies of multivalent minibinder constructs against SARS-CoV-2 pseudovirus variants. N/A indicates an IC₅₀ value above the tested concentration range and an IC₅₀ greater than 50,000 pM. (F)

Summary of neutralization potencies of multivalent minibinder constructs against authentic SARS-CoV-2 variants. N/A indicates an IC₅₀ value above the tested concentration range and an IC₅₀ greater than 30,000 pM. NT indicates not tested. (G) Quantification of replicating authentic B.1.351 virus in the presence of minibinder constructs (0.3 μM) in human kidney organoids (mean ± SEM, n = 4 biological replicates with 2-3 technical replicates per experiment, Kruskal-Wallis test with Dunn's post-hoc analysis: ** P < 0.01, *** P < 0.001). Dashed line indicates lower limit of detection of plaque assay. (H) Relative gene expression of SARS-CoV-2 envelope protein (SARS-CoV2-E) in kidney organoids post viral infection with and without multivalent minibinders (0.3 μM) (mean ± SEM, n = 4 biological replicates with 2-3 technical replicates per experiment, Kruskal-Wallis test with Dunn's post-hoc analysis: * P < 0.05, *** P < 0.001).

4.3.6 Multivalent minibinders potently neutralize circulating SARS-CoV-2 variants

To investigate the efficacy of the multivalent minibinders for preventing viral infection, we performed neutralization assays with the inhibitors using both pseudotyped lentiviruses and authentic SARS-CoV-2 variants. Against pseudoviruses displaying S proteins corresponding to the B.1.1.7, B.1.351, P.1, B.1.617.1, B.1.617.2 (Delta), and B.1.617.2.1 (Delta plus, AY.1) variants, all three monomer minibinders showed reduced neutralization capacity compared to the Wuhan-Hu-1 D614G strain, whereas many of the multivalent minibinders were less affected in an ACE2 overexpressing cell line (Fig. 4.4C and E, and Fig. 9.14A and C). The same proteins also were evaluated against pseudoviruses containing the E406W, L452R, and Y453F mutations which again had little impact on neutralization for most multivalent minibinders tested (Fig. 9.14A and C). This suggests that the increase in affinity from multivalency improved neutralization breadth. The top neutralizing minibinders from this screen were tested for neutralization of a panel of authentic SARS-CoV-2 viruses including a historical WA1/2020 strain, B.1.1.7, B.1.526 (Iota), B.1.1.529 (Omicron), B.1.617.1, B.1.617.2, and B.1.617.2.1 natural isolates, and chimeric WA1/2020 strains encoding spike genes corresponding to those of B.1.351 (Wash-B.1.351), and P.1 (Wash-P.1) variants. Again, the top candidates maintained pM-range IC₅₀ values (Fig. 4.4D and F, and Fig. 9.14B and D), with the exception of the FUS231 proteins, which did not fully neutralize the B.1.1.529 variant in the tested concentration range. The TRI2 proteins maintained potent neutralization across all tested variants, notably including the B.1.1.7, Wash-B.1.351, Wash-P.1, B.1.617.2, and B.1.1.529 variants.

While Vero cells are useful for neutralization studies, they likely do not fully reflect the human cell infectivity. Recent findings highlight the relevance of using non-transformed human organoid models for SARS-CoV-2 research ⁴⁰². SARS-CoV-2 can infect and readily replicate in human kidney organoids, specifically targeting kidney tubular epithelial cells expressing ACE2 receptors, responsible for viral entry ⁴⁰³. Therefore, we generated kidney organoids from (H9) human embryonic stem cell line ⁴⁰⁴ (Fig. 9.15) and

evaluated the ability of the multivalent minibinders to prevent SARS-CoV-2 viral entry and replication. Replication of the B.1.351 variant was inhibited when the virus was pre-incubated with designed multivalent minibinders FUS231-G10 and TRI2-2, but not with MON1 (Fig. 4.4G). RT-qPCR analysis of viral RNA from the kidney organoids also showed reduced SARS-CoV-2 envelope protein (SARS-CoV2-E) gene expression in the presence of either FUS231-G10 or TRI2-2 (Fig. 4.4H). These data show that designed multivalent minibinders are potent neutralizers of the B.1.351 variant in a human organoid system.

4.3.7 Multivalent minibinders resist viral escape

Given the promising neutralization data, we tested the multivalent minibinders for resistance against viral escape mutations in the S trimer (Fig. 4.5A and B)³²⁷. Plaque assays were performed with a VSV-SARS-CoV-2 S chimera on Vero E6 cells with minibinders included in the overlay to halt spread of non-resistant viruses. In positive control wells, inclusion in the overlay of 2B04, a potent neutralizing antibody targeting the RBD^{327,405–407}, resulted in multiple escape mutants in each plate similar to previously reported escape mutants³²⁷. In contrast, for both FUS231-P12 and TRI2-2, escape mutants were not isolated in 36 replicate wells for each protein (Fig. 9.16). These data indicate that both the FUS231-P12 and TRI2-2 proteins are more difficult to escape than 2B04. Given the known mutation rate of the VSV RNA polymerase L⁴⁰⁸ and the number of viral particles screened, we estimate (Table 9.5) that for the multivalent minibinders the screened pool of viral mutants contains a large fraction of the possible single amino acid substitutions (34% to 88%) and a small fraction of the possible double amino acid substitutions (0.4% to 9.6%) within the region of the RBD that contacts the minibinders. Taken together with the results of the single site saturation mutagenesis studies for the monovalent minibinders (Fig. 9.1) these findings indicate that at least two or more mutations in the RBD are likely necessary to escape binding of the multivalent minibinders.

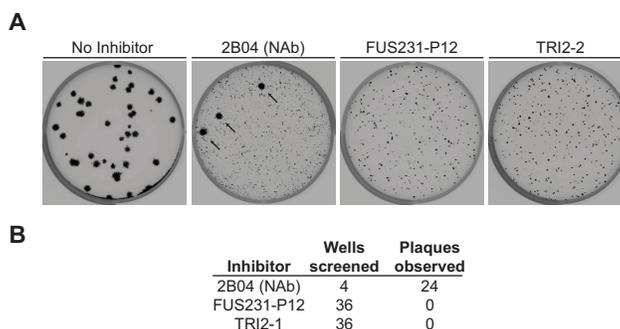


Figure 4.5. Top multivalent minibinder candidates are escape resistant. (A) Plaque assays were performed to isolate VSV-SARS-CoV-2 S chimera virus escape mutants against a control neutralizing antibody (2B04) and the FUS231-P12 and TRI2-2 multivalent minibinders. For each inhibitor tested, Vero cells were incubated with VSV-SARS-CoV-2 S chimera virus for one hour, followed by addition of the inhibitor protein at a fully neutralizing concentration, and further incubation to allow for replication and spread of resistant viruses. 36 independent selections were carried out for each minibinder compound in a single experiment; representative examples are shown in the images. Small plaques are indicative of inhibited viral spreading and large plaques, highlighted by black arrows, are indicative of viral escape mutants spreading. (B) Summary of the results of the viral escape screen.

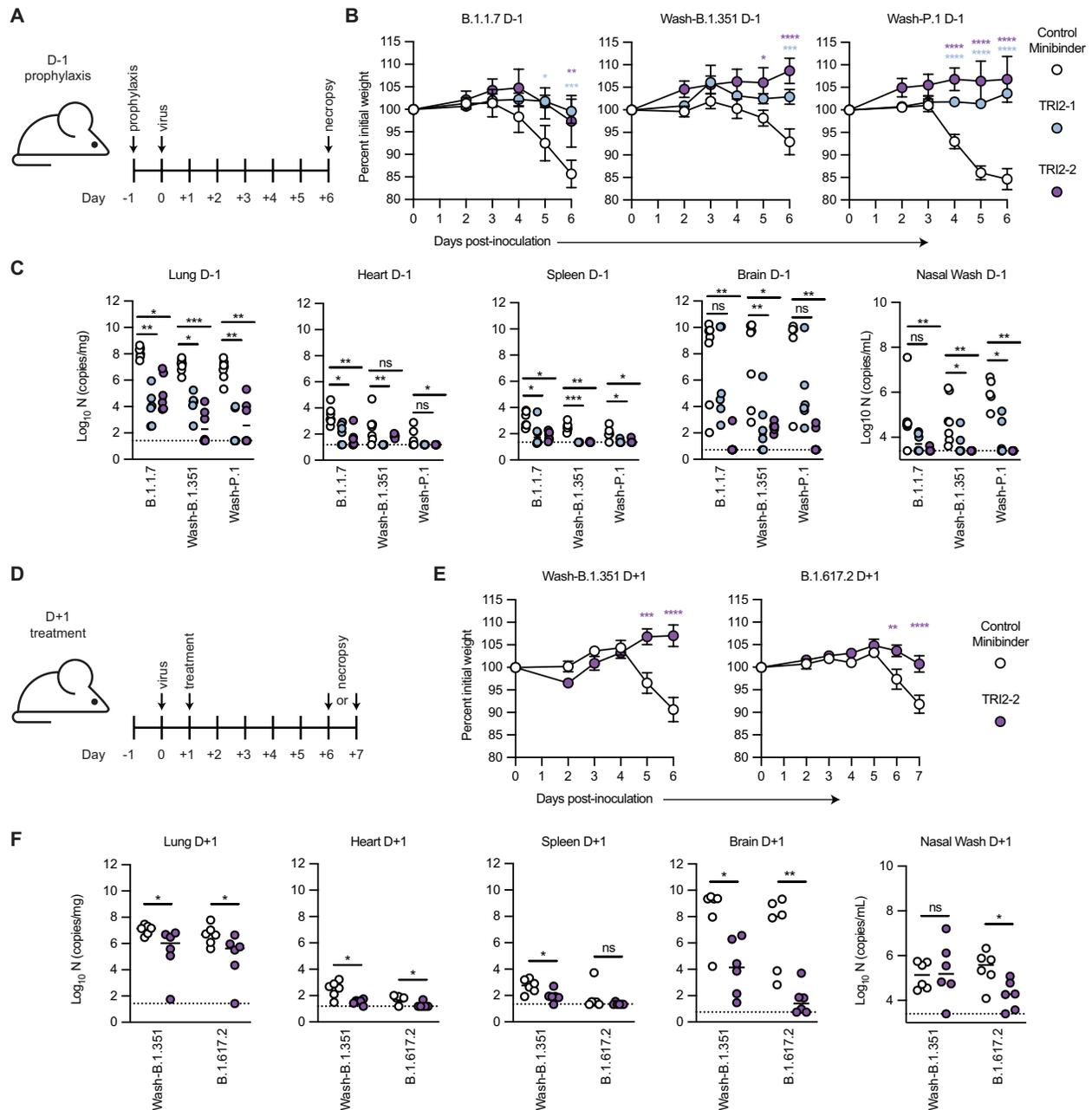


Figure 4.6. Top multivalent minibinder candidates protect mice from SARS-CoV-2 challenge. (A) K18-hACE2-transgenic mice ($n = 6$ from two independent experiments) were dosed with $50 \mu\text{g}$ of the indicated minibinder by i.n. administration ($50 \mu\text{l}$ total) 24 h prior (D-1) to infection with 10^3 focus forming units (FFU) of SARS-CoV-2 variants B.1.1.7, Wash-B.1.351, or Wash-P.1 i.n. on Day 0. (B) Daily weight-change following inoculation (mean \pm SEM; two-way ANOVA with Sidak's post-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). (C) At 6 days post infection (dpi) animals ($n = 6$ from two independent experiments) were sacrificed and analyzed for SARS-CoV-2 viral RNA by RT-qPCR in the lung, heart, spleen, brain, or nasal wash (line at median; Kruskal-Wallis test with Dunn's post-hoc analysis: ns, not significant, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). (D) K18-hACE2-transgenic mice ($n = 6$ from two independent experiments) were dosed with $50 \mu\text{g}$ of the indicated minibinder by i.n. administration ($50 \mu\text{l}$ total) 24 h after (D+1) infection with 10^3 FFU of the SARS-CoV-2 Wash-B.1.351 or B.1.617.2 variant on Day 0. (E) Daily weight-change following inoculation (mean \pm SEM; two-way ANOVA with Sidak's post-test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). (F) At 6 dpi (B.1.351) or 7 dpi (B.1.617.2), animals

(n = 6 from two independent experiments) were sacrificed and analyzed for SARS-CoV-2 viral RNA by RT-qPCR in the lung, heart, spleen, brain, or nasal wash (line at median: two-tailed Mann-Whitney test: ns, not significant, * P < 0.05, ** P < 0.01).

4.3.8 Multivalent minibinder confers protection in human ACE2-expressing transgenic mice

To determine whether the multivalent minibinders can prevent or treat SARS-CoV-2 infection *in vivo*, we performed pre-exposure prophylaxis or post-exposure therapy studies in highly susceptible K18-hACE2 transgenic mice⁴⁰⁹ with TRI2 multivalent minibinders, which retained the most consistent binding to all S trimer variants tested. For prophylaxis, a single 50 µg dose (~2.5 mg/kg) of TRI2-1 or TRI2-2 was administered directly to the nasal cavity (intranasal administration) one day prior to inoculation with 10³ focus forming units (FFU) of the indicated SARS-CoV-2 VOC (Fig. 4.6A). In all cases, intranasal administration of TRI2-1 or TRI2-2 protected mice against SARS-CoV-2-induced weight loss (Fig. 4.6B). At 6 days post infection, viral burden in tissues was reduced at almost all primary (lung and nasal wash) and secondary sites (heart, spleen, brain) of viral replication in TRI2-1 and TRI2-2 treated animals (Fig. 4.6C). To determine the therapeutic potential of TRI2-2, we inoculated K18-hACE2 mice with 10³ FFU of Wash-B.1.351 or B.1.617.2 and one day later, administered a single 50 µg dose of minibinder (Fig. 4.6D). Treatment with TRI2-2 protected against weight loss, and reduced viral burden in all tissues except nasal washes (Wash-B.1.351) or the spleen (B.1.617.2) (Fig. 4.6E and F). TRI2-2 therapy at D+1 reduced infectious virus titers in the lungs of Wash-B.1.351- and B.1.617.2-infected mice (Fig. 9.17). We determined the pharmacokinetics of TRI2-2 after intranasal administration by quantitative competition ELISA. Substantial concentrations of TRI2-2 were detected in the lung lysate and serum 48 hours after administration (Fig. 9.18), but was too low for confident quantification in nasal turbinates after the first time point and for confident quantification in nasal washes at all time points. These results indicate that intranasal administration of TRI2-1 or TRI2-2 can protect as both pre-exposure prophylaxis and post-exposure therapy against SARS-CoV-2 infection in a stringent model of disease.

4.4 Discussion

Both strategies for generating multivalent S binders from miniproteins — self-assembling homotrimers (TRI) and multi-domain fusions (FUS) — yielded designs with improved affinity, neutralization of current and historical VOC, and resistance to escape mutants over their monovalent counterparts^{377,378}.

The TRI2 proteins maintained the strongest binding across all S trimer variants tested, likely because MON2 is an ACE2 mimic, similar to the recently reported S2K146 mAb^{370,376}. This combination of trivalency and receptor mimicry could be a useful general approach for combating viral escape and antigenic drift^{334,370,376,388,410,411}.

The designs also have potential advantages as therapeutics over ACE2 receptor traps and mAbs. When compared to receptor traps^{403,412–414}, TRI2-2 has a low risk of eliciting host-directed anti ACE2 responses due to low sequence similarity between MON2 and ACE2 (Fig. 9.19). On a per mass basis the TRI2 proteins are more potent neutralizers than all currently authorized mAbs for the treatment of COVID-19, and, unlike most clinical mAbs^{370,371}, they maintain activity against Omicron (Table 9.6). The multivalent minibinders are amenable to large-scale production in microorganisms like *E. coli*, making them more cost-effective to manufacture than mAbs³⁶³. Furthermore, their small size and stability may enable direct nebulization into the human upper respiratory tract^{3,415–417}, a strategy that could increase accessibility for patients over the typical intravenous or subcutaneous routes used for neutralizing mAbs.

The high potency of the multivalent constructs, in particular TRI2-2 against Omicron, Delta, and the other tested VOC makes them promising candidate SARS-CoV-2 therapeutics, and they are currently undergoing further preclinical development and IND enabling studies. These efforts will address limitations in our current study. First, anti-drug antibodies are a concern with non-human proteins and while MON1 and other minibinders^{378,418} elicit little or no immune response, additional studies are required to determine the immunogenicity of the multivalent constructs. Second, it will be important to assess the pharmacokinetics following different modes of administration; in humans it may be necessary to distribute the compound deeper into the respiratory system for post infection efficacy. Third, as with any new drug candidate going through the drug development pipeline, it will be necessary to assess its stability as well as its potency, and toxicity after prolonged administration.

In summary, our integration of structure-guided computational protein design, cell-free DNA assembly, cell-free expression, and a competition-based off-rate screen enabled the rapid design and optimization of S trimer-engaging multivalent minibinders. Scaling cell-free expression to manufacture mg quantities of endotoxin-free protein for cell-based neutralization assays further reduced the time required to evaluate lead molecules. The integrated pipeline has direct relevance to diagnostics as well: the FUS231-

based BRET sensor is easy to use, fast, and potentially cheaper than state-of-the-art lateral flow assay-based antigen tests^{419,420}. Looking forward, our integrated computational and experimental pipeline should enable the rapid generation of potent protein-based medical countermeasures and diagnostic reagents against newly emerging pathogens.

4.5 Materials and Methods

4.5.1 Study design

The objective of this study was to design and evaluate multivalent minibinders that neutralize SARS-CoV-2 variants containing mutations within the RBD. At the outset, we hypothesized that multivalency would overcome mutations that reduce binding for individual monomeric minibinders. Designed proteins were evaluated in controlled laboratory experiments, first using biophysical methods with purified proteins (AlphaLISA and ELISA competition assays) followed by *in vitro* methods requiring cell culture (pseudovirus and authentic virus neutralization assays). The top candidates from neutralization assays were screened via electron microscopy for cross-linking multiple S trimers and the best-behaved candidates were subjected to structural analysis via cryoEM. The most promising proteins were evaluated *in vivo* in mice. In all studies where cell lines were used, the cell line is noted in the corresponding methods section. The total number and type of experimental replicates is noted in each figure legend. Details on the *in vivo* mouse study compliance with best practices can be found in the corresponding methods section. No sample-size calculations were performed to power each *in vivo* study. Instead, sample sizes and study endpoints were determined on the basis of previous *in vivo* virus challenge experiments. For all other experiments, sample size was selected based on previous literature and previous experience. In the animal studies, mice were randomly assigned to the control and treatment groups. Animal caretakers and researchers were not blinded to the study groups or during the assessment of the outcomes.

4.5.2 Cell lines and cell culture

For production of S protein variants, Expi293F cells (Thermo Fisher, A14527) were grown in Expi293 Expression Medium (Gibco), cultured at 37°C with 8% CO₂ and shaking at 130 rpm.

For pseudovirus neutralization studies, the HEK293T/17 female human embryonic kidney cell line was obtained from ATCC (CRL-11268). The HEK-ACE2 adherent cell line was obtained through BEI

Resources, NIAID, NIH: NR-52511⁴²¹. All adherent cells were cultured at 37°C with 8% CO₂ in flasks with DMEM + 10% FBS (Hyclone) + 1% penicillin-streptomycin. Cell lines were not tested for mycoplasma contamination nor authenticated.

For authentic SARS-CoV-2 studies, Vero CCL81 (ATCC, CCL-81; RRID:CVCL_0059), Vero-TMPRSS2, and Vero-hACE2-TMPRSS2 (a gift of A. Creanga and B. Graham, NIH) were cultured at 37°C in Dulbecco's Modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 10 mM HEPES pH 7.3, 1 mM sodium pyruvate, 1× non-essential amino acids, and 100 U/ml of penicillin–streptomycin. Additionally, Vero-TMPRSS2 and Vero-hACE2-TMPRSS2 cells were cultured in the presence of 5 µg/mL of blasticidin or puromycin, respectively. The WA1/2020 (2019n-CoV/USA_WA1/2020) isolate of SARS-CoV-2 was obtained from the US Centers for Disease Control (CDC). WA1/2020 stocks were propagated on Vero CCL81 cells and used at passage 6. The B.1.1.7, Wash-B.1.351, and Wash-P.1 (previously described as Wash-B.1.1.28 and Wash BR-B.1.1.248) viruses have been described previously^{367,422}. B.1.526 (S477N), B.1.617.1, B.1.617.2, B.1.617.2.1 (AY.1), B.1.1.529 viruses were isolated from infected individuals. For all strains, infectious stocks were propagated by inoculating Vero CCL81 or Vero-TMPRSS2 cells. Supernatant was collected, aliquoted, and stored at -80°C. All work with infectious SARS-CoV-2 was performed in Institutional Biosafety Committee-approved BSL3 and A-BSL3 facilities at Washington University School of Medicine using positive pressure air respirators and protective equipment. All virus stocks were deep-sequenced after RNA extraction to confirm the presence of the anticipated substitutions.

For VSV SARS-CoV-2 chimera escape selections Vero CCL81 (ATCC, CCL-81; RRID: CVCL_0059) and MA104 (gift from Harry Greenberg) were used as described previously⁴²³.

4.5.3 LCB1 (MON1) and LCB3 (MON3) optimization

Site saturation mutagenesis data were collected for LCB1 and LCB3 using the method described previously³⁷⁷. Beneficial mutations that showed increased binding with RBD were selected and used to construct a Resfile⁴²⁴. 10,000 sequence design trajectories were performed using the Rosetta FastDesign protocol, constrained by the Resfile. Eight sequences were selected based on rosetta score and binding ddG for LCB1 and LCB3 respectively, while keeping the sequence diversity. Genes encoding the selected sequences were cloned into modified pET-29b(+) E. coli plasmid expression vectors (GenScript, N-terminal

8 His-tag followed by a TEV cleavage site) for checking the expression yield (the expression yield is often correlated with protein solubility and stability) and the binding with RBD was characterized with the AlphaLISA assay (see below), the sequences with high expression yield, as well as tight binding affinity with RBD were selected for downstream multivalent constructs design, with the names as LCB1_v2.2 and LCB3_v2.2.

4.5.4 Multivalent fusion constructs design

The CryoEM structures of LCB1 and LCB3 in complex with the spike protein, and the design model of AHB2 were used to determine the order of the monomers in the multivalent fusion constructs. The structures/model were firstly aligned using the spike protein as the reference and the order was determined with the criterion that a shorter linker had to be short when the two linked binders bound with the same spike simultaneously. Four different combinations, AHB2-LCB1, LCB3-LCB1, AHB2-LCB3, LCB3-AHB2, LCB3-AHB2-LCB1, and AHB2-LCB2-LCB1, were tested with either a flexible glycine-serine (G) linker or a proline-alanine-serine (P) linker ³⁹⁵. The linker length was optimized using the AlphaLISA assay and screening by negative stain EM (see below).

4.5.5 Self-assembling homotrimer design: scaffold selection and backbone generation

De novo designed C3 symmetric protein scaffolds and *de novo* designed helical repeat proteins (DHRs) are the basis for the generation of trimeric mini-binders. 34 designed C3 scaffolds and 62 DHRs were selected after preliminary screening based on the geometry matching with the trimeric SARS-CoV-2 S protein and scaffold quality. Fusing the three components, i.e., C3 scaffolds, DHRs and designed mini-binders together through a modified version of the WORMS ⁴²⁵ software, 1096 trimeric binder backbones were generated. This was followed by steric clashing filtering (Rosetta centroid energy < 10), visual inspection in a molecular graphics viewer (Pymol), and sequence design for residues within 9 Å from the fused junction. 56 designs were selected for further validation. To ensure the matching geometry between the fused constructs and the SARS-CoV-2 S trimer with three RBD in the open state (PDB: 7CAK), the WORMS package was customized with the "stack" orientation implemented, where the two cyclic symmetrical axes from both C3 scaffolds and Spike trimer are aligned along Z axis.

4.5.6 Self-assembling homotrimer design: junction modification and flexibility insertion

Considering the strict geometric requirement, the dynamic nature of the S trimer and the low expression successful rate, two new approaches have been pursued to simplify the constructs and enhance the flexibility.

The first approach was creating semi-flexible binders. Starting from the high quality rigid fusion models with a relatively small size (<250 amino acids in total), DHR junctions were omitted. Instead, the last one or two helices from the C3 scaffolds were modified, truncated, or extended with one or two helices using blueprint-based backbone generation^{426,427}. The newly generated last helix was required to be within 12 Å to the first helix from mini-binders in models. (Glycine-Serine)_{x1} to (Glycine-Serine)_{x6} were modeled in Pymol as the flexible linker between the modified C3 scaffolds and mini-binders. Based on the success of this approach, the second was introduced to simplify the constructs further with short flexible linkers only. Four *de novo* C3 scaffolds and six native C3 scaffolds were selected as final candidates with the right geometry, some of which confirmed by the solved cryo-electron microscopy structures. The modification on the last helix was allowed within two to ten residues to be as geometrically compatible as previously. The sequence was optimized based on Rosetta combinatorial sequence optimization packages⁴²⁸⁻⁴³⁰ and homology models.

4.5.7 Deep mutational scanning profiles of minibinder escape

Minibinder escape mapping experiments were performed in biological duplicate using a deep mutational scanning approach as previously described^{362,431}. Briefly, yeast-surface display libraries expressing 3,804 of the 3,819 possible amino acid mutations in the SARS-CoV-2 RBD (Wuhan-Hu-1 sequence, Genbank MN908947, residues N331-T531) were constructed in³⁸⁹, and sorted to purge mutations that abolish RBD folding or ACE2 binding. Libraries were induced for RBD surface expression and labeled with minibinder at the concentration of the EC90 for binding to yeast-displayed RBD (“sensitive” selection) or 400 ng/mL (“stringent” selection), the concentration used for prior selections of clinical antibodies^{362,432}, followed by secondary labeling with 1:100 FITC-conjugated anti-Myc (Immunology Consultants Lab, CMYC-45F), and 1:100 iFluor-647-conjugated mouse anti-His tag antibody (Genscript A01802). Library variants that escape miniprotein binding were sorted on a BD FACSAria II based on gates drawn from control populations labeled at 0.1x (“sensitive” selection) or 0.01x (“stringent” selection) the

selection concentration, as shown in Fig. 9.1 A-C. For each sample, >10 million RBD⁺ cells were processed, minibinder-escape cells were sorted and grown overnight, plasmid purified, and sequenced on a HiSeq 2500. Escape fractions were computed from sequencing counts exactly as described in Starr et al.³⁶². Illumina sequencing counts are available from the NCBI SRA (BioProject SAMN19925005). All code and analysis steps are described on GitHub: https://github.com/jbloombloom/SARS-CoV-2-RBD_MAP_minibinders. A table reporting all mutation escape fractions is available: https://github.com/jbloombloom/SARS-CoV-2-RBD_MAP_minibinders/blob/main/results/supp_data/IPD_ligands_raw_data.csv.

4.5.8 Cell-free DNA assembly and CFPS template preparation

Proteins to be manufactured via CFPS were codon optimized using the IDT codon optimization tool and ordered as gblocks containing the pJL1 5' (ttgtttaactttaagaaggagatatacat) and 3' (gtcgaccggctgctaacaagcccgaaagg) Gibson assembly overhangs. DNA was resuspended at a concentration of 50 ng/μL.

Linearized pJL1 plasmid backbone (Addgene plasmid # 69496) was ordered as a gblock from IDT (see below) and amplified using the pJL1_F (gtcgaccggctgcta) and pJL1_R (atgtatatctccttctaaagtaaacaaaattatttcta) primers via PCR using Q5 Hot Start DNA polymerase (NEB, M0493L) following manufacturer instructions. Amplified pJL1 backbone was purified using the DNA Clean and Concentrate Kit (Zymo Research, D4006) and diluted to a concentration of 50 ng/μL.

Linearized pJL1 plasmid backbone:

```
gtcgaccggctgctaacaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacccttggggcctctaaacg
ggctctgaggggttttctgtaaaagccaattctgattagaaaaactcatcgagcatcaaatgaaactgcaattattcatatcaggattatcaataccat
atctttgaaaaagccgtttctgtaaatgaaggagaaaactcaccgaggcagttccataggtggcaagatcctggatcggtctgagattccgactcgtc
caacatcaatacaacctattaatttcccctgtaaaaaataaggttatcaagtgagaaatcacatgagtgacgactgaatccggtgagaatggcaa
aagcttatgcatttctccagactgttcaacaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgc
ctgagcgagacgaaatacgcgatcgtgttaaaaggacaattacaaacaggaatcgaatgcaaccggcgaggaacactgccagcgcatcaa
caatattttcacctgaatcaggatatttcttaatacctggaatgctgtttcccggggatcgagtggtgagtaaccatgcatcatcaggagtacggata
aaatgcttgatggtcggaagaggcataaattccgctagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttccatgtttcag
aaacaactctggcgcatgggctccatacaatcgatagattgtcgcacctgattgccgacattatcgagcccatttatacccatataaatcagc
```

atccatgttgaatttaatcgcggcttcgagcaagacggttcccgttgaatatggctcataacaccccctgtattactgtttatgtaagcagacagttttattg
 ttcgatgatataatcttctgtgcaatgtaacatcagagattttgagacacaacgtgagatcaaaggatcttcttgagatcctttttctgcgcgtaatct
 gctgcttgcacaacaaaaaacaccgctaccagcgggtgtttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcaga
 gcgagatacacaataactgttcttagttagccgtagttaggcccaccactcaagaactctgtagcaccgctacatacctcgctctgtaaatcctgtt
 accagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctgggctgaacgg
 ggggttcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaaagcggccagcttcccga
 agggagaaaaggcggacaggtatccggaagcggcagggctggaacaggagagcgcacgagggagctccagggggaaaacgctggtatctt
 tatagtctgtcgggttcgccacctctgacttgagcgtcgattttgtgatgctcgtcagggggcgaggcctatggaaaaacgccagcaacgcgatc
 ccgcaaattaatcagactcactataggagaccacaacggttccctctagaataatttgtttaactttaagaaggagatatacat

Gibson assembly was used to assemble protein open reading frame DNA with the pJL1 backbone following the published protocol with the addition 3.125 µg/mL of ET SSB (NEB, product no. M2401S)^{357,358}. 20 ng of purified, linearized pJL1 backbone and 20 ng of the protein open reading frame insert were combined in 2 µL Gibson assembly reactions and incubated at 50°C for 30 minutes. Similar to other published methods^{53,55}, the unpurified assembly reactions were diluted in 40 µL of nuclease free water (Fisher Scientific, AM9937) and 1 µL of the diluted reaction were used as the template for a PCR to generate linear expression templates (LETs) for CFPS. Linear expression templates were amplified via PCR using the pJL1_LET_F (ctgagatacctacagcgtgagc) and pJL1_LET_R (cgctactcatgggtgatttctcacttg) primers and the Q5 Hot Start DNA polymerase (NEB, M0493L) following manufacturer instructions.

4.5.9 CFPS cell extract preparation

Cell extracts for CFPS reactions were prepared from BL21 StarTM (DE3) (Thermo Fisher, C601003) as previously described³³.

4.5.10 CFPS reactions and purification

CFPS reactions were assembled as previously described except for the DNA template³³. Unpurified linear expression templates in PCR reaction buffer were added to the CFPS reaction at 6.66% v/v to drive protein synthesis. CFPS reactions were run in 24 well plates at 30°C for 14-20 h. 100-200 µL volume reactions, to produce protein for AlphaLISA off-rate screening, were run in 24-well plates (Falcon, 351147) and purified via StrepTactinXT spin columns (IBA, 2-4151-000) following manufacturer

instructions. Post purification, proteins were extensively dialyzed (Fisher Scientific, 69552) into 50 mM HEPES pH 7.4, 150 mM NaCl.

2 mL reactions, to produce protein for ELISA, neutralization, and VSV escape studies, were run in 6-well plates (Costar, 3736) and purified via gravity flow using StrepTactinXT Superflow high capacity resin (IBA, 2-4030-010). CFPS reactions were incubated with resin on an end-over-end rotator for 30 min at room temperature. The resin was spun down at 2,500 x *g* for 2 min and the supernatant was removed. Resin was resuspended in Buffer W (IBA, 2-1003-100) and loaded onto a gravity flow column (BioRad, 7321010). Resin was washed with 20X resin volumes of Buffer W and proteins were eluted via the addition of Buffer BXT (IBA, 2-1042-025). Following purification, EDTA and CHAPS were added to a concentration of 10 mM and 4 mg/mL, respectively to aid in the removal of endotoxin by dialysis⁴³³. Proteins were dialyzed (Fisher Scientific, 87724) into 1L of endotoxin-free 1x PBS (Fisher Scientific, SH30256LS) in a 4 L glass beaker. Glass beakers were baked at 240°C for 12 h to degrade endotoxin prior to use. Samples were dialyzed at 4°C for > 6 h. At least two additional rounds of dialysis with the addition of EDTA and CHAPS to concentrations of 10 mM and 4 mg/mL, respectively, were performed to remove endotoxin. Endotoxin was quantified using the Pierce Chromogenic Quant Kit (Fisher Scientific, A39553) and samples with less than 10 EU/mL were used for cell-based assays.

4.5.11 Expression and purification of competitor proteins for AlphaLISA experiments

Proteins used as competitors in the AlphaLISA off-rate screen were cloned into pET28a, transformed into BL21 Star™ DE3 (Thermo Fisher, C601003), plated on LB agar, and cultured overnight at 37°C. 1 L of Overnight Express TB (Fisher Scientific, 71491-4) was inoculated by scraping all colonies on a transformation plate and cultured at 37°C in 2.5 L tunair flasks (IBI Scientific, SS-8003) at 220 rpm overnight. Cells were harvested, resuspended at a ratio of 1 g cell mass to 4 mL resuspension buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 1X HALT protease inhibitor without EDTA (Fisher Scientific, 78429), 1 mg/mL lysozyme, 62.5 U/mL cell suspension of benzonase (Sigma Aldrich, E1014-25KU)) and lysed using an Avestin B15 homogenizer at 21,000 PSI. Lysate was spun down 14,000 x *g* for 10 min and the clarified supernatant was incubated with Ni-NTA Agarose (Qiagen, 30230) for 60 min on an end-over-end shaker. Resin was spun down 2,500 x *g* for 2 min, supernatant removed, resuspended in wash buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 50 mM Imidazole), loaded on a gravity flow column, and subsequently

washed with 20X resin volumes of wash buffer. Protein was eluted using elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM Imidazole) and exchanged into 50 mM HEPES pH 7.4, 150 mM NaCl using PD-10 desalting columns (Cytiva, 17-0851-01).

His tags were removed via cleavage by ProTEV Plus (Promega, V6102). Prior to cleavage, 10% v/v glycerol was added to the protein. ProTEV Plus was added to a concentration of 0.5 U/ μ g purified protein and DTT was added to a concentration of 1 mM. Cleavage reactions were carried out at 30°C for 4 h. Free his tag and ProTEV Plus were removed by incubating with Ni-NTA Agarose for 1 h at 4°C and collecting the supernatant. Proteins were subsequently concentrated to > 1mg/mL (Millipore, UFC800396). His tag removal was validated via SDS PAGE and the AlphaScreen Histidine (Nickel Chelate) Detection Kit (Perkin Elmer, 6760619C).

4.5.12 Expression and purification of SARS-CoV-2 S proteins

S2P^{361,399,434,435}, S6P³⁹¹, and S6P variants were produced in Expi293F cells (Thermo Fisher, A14527) grown in suspension using Expi293F expression medium (Life Technologies) at 33°C, 70% humidity, 8% CO₂ rotating at 150 rpm. The cultures were transfected using PEI-MAX (Polyscience) with cells grown to a density of 3.0 million cells per mL and cultivated for 3 days. Supernatants were clarified by centrifugation (5 min at 4000 rcf), addition of PDADMAC solution to a final concentration of 0.0375% (Sigma Aldrich, #409014), and a second spin (5 min at 4000 rcf).

Proteins were purified from clarified supernatants via a batch bind method where each clarified supernatant was supplemented with 1 M Tris-HCl pH 8.0 to a final concentration of 45 mM and 5 M NaCl to a final concentration of ~310 mM. Talon cobalt affinity resin (Takara) was added to the treated supernatants and allowed to incubate for 15 min with gentle shaking. Resin was collected using vacuum filtration with a 0.2 μ m filter and transferred to a gravity column. The resin was washed with 20 mM Tris pH 8.0, 300 mM NaCl, and the protein was eluted with 3 column volumes of 20 mM Tris pH 8.0, 300 mM NaCl, 300 mM imidazole. The batch bind process was then repeated and the first and second elutions combined. SDS-PAGE was used to assess purity. IMAC elutions were concentrated to ~1 mg/mL and dialyzed three times into 50 mM Tris pH 8, 150 mM NaCl, 0.25% L-Histidine in a hydrated 10K molecular weight cutoff dialysis cassette (Thermo Scientific). Due to inherent instability, S2P was immediately flash frozen and stored at -80°C.

4.5.13 Competition-based off-rate screening via AlphaLISA

AlphaLISA reactions were carried out in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mg/mL BSA, and 0.015% v/v TritonX-100 (hereafter referred to as Alpha buffer). All components were dispensed using an Echo 525 liquid handler from a 384-Well Polypropylene 2.0 Plus microplate (Labcyte, PPL-0200) using the 384PP_Plus_GPSA fluid type. All AlphaLISA reactions were performed in a ProxiPlate-384 Plus (Perkin Elmer, 6008280). AlphaLISA StrepTactin donor beads, to capture StrepII or TwinStrep-tagged minibinder variants, (Perkin Elmer, AS106) and AlphaLISA Anti-6x-his, to capture 6xhis-tagged S6P or RBD, (Perkin Elmer, AL178C) were combined to prepare a 4X stock in Alpha buffer immediately prior to use and added to the proteins to yield a concentration of 0.08 mg/mL donor beads and 0.02 mg/mL acceptor beads in the final reaction.

Multivalency screening experiments were carried out at a final concentration of 2.5 nM S6P, 2.5 nM minibinder variant, and 250 nM of the specified untagged competitor in Alpha buffer. First, the minibinder variant and S6P were diluted to 1.33x final concentration (adjusted for the later addition of competitor) in 120 μ L in a 384-Well Polypropylene 2.0 Plus plate and sealed (BioRad, MSB1001). Samples were allowed to fully associate for 12-16 h at 20°C. Next, conditions were split in half and the same volume of either buffer or 250 nM (100x molar excess) untagged competitor were added to achieve 1.33x final concentration of all components (competitor was previously concentrated to achieve less than 5% volume change at this step). Samples were then incubated for the specified time, with replicates measured by dispensing 1.5 μ L of each condition and 0.5 μ L of 4X Alpha bead stock via the Echo 525 liquid handler. Plates were immediately spun down following the dispense and sealed (BioRad, MSB1001). Reactions were incubated with beads for 1 h for 2 h dissociation time points and up to 2 h for longer dissociation time points before measurement (bead incubation time was included in the specified timepoints). AlphaLISA measurements were taken on a Tecan Infinite M1000 Pro using the AlphaLISA filter with an excitation time of 100 ms, an integration time of 300 ms and a settle time 20 ms. Prior to measurement, plates were allowed to equilibrate inside the instrument for 10 min. Fraction of protein bound was determined by subtracting the average background bead signal and then dividing the plus competitor by the minus competitor condition. Values below zero after background subtraction were set to zero. Conditions with signal in the no competitor condition within 3 standard deviations of the background were set to zero. After normalization, conditions

with values less than 0.01 were set to zero to account for the typical max dynamic range of the measurement. Prism 9 (GraphPad) was used to plot the data.

Monovalent optimization experiments were performed in the same manner as multivalency screening experiments. S6P and RBD comparison experiments were carried out in the same manner as the multivalency screening experiments except for concentrations of 5 nM of S6P or RBD (Sino Biological, 40592-V08H), 5 nM minibinder variant, and 500 nM (100x molar excess) of the specified untagged competitor in the final reactions. In S6P variant experiments, reactions were carried out in the same manner as the multivalency screening experiments.

4.5.14 Negative stain electron microscopy

The SARS-CoV-2 HexaPro Spike protein (S6P) was produced in HEK293F cells grown in suspension using FreeStyle 293 expression medium (Life technologies) at 37°C in a humidified 8% CO₂ incubator rotating at 130 rpm. The cultures were transfected using PEI (9 µg/mL) with cells grown to a density of 2.5 million cells per mL and cultivated for 3 d. The supernatants were harvested and cells resuspended for another 3 d, yielding two harvests. Spike proteins were purified from clarified supernatants using a 5 mL Cobalt affinity column (Cytiva, HiTrap TALON crude), concentrated and flash frozen in a buffer containing 20 mM Tris pH 8.0 and 150 mM NaCl prior to analysis.

10 µM SARS-CoV-2 Spike was incubated with 13 µM minibinders for 1 h at room temperature. Samples were diluted to 0.01 mg/mL immediately prior to adsorption to glow-discharged carbon-coated copper grids for ~30 sec prior to a 2% uranyl formate staining. Micrographs were recorded using the Leginon on a 120 KV FEI Tecnai G2 Spirit with a Gatan Ultrascan 4000 4k x 4k CCD camera at 67,000 nominal magnification. The defocus ranged from -1.0 to -2.0 µm and the pixel size was 1.6 Å.

4.5.15 Cryo-electron microscopy

SARS-CoV-2 HexaPro S (S6P)³⁹¹ at 1.2 mg/mL was incubated with 1.2 fold molar excess of recombinantly purified TRI2-2, FUS31-G10, or FUS231-P24 at 4°C before application onto a freshly glow discharged 2.0/2.0 UltraAuFoil grid (200 mesh)⁴³⁶. Plunge freezing used a vitrobot MarkIV (ThermoFisher Scientific) using a blot force of 0 and 6.5 second blot time at 100% humidity and 23°C.

For the S6P/TRI2-2 data set, data were acquired using an FEI Titan Krios transmission electron microscope operated at 300 kV and equipped with a Gatan K3 direct detector and Gatan Quantum GIF

energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon⁴³⁷ at a nominal magnification of 105,000x with a pixel size of 0.4215 Å. The dose rate was adjusted to 15 counts/pixel/s, and each movie was acquired in super-resolution mode fractionated in 75 frames of 40 ms. 5,991 micrographs were collected with a defocus range between -0.5 and -2.5 µm. Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking, and extraction were carried out using Warp⁴³⁸ (Fig S7 and S8).

For the S6P/FUS31-G10 data set, data were acquired on an FEI Titan Krios transmission electron microscope operated at 300 kV equipped with a Gatan K2 Summit direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon⁴³⁷ at a nominal magnification of 130,000x with a pixel size of 0.525 Å. The dose rate was adjusted to 8 counts/pixel/s, and each movie was acquired in counting mode fractionated in 50 frames of 200 ms. 1000 micrographs were collected in a single session with a defocus range between -0.5 and -2.5 µm (Fig S9).

For the S6P/FUS231-P24 data set, data were acquired on an FEI Glacios transmission electron microscope operated at 200 kV equipped with a Gatan K2 Summit direct detector. Automated data collection was carried out using Leginon⁴³⁷ at a nominal magnification of 36,000x with a pixel size of 1.16 Å. The dose rate was adjusted to 8 counts/pixel/s, and each movie was acquired in counting mode fractionated in 50 frames of 200 ms. 1,663 micrographs were collected in a single session with a defocus range between -0.5 and -2.5 µm (Fig S10).

For the S6P/TRI2-2, S6P/FUS31-G10, and S6P/FUS231-P24 datasets, two rounds of reference-free 2D classification were performed using CryoSPARC⁴³⁹ to select well-defined particle images. These selected particles were subjected to two rounds of 3D classification with 50 iterations each (angular sampling 7.5° for 25 iterations and 1.8° with local search for 25 iterations), using our previously reported closed SARS-CoV-2 S structure as initial model (PDB 6VXX) in Relion⁴⁴⁰. 3D refinements were carried out using non-uniform refinement⁴⁴¹ along with per-particle defocus refinement in CryoSPARC. Selected particle images were subjected to the Bayesian polishing procedure implemented in Relion3.0⁴⁴² before performing another round of non-uniform refinement in CryoSPARC followed by per-particle defocus refinement and again non-uniform refinement. To further improve the density of the TRI2-2, the particles

were then subjected to focus 3D classification without refining angles and shifts using a soft mask comprising the RBD and TRI2-2 region with a tau value of 60 in Relion. Particles belonging to classes with the best resolved local density were selected and subject to local refinement using CryoSPARC. Local resolution estimation, filtering, and sharpening were carried out using CryoSPARC. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution^{443,444}. UCSF Chimera⁴⁴⁵ and Coot⁴⁴⁶ were used to fit atomic models into the cryoEM maps. Spike-RBD/TRI2-2 model was refined and relaxed using Rosetta^{447,448} using sharpened and unsharpened maps (Table 9.2, Fig 9.8, 9.9, and 9.10).

4.5.16 BRET sensor for SARS-CoV-2 S detection

Synergy Neo2 plate reader (Biotek) was used for all luminescent assays. 10 μ L of 1 nM teluc-FUS231-P12-mCyRFP3, 10 μ L of serial diluted 10X spike protein (final concentrations range from 1.85 nM to 0.85 μ M), and 30 μ L of buffer (25 mM Tris, pH8, 50mM NaCl) were mixed and incubated for 10 min at room temperature. Diphenylterazine stock solution was prepared as previously described⁴⁰⁰. Then, 50 μ L of diluted diphenylterazine solution (60 μ M) was added to each well. The luminescence spectra were collected under the monochromator mode with 0.1 s integration and 5 nm increments from 400 to 750 nm. The luminescent image was taken by an iPhone8 camera. To record the emission ratio, luminescence signals were acquired by a filter mode with a two channel cube (470/40 nm and 590/35 nm). The emission ratios were calculated from 590 nm/470 nm channels directly. The linear region of ratiometric responses was extracted and a linear regression curve was plotted, which was used to derive the standard deviation (s.d.) of the response and the slope of the calibration curve (S). The limit of detection was determined as 3 s.d. above background signal. The limit of detection of the sensor was compared with the performance of commercial lateral flow n-protein antigen tests in Corman *et al.*⁴⁴⁹ on purified n-protein assuming a 10:1 molar ratio of n-protein to S trimer per virion⁴⁵⁰.

BRET sensor protein sequence [His-TEV][teluc][FUS231-P12][mCyRFP3]:

[MGSHHHHHHGSSENLYFQGSG][VFTLEDFVGDWRQTAGYNLSQVLEQGGVSSLFQNLGVSVTPIQRI
VLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDNHHFKVILHYGTLVIDGVTPNMIDYFGRPYEGI
AVFDGKKITVTGLWNGNKIIDERLINPDGSLFRVTVINGVTGWRLHERILASGSSG][ELEEQVMHVLVDQV

SELAHELLHKL TGEELERAAYFNWWATEMMLELIKSDDEREIREIEEEAARILEHLEELARTGGASPAAPA
 PGGNLDELHMQMTDLVYEALHFAKDEEFQKHVFQLFEKATKAYKNKDRQKLEKVVEELKELLERLLSGG
 ASPAAPAPGGDKENVLQKIYEIMKELERLGHAEASMQVSDLIYEFMKTCDENLLEEAERLLEEVRK][GEEL
 IKENMRSKLYLEGSVNGHQFKCTHEGEGKPYEGKQTARIKVVVEGGPLPFAFDILATMFMYGSKVFIKYP
 DLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTSLQDGELIYNVKL RGVNFPANGPVMQKKT LGWEP
 STETMTPADGGLEGRCDKVLKLVGGGHLHVNFKTTYKSKKPKMPGVHYVDRRLERIKEADNETYVEQY
 EHAVARYSNLGGGMDELYK]

BRET Sensor DNA sequence:

atgggcagccatcatcatcaccatcatggtagcggcagcgaaaactgtatcttcaggggagcggcgtgttaccctggaagatcttggggcagattg
 cgccagaccgcggtataacctgagccaggtctggaacaggggtgtgagcagcctgttcagaatctggcgtagcgttaccatccag
 cgattgtactgtctggtgaaaacggcctgaaaattgatccatgtgattattccgtatgaaggcctgagcggcagatcagatgggcccagattgaaaa
 gatctttaaagtggtatccggtggataaccatcattcaaagtattctgattatggaccctggtattgatggggttacgccgaacatgattgattat
 ttggccgcccgtacgaaggcattgcccgtttgatggcaaaaagattaccgtgaccggcagcgtgtggaacggtaacaaaattattgacgaacgc
 ctgattaaccggatggctctctgctgtttcgctgaccattaatggcgtcaccggctggcgtctgcacgaacgtatcttggcaagcggtagcagtggt
 gaactggaagaacaagtgtatgctgtggaactgagcgaactggccatgagctgtgcataaactgaccggtaagaactgaacgcg
 cggcgtatttaactggtgggcaaccgaaatgatgctggaactgatcaaaagcagatgatgaacgtgaaattcgcgaaattgaagaagaggcggcc
 cgtattctgaacatctggaagaattggcgagaacaggtggtctcactgctgctcctgcgcaggtggttaatttagatgaattgcacatgcagatg
 accgatctggttatgaagcgtgcatcttccaaagatgaagaattcagaacatgtttccagctgtttgaaaaagcgacgaaagcgtacaaaaa
 caaagaccgcccagaaattgaaaaagttgtcgaagaactgaaagaattactggaacgtctgttgagcgggtggagcaagcccagctgcaccagct
 cctggtggtgacaaagaaaacgtcctccagaaaattacgaaattatgaaagagctggaaaggctggccatgctgaagcagatgcaggttag
 cgatctgatctatgaattcatgaaaaccaagacgaaaacctgttgaagaagccgaacgtttgttagaagaagtgaacgcggcgaagaattaa
 ttaagaaaacatgcgtagcaactgtatttgaaggcagcgtgaacggccatcagttaaatgcacccatgaagggaaggcaaaccttacgaa
 ggtaaacagacggcgcattaaagtcgtggaagggtggcccgtcggttgctttgatattctggcgaccatgtttatgtatggcagcaaaagtgtta
 ttaaatatccggcggatctgccgattatttcaaacagagctttccggaaggctttacctgggaacgcgtgatggttccaagatggcgggtgtgctga
 ccgagaccagataaccagcctgcaagatggggaactcattataacgttaaactgcgggtgtaacttccggcgaacggctccggtatgcaga
 agaaaacctgggtgggaaccgagcaccgaaacctgacccggcagatggtggttagaaggccgctgcgataaagtctgaaactggttg
 cggcggccacctgcatgcaactttaaaccacgtataaaagcaaaaagccggttaaatgcccggcgtgcactatgtggatgcggctctggaac

gcatcaaagaagcggataacgaaacctatggtgaacagtatgaacacgccgtggcgcttatagcaacctcgggtggcatggatgaactgtac
aaataa

4.5.17 Minibinder/RBD interaction kinetics via bio-layer interferometry (BLI)

The Octet® HTX instrument was used to determine affinities of minibinders to monovalent soluble RBD (sol-RBD) (obtained from the Institute of Protein Design at University of Washington). Prior to sol-RBD affinity measurements, Amine Reactive 2nd Generation (AR2G) biosensors (Sartorius, #18-5094) were pre-hydrated for 10 min in water before being activated using EDC/NHS chemistry (solution containing 20 mM EDC and 10 mM NHS in ddH₂O) for 5 min. Sol-RBD in 10 mM Acetate buffer pH 5.0 was reacted with activated AR2G biosensors for 10 min and quenched for 5 min in 1M ethanolamine pH 8.5. Sol-RBD linked AR2G biosensors were equilibrated for 60 s in running buffer (10 mM Tris, 150 mM NaCl, 1 mM CaCl₂, 0.1 mg/mL BSA, and 0.1% Triton-X 100, pH7.4) before measuring association of minibinders at concentrations ranging from 20 nM to 0.25 nM for 10 min, followed by 20 min of dissociation in running buffer. Minibinders were prepared in the running buffer. For data evaluation, the ForteBio Data Analysis v11.0 software was used. The kinetic rate constants, association rate constant (k_a , M⁻¹s⁻¹), dissociation rate constant (k_d , s⁻¹), and the equilibrium rate constant (K_D , M) were determined by using a 1:1 Langmuir model. Following standard best practice, 5% dissociation during the experiment was used as a threshold to determine the lower limit of detection for the dissociation rate constant⁴⁵¹.

4.5.18 Multivalent Minibinder/S6P variant interaction kinetics via surface plasmon resonance (SPR)

The Catterra LSA instrument was used to perform HT-SPR of minibinder affinity to Spike protein variants. To prepare the surfaces, the Single Flow Channel (SFC) and 96-Print Head (96PH) were primed with running buffer (HBS-T; 50 mM HEPES pH 7.5, 150 mM NaCl, 0.1% Tween 20). The capture surface was prepared in the 96PH by standard amine-coupling. A HC30-M chip (Catterra LSA cat# 4279) was activated with a 10-min injection of freshly prepared 1:1:1 (v/v/v) mixture of 0.4 M EDC + 0.1 M NHS + 0.1 M MES pH 5.5. Spike protein trimers were diluted to 12.5 µg/ml in 10 mM sodium acetate pH 4.5 (Catterra, #3628) and coupled for 20 min. Excess reactive esters were blocked with a 7-min injection of 1 M ethanolamine HCl pH 8.5 (Catterra, #3626). Final coupling amounts were greater than 1000RU. Minibinders were prepared in HBS-T buffer, at a three-fold dilution series for 6 points starting at 20nM. Association was for 20 minutes, with a 60-minute dissociation time. Samples were injected in ascending

concentration without any regeneration. The data was double referenced in that both a local reference and a zero nanomolar analyte concentration (e.g. buffer) were subtracted. The double-referenced data were fit globally to a 1:1 Langmuir binding model in Carterra's Kinetic tool, allowing each spot its own k_a and k_d value to determine KD. Following standard best practice, 5% dissociation during the experiment was used as a threshold to determine the lower limit of detection for the dissociation rate constant ⁴⁵¹.

4.5.19 Competition ELISA of minibinders and SARS-CoV-2 S6P variants for immobilized hACE2-Fc

A 0.003 mg/mL solution of hACE2-Fc in 20mM Tris pH 8 and 100mM NaCl was immobilized to a Maxisorp 384-well plate (Thermo Scientific 464718) overnight at 4°C. Plates were slapped dry and blocked with Blocker Casein in TBS (Thermo Scientific 37532) for one hour at 37°C. 20nM mini binders were serially diluted 1:3 in avi-tagged prefusion-stabilized SARS-CoV-2 S6P ³⁹¹ variants at their EC50 concentrations of 1.2nM for WT, 6.3nM for E406W, 1.7nM for K417N, 0.5nM for Y453F, 0.6nM for Y453R, 1.3nM for L455F, 1.1nM for F456L, 0.1nM for E484K, 0.2nM for N501Y, 0.3nM for B.1.1.7, 0.2nM for B.1.351, or 0.2nM for P.1 and incubated for 30 min on a non-binding plate (Greiner 781901) at 37°C. The plates with blocking buffer were slapped dry and the pre-incubated mini binders and spikes were added. Plates were incubated for 1 h at 37°C then washed 4x with TBST using a 405 TS Microplate Washer (BioTek) followed by addition of 30 µL avi-tag pAb (GenScript A00674) at 0.2 µg/mL. Plates were incubated for 1 h at 37°C then washed 4x with TBST using a 405 TS Microplate Washer (BioTek) followed by addition of 30 µL 1:2000 Goat anti-Rabbit HRP (Invitrogen 656120) and a 1 h incubation at 37°C. Plates were washed 4x and TMB Microwell Peroxidase (Seracare 5120-0083) was added. The reaction was quenched after 2-3 min with 1 N HCl and the A450 of each well was read using a BioTek plate reader (BioTek). Data were plotted and fit in Prism (GraphPad) using nonlinear regression sigmoidal, 4PL, X is log(concentration) to determine IC50 values from curve fits. IC50 values less than two-fold below the concentration of S6P in that condition were not considered different from two-fold below the concentration of S6P.

4.5.20 Pseudovirus production

HIV-based pseudotypes were prepared as previously described ⁴²¹. Briefly, HEK293T cells were co-transfected using Lipofectamine 2000 (Life Technologies) with an S-encoding plasmid with full suite of variant mutations, an HIV Gag-Pol, Tat, Rev1B packaging construct, and the HIV transfer vector encoding a luciferase reporter according to the manufacturer's instructions. Cells were washed 3x with Opti-MEM

and incubated for 5 h at 37°C with transfection medium. DMEM containing 10% FBS was added for 60 h. The supernatants were harvested by spinning at 2,500 x *g*, filtered through a 0.45 µm filter, concentrated with a 100 kDa membrane for 10 min at 2,500 x *g* and then aliquoted and stored at -80°C.

4.5.21 Pseudovirus neutralization

HEK-hACE2 cells were cultured in DMEM with 10% FBS (Hyclone) and 1% PenStrep with 8% CO₂ in a 37°C incubator (ThermoFisher). One day prior to infection, 40 µL of poly-lysine (Sigma) was placed into 96-well plates and incubated with rotation for 5 min. Poly-lysine was removed, plates were dried for 5 min then washed 1x with water prior to plating with 40,000 cells. The following day, cells were checked to be at 80% confluence. In an 80 µL final volume, minibinders were serially diluted in DMEM 1:3 starting at 100 nM. Pseudovirus was added 1:1 to the diluted minibinders and allowed to incubate for 30-60 min at room temperature. After incubation, the mixture of minibinder and virus was added to the cells at 37°C and allowed to incubate for 2 h. Post infection, 160 µL of 20% FBS-2% PenStrep DMEM was added. After 48 h, 40 µL/well of One-Glo-EX substrate (Promega) was added to the cells and incubated in the dark for 5-10 min prior reading on a BioTek plate reader. Measurements were done in at least duplicate. Relative luciferase units were plotted and normalized in Prism (GraphPad). Nonlinear regression of log(inhibitor) versus normalized response was used to determine IC₅₀ values from curve fits.

4.5.21 SARS-CoV-2 neutralization

Serial dilutions of minibinders were incubated with 10² FFU of SARS-CoV-2 for 1 h at 37°C. Binder-virus complexes were added to Vero-hACE2-TMPRSS2 cell monolayers in 96-well plates and incubated at 37°C for 1 h. Subsequently, cells were overlaid with 1% (w/v) methylcellulose in MEM supplemented with 2% FBS. Plates were harvested 20-24 h later by removing overlays and fixed with 4% PFA in PBS for 20 min at room temperature. Plates were washed and sequentially incubated with an oligoclonal pool of SARS2-2, SARS2-11, SARS2-16, SARS2-31, SARS2-38, SARS2-57, and SARS2-71 anti-spike protein antibodies³⁷⁵ and HRP-conjugated goat anti-mouse IgG in PBS supplemented with 0.1% saponin and 0.1% bovine serum albumin. SARS-CoV-2-infected cell foci were visualized using TrueBlue peroxidase substrate (KPL) and quantitated on an ImmunoSpot microanalyzer (Cellular Technologies). Data were processed using Prism software (GraphPad Prism 8.0).

4.5.22 Kidney organoid differentiation and infection with SARS-CoV-2

Kidney organoids were differentiated from H9 human embryonic stem cells (WiCell, WA09) for 21 days prior to infection in adherent, thin-layer Matrigel sandwich cultures induced for 36 hours with CHIR99021 as described previously⁴⁰⁴. SARS-CoV-2 strain B.1.351-HV001 containing E484K/N501Y/D614G mutations along with furin cleavage site point mutation were obtained directly from B.1.351 clinical isolates. All experiments using live viruses were performed at Biosafety Level 3 (BSL-3) facilities at the University of Washington in compliance with BSL-3 laboratory safety protocols (CDC BMBL 5th ed.) and the recent CDC guidelines for handling SARS-CoV-2. Virus stocks were generated and titrated using plaque forming assays in Vero cells (USAMRIID). Minibinders FUS231-10GS, TRI2-2, and MON1 at 0.3 μ M were diluted in serum-free DMEM and pre-incubated with virus (10 multiplicity of infection) for 1 h at 37°C. The virus minibinder mix was then added to kidney organoids for 1 h at 37°C. After 1 h, kidney organoids were washed with 1X PBS and fresh organoid growth media (Advanced RPMI + 1X Glutamax + 1X B27 Supplement, Thermo Fisher Scientific) was added and incubated for 72 h. Supernatants were harvested from the infected organoids for plaque forming assays and cells were lysed using Trizol RNA extract for gene expression analysis.

4.5.23 Kidney organoid plaque assay

Vero cells were plated at 80% confluency, washed with 1X PBS, and incubated with serially diluted supernatant from infected organoids for 1 h at 37°C. Cells were then overlaid with a 1:1 mixture of 1.8% cellulose in water:2X DMEM supplemented with 4% heat-inactivated FBS, L- glutamine, 1X antibiotic-antimycotic (Gibco), and 220 mg/mL sodium pyruvate was layered on top of the cells and incubated for 48 h at 37°C. Cells were then fixed using 10% formaldehyde and incubated for 30 min at room temperature. Overlay was removed carefully and cells were stained with 0.5% crystal violet solution in 20% ethanol. Plaques were counted and the virus titer in the original sample was calculated as plaque-forming units per mL (PFU/mL). Data was analyzed using GraphPad Prism (8.0). Limit of detection for the plaque assay was determined by calculating the concentration (PFU/mL) corresponding to a single observed plaque. Conditions where zero plaques were observed were set to the value of the limit of detection.

4.5.24 Kidney organoid gene expression analysis

RNA from infected kidney organoids was harvested using Trizol reagent. cDNA was generated using iScript™ cDNA Synthesis Kit (catalog #1708890, Bio-rad). qRT-PCR was performed using SYBR green master mix (catalog #4309155, Applied biosystems) and the following primers: SARS-CoV2-E F (GAACCGACGACGACTACTAGC), SARS-CoV2-E R (ATTGCAGCAGTACGCACACA), β -ACTIN F (GCAAAGACCTGTACGCCAACA), β -ACTIN R (ACACGGAGTACTTGCGCTCAG).

4.5.25 Selection of escape mutants in SARS-CoV-2 S using VSV-SARS-CoV-2 chimera

VSV-SARS-CoV-2 chimera (S from Wuhan-Hu-1 D614G strain) was used to select for SARS-CoV-2 S minibinder resistant mutants as described previously for selection of SARS-CoV-2 mAb escape mutants^{324,327}. Briefly, VSV-SARS-CoV-2 chimera virus was allowed to infect Vero cells for 1 hour at 37 °C. For the no inhibitor control, approximately 40 PFU of VSV-SARS-CoV-2 chimera were added per well. For the inhibitor containing conditions, 10⁶ PFU of VSV-SARS-CoV-2 chimera were added per well. Following infection, inhibitor protein (minibinder or mAb) was added to the culture medium and incubated for 72 hours to allow for the replication and spreading of resistant viruses. Neutralization resistant mutants were recovered by plaque isolation on Vero cells with the indicated minibinder in the overlay. The concentration of minibinder in the overlay was determined by neutralization assays at a multiplicity of infection (MOI) of 100. The concentration of 2B04, FUS231-P12 and TRI2-2 added in the overlay was 102 ng/ml, 58 ng/ml and 58 ng/ml respectively. Escape clones were plaque-purified on Vero cells in the presence of peptide binder or monoclonal antibody 2B04 respectively, and plaques in agarose plugs were amplified on MA104 cells with the peptide binder present in the medium. Viral stocks were amplified on MA104 cells at an MOI of 0.01 in Medium 199 containing 2% FBS and 20 mM HEPES pH 7.7 (Millipore Sigma) at 34°C. Viral supernatants were harvested upon extensive cytopathic effect and clarified of cell debris by centrifugation at 1,000 x g for 5 min. Aliquots were maintained at -80°C. Viral RNA was extracted from VSV-SARS-CoV-2 mutant viruses using an RNeasy Mini kit (Qiagen), and S was amplified using OneStep RT-PCR Kit (Qiagen). The mutations present in 2B04 escape variants were identified by Sanger sequencing (GENEWIZ).

To determine the diversity present in the pool of viral particles used for the escape selection, the Pedel-AA⁴⁵² web tool for predicting protein diversity in randomized libraries was used

(<http://guinevere.otago.ac.nz/aef/STATS/index.html>). Upper (10^{-3} errors per nucleotide) and lower (10^{-4} errors per nucleotide) estimates of the VSV RNA polymerase L error rate⁴⁰⁸ were evaluated separately to provide a plausible range of library diversities. Only the continuous amino acid sequence that contacts the minibinders (Genbank accession# MN908947.3, surface glycoprotein residues 400 to 508) within the SARS-CoV-2 RBD was considered. The library diversity was calculated using an equal probability of mutating from any nucleotide to any other, a zero probability of insertions and deletions, and a poisson distribution to estimate the mean number of mutations per sequence.

4.5.26 Protein production for animal studies

Protein was produced by fermentation in the E. coli BL21 pLysS (Thermo Fisher, C602003) strain using pET vectors induced with IPTG. The 6xHis tagged proteins were purified from clarified cell lysates by immobilized metal chelate chromatography (IMAC, Ni-NTA resin) and step eluted with 300 mM imidazole. Proteins were polished by size exclusion chromatography using an S75 Increase Column into a final buffer of 20mM NaPO₄ 150mM NaCl pH 7.4. Proteins analyzed by SDS-PAGE after heating 95°C without (-) reducing agent Dithiothreitol (DTT) followed by Coomassie blue staining (molecular weight standards shown). Protein endotoxin amounts are < 10 E.U./mg.

4.5.27 Mouse studies

Animal studies were carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocols were approved by the Institutional Animal Care and Use Committee at the Washington University School of Medicine (assurance number A3381-01). Virus inoculations and intranasal minibinder treatments were performed under anesthesia that was induced and maintained with ketamine hydrochloride and xylazine, and all efforts were made to minimize animal suffering.

Heterozygous K18-hACE2 C57BL/6J female mice (strain: 2B6.Cg-Tg(K18-ACE2)2PrImn/J) were obtained from The Jackson Laboratory. Animals were housed in groups and fed standard chow diets. For inoculation, 8-week-old mice were administered 10^3 PFU of the indicated SARS-CoV-2 strain in 50 μ L total volume via intranasal administration. Briefly, mice were anesthetized and the virus inocula was slowly pipetted into both nostrils of each mouse. Following administration, animals were placed on their back and observed for 5 min to ensure aspiration into the lungs. This procedure was performed identically for

administration of minibinders at D-1 or D+1 relative to virus inoculation. A previously described influenza minibinder was used as a negative control⁴¹⁸.

4.5.28 Measurement of viral burden in mouse studies

Tissues were weighed and homogenized with zirconia beads in a MagNA Lyser instrument (Roche Life Science) in 1 mL of DMEM media supplemented with 2% heat-inactivated FBS. Tissue homogenates were clarified by centrifugation at 10,000 rpm for 5 min and stored at -80°C .

To measure viral RNA, RNA was extracted using the MagMax mirVana Total RNA isolation kit (Thermo Scientific) on a Kingfisher Flex extraction robot (Thermo Scientific). RNA was reverse transcribed and amplified using the TaqMan RNA-to-CT 1-Step Kit (ThermoFisher). Reverse transcription was performed at 48°C for 15 min followed by 2 min at 95°C . Amplification was accomplished over 50 cycles as follows: 95°C for 15 s and 60°C for 1 min. Copies of SARS-CoV-2 *N* gene RNA in samples were determined using a previously published assay^{409,453}. Briefly, a TaqMan assay was designed to target a highly conserved region of the *N* gene (Forward primer: ATGCTGCAATCGTGCTACAA; Reverse primer: GACTGCCGCCTCTGCTC; Probe: /56-FAM/TCAAGGAAC/ZEN/AACATTGCCAA/3IABkFQ/). This region was included in an RNA standard to allow for copy number determination down to 10 copies per reaction. The reaction mixture contained final concentrations of primers and probe of 500 and 100 nM, respectively.

To measure viral titers, Vero-TMPRSS2-hACE2 cells were seeded at a density of 1×10^5 cells per well in 24-well tissue culture plates. The following day, medium was removed and replaced with 200 μL of material to be titrated diluted serially in DMEM supplemented with 2% FBS. One hour later, 1 mL of 1:1 DMEM:2% (w/v) methylcellulose overlay was added. Plates were incubated for 72 h, then fixed with 4% paraformaldehyde (final concentration) in PBS for 20 min. Plates were stained with 0.05% (w/v) crystal violet in 20% methanol and washed twice with distilled, deionized water.

4.5.29 Pharmacokinetics of TRI2-2 in mouse studies

For pharmacokinetic studies, 6 week-old C57BL/6J mice were administered TRI2-2 intranasally as described above for mouse protection studies. At the indicated time points post-administration, tissues were collected, weighed, and homogenized with zirconia beads in a MagNA Lyser instrument (Roche Life Science) in 1 mL of DMEM medium supplemented with 2% heat-inactivated FBS. Tissue homogenates

were clarified by centrifugation at 10,000 rpm for 5 min and stored at -80°C . To determine TRI2-2 binder concentrations in tissues, a competitive ELISA was performed in the same manner as previously reported³⁷⁸. Maxisorp 96-well plates (Thermo Scientific #44-2404-21) were coated with 4 ng/uL Hexapro in 1X TBS-T (20X TBS-T: Thermo Scientific #28360) for 1 h at 37°C . Plates were washed four times with 300 uL of 1X TBS-T before blocking with 200 uL SuperBlock blocking buffer (Thermo Scientific #37515) for 1 h at 37°C . Plates were washed four times with 300 uL of 1X TBS-T. 100 uL 2 nM hACE2-Fc in 1X TBS-T was added to all wells. 16.7 uL of 300 nM LCB1v1.3 was added to the first row of columns 1-3 as a standard, and 16.7 uL of tissue sample was added to the first row of other columns. All samples were processed in triplicate. Samples and standards were serially diluted in a 1:3 fashion and then incubated for 1 h at 37°C . Plates were washed four times with 300 uL of 1X TBS-T. 50 uL of 1:5000 anti-hFc-HRP secondary antibody (Invitrogen #A18817) was added to each well and incubated for 1 h at 37°C . Plates were washed four times with 300 uL of 1X TBS-T. 100 uL TMB (Thermo Scientific #34028) was added to each well, and plates were placed on an orbital shaker for 30 sec before addition of 100 uL of 1 N HCl to stop the reaction. Plates were read for absorbance at 450 nm on a SpectraMax Plus 384 plate reader. Data were analyzed by nonlinear regression in GraphPad Prism to obtain IC50 values. Sample TRI2-2 concentrations were determined by using the ratio of standard IC50/standard concentration to sample IC50/sample concentration.

4.5.30 Statistical analysis

Statistical significance was determined by a P value < 0.05 using the GraphPad Prism 9 software. Only non-parametric tests were used throughout this manuscript. Analysis of mouse weight changes was performed using a two-way ANOVA with Sidak's post-test for multiple comparisons. Statistical analysis of viral load between two groups was performed using either a Kruskal-Wallis test with Dunn's post-hoc analysis for multiple comparisons or a two-tailed Mann-Whitney test as noted in the corresponding figure captions.

Chapter 5 - Rapid and Sensitive Detection of Antigen from SARS-CoV-2 Variants of Concern by a Multivalent Minibinder-Functionalized Nanomechanical Sensor

5.1 Abstract

New platforms for the rapid and sensitive detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern are urgently needed. Here we report the development of a nanomechanical sensor based on the deflection of a microcantilever capable of detecting the SARS-CoV-2 spike (S) glycoprotein antigen using computationally designed multivalent minibinders immobilized on a microcantilever surface. The sensor exhibits rapid (< 5 min) detection of the target antigens down to concentrations of 0.05 ng/mL (362 fM). Validation of the sensor with clinical samples from 15 patients, including 5 patients infected with the Omicron (B.1.1.529) variant observed detection of antigen from nasopharyngeal swabs with cycle threshold (Ct) values as high as 39, suggesting a limit of detection similar to that of the quantitative reverse transcription polymerase chain reaction (RT-qPCR). Our findings demonstrate the use of minibinders for rapid, sensitive, and point-of-care detection of SARS-CoV-2 and potentially other disease markers.

The following text appeared in:

Agarwal*, D. K.; Hunt*, A. C.; *et al.* Rapid and sensitive detection of antigen from SARS-CoV-2 variants of concern by a multivalent minibinder-functionalized nanomechanical sensor. *Analytical Chemistry* (2022). (* Indicates equal contribution)

<https://doi.org/10.1021/acs.analchem.2c01221>

5.2 Introduction

The coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has highlighted the importance of the rapid and accurate detection of pathogens for disease identification and pandemic mitigation. The gold standard techniques for the detection of viral pathogens are the detection of viral nucleic acid by quantitative reverse transcription polymerase chain reaction (RT-qPCR) or viral antigen detection through lateral flow immunoassays (LFIA)^{454–456}. RT-qPCR based techniques are highly sensitive, but are relatively expensive, take hours to days to get results, and require a centralized laboratory with trained technicians^{454,457}. LFIAs, frequently called rapid antigen tests, are cheaper, faster, useful in point-of-care settings⁴⁵⁴, and have shown beneficial impact on population-level disease spread in widespread testing campaigns^{458,459}. However, SARS-CoV-2 LFIAs have moderate to low sensitivities at viral loads below 10^7 RNA copies per mL^{449,460}, which doesn't cover the range of viral loads where infected individuals transmit the virus^{449,459,461,462}, an issue during the emergence of the B.1.1.529 (Omicron) variant of concern^{463,464}. It is clear there remains a need for viral detection approaches that are rapid, sensitive, and clinically useful in point-of-care settings.

Towards addressing this need, we previously designed a nanomechanical microcantilever sensor platform that enables rapid and sensitive detection of SARS-CoV-2⁴⁶⁵. Microcantilevers (Fig. 10.1) have been utilized for sensing biomolecular binding in many contexts and are promising for disease diagnostics due to their rapid and sensitive detection of biomolecular targets and potential for point-of-care use^{466–468}. Binding events on the receptor-functionalized microcantilever surface result in surface stress that causes physical bending of the cantilever^{469,470}, which can be read out using optical⁴⁶⁵ or electronic^{471,472} means (Fig. 5.1A). Monoclonal antibodies are traditionally used for mediating specific binding of the target antigen of choice to the cantilever surface. However, they are typically non-specifically labelled, resulting in randomly oriented proteins on the cantilever surface⁴⁶⁵, and they are sensitive to point mutations in their target epitope, here the SARS-CoV-2 spike (S) protein antigen^{327,362}.

As an alternative to antibodies, we previously developed multivalent minibinders — small, computationally designed binding proteins^{377,382,418} — targeting the SARS-CoV-2 S glycoprotein trimer³⁵. The TRI2-2 multivalent minibinder, a trivalent version of the monovalent AHB2³⁷⁷ minibinder, simultaneously engages all three RBDs on a single S trimer and exhibits very high avidity and tight binding

to all tested SARS-CoV-2 variants³⁵. The minibinders are small (5-15x smaller than an antibody), can be economically produced in *Escherichia coli*³⁵, and can be site specifically functionalized with a cysteine residue to enable simple, oriented, and high density immobilization on sensor surfaces⁴⁷³⁻⁴⁷⁵. Additionally, multivalent proteins containing multiple binding domains can be easily designed to maximize binding avidity^{35,382} and reduce the minibinder's vulnerability to point mutations that result from antigenic drift in their target epitope³⁵.

In this work, we evaluated the performance of microcantilever sensors functionalized with the monovalent AHB2 and trivalent TRI2-2 minibinders^{35,377} for the detection of SARS-CoV-2 S antigen. We first functionalized gold-coated microcantilever sensors with minibinders containing C-terminal cysteine residues produced via cell-free protein synthesis^{38,42,87,292,317,476}. Using these sensors, we observed rapid (< 5 min) and sensitive detection (<0.05 ng/mL) of purified SARS-CoV-2 antigens from Wuhan-Hu-1, Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529) variants as well as detection of SARS-CoV-2 in patient nasopharyngeal swabs. Confident detection of S antigen is achieved at a concentration corresponding to 96 genome copies per mL. The observed limit of detection (LOD) is on-par with RT-qPCR-based tests⁴⁷⁷ and the developed technology is a promising point-of-care diagnostic platform.

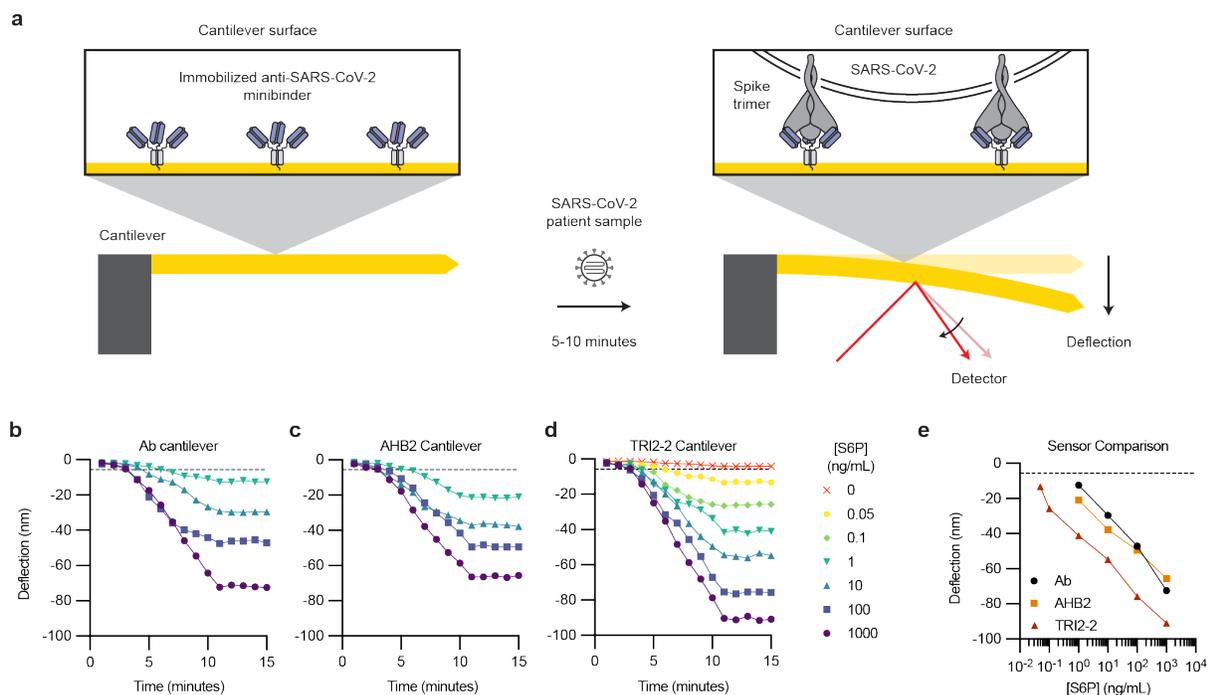


Figure 5.1. Cantilever-based sensing of SARS-CoV-2 spike protein (S6P). (A) Cantilever sensing mechanism; deflection is observed upon binding of the SARS-CoV-2 spike trimer to the captured binder immobilized on the cantilever. (b-d) Deflection of cantilever sensors over time with anti-S antibody (B), AHB2 (C), and TRI2-2 (D) immobilized on the cantilever (mean \pm SEM, $n = 3$). (E) Comparison of deflection between antibody, AHB2, and TRI2-2 cantilevers after 15 minutes of equilibration (mean \pm SEM, $n = 3$). For all plots, the dashed horizontal line indicates the deflection LOD cut-off (average of the combined negative control measurements + 3 standard deviations) and an absence of error bars indicates error within the marker.

5.3 Results

5.3.1 Comparison of antibody, AHB2, and TRI2-2 sensing of SARS-CoV-2 S trimer

We first compared the response of cantilevers functionalized with antibody, monomeric minibinder AHB2, and trimeric minibinder TRI2-2 to detect the Wuhan-Hu-1 SARS-CoV-2 pre-fusion stabilized hexaprotein spike protein (S6P) (Fig. 5.1)³⁹¹. Cantilever measurements were conducted in a small microfluidic chamber and deflection was monitored using an optical liquid atomic force microscopy (AFM) setup. Cantilevers were functionalized with the desired binder and then incubated with different concentrations of analytes in the sample chamber and monitored for 15 minutes. For all binders, the concentration of the S6P analyte exhibited a log-linear relationship with deflection and the system reached equilibrium after approximately 10 minutes of incubation (Fig. 5.1B-D). Replicates were highly concordant and exhibited low standard deviations (Fig. 10.2, mean standard deviation = 0.65 nm). TRI2-2 cantilevers exhibited more sensitive detection than either AHB2 or the antibody-functionalized cantilevers by more than an order of magnitude (Fig. 5.1E). At the lowest tested concentration of S6P (0.05 ng/mL or 362 fM) TRI2-2 exhibited signal significantly different from the negative control after 4 minutes ($p = 0.032$, 2way ANOVA with Sidak's multiple comparisons test) (Fig. 5.1D). The observed difference in response with the TRI2-2 cantilevers is likely due to the high avidity (it engages all three RBDs within an S trimer simultaneously)³⁵. When compared to the antibody cantilevers, the improvement may also be influenced by a greater immobilization density on the cantilever. This results indicates that TRI2-2 functionalized cantilevers are suitable sensors for the SARS-CoV-2 S trimer antigen.

5.3.2 Detection of S trimer from SARS-CoV-2 variants of concern

We next evaluated the ability of TRI2-2 cantilevers to sense S trimer corresponding to the Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529) variants of concern

(Fig. 5.2). We observed successful detection of 0.05 ng/mL for the Alpha, Gamma, Delta, and Omicron S trimer, and 0.1 ng/mL of the Beta S trimer, with the deflection crossing the limit of detection cut-off around 5 minutes at the lowest detectable concentration (Fig. 2A-E). We observed varying levels of deflection for the different S trimer variants (Fig. 2F), possibly due to heterogeneity in the different preparations of recombinant S trimer. These measurements demonstrate the ability of TRI2-2 functionalized cantilevers to detect antigens from different SARS-CoV-2 variants of concern.

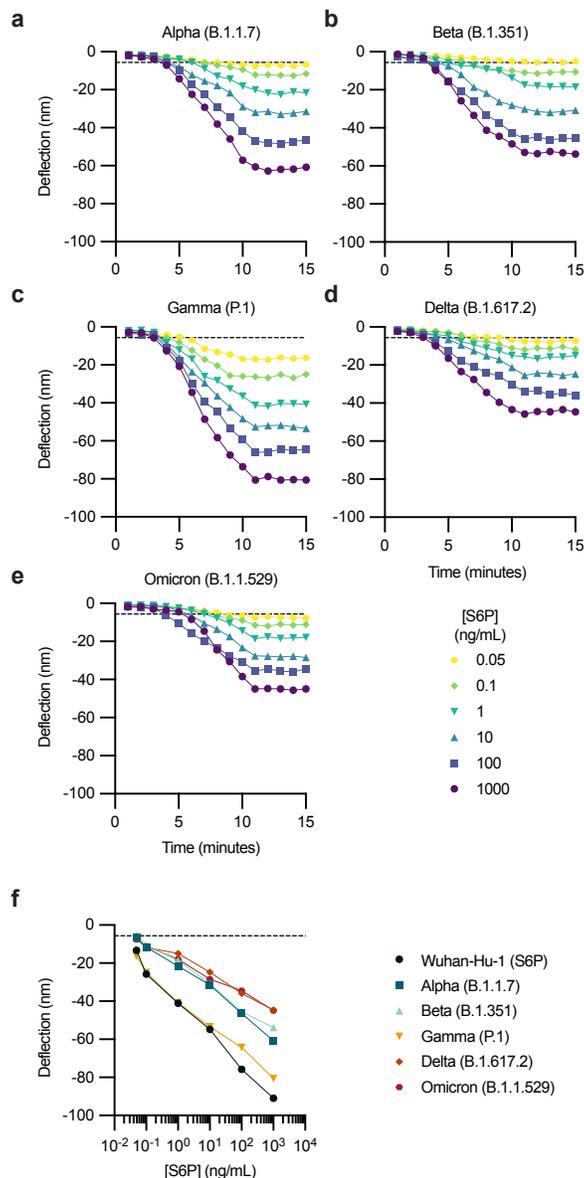


Figure 5.2. Detection of purified S trimer representing current and historical SARS-CoV-2 variants of concern. (A-E) Deflection observed over time in response to purified S6P variants corresponding to the Alpha (A), Beta (B), Gamma (C), Delta (D), and Omicron (E) variants (mean \pm SEM, $n = 3$). (F) Comparison

of cantilever deflection after 15 minutes of equilibration (mean \pm SEM, $n = 3$). For all plots, the dashed horizontal line indicates the deflection LOD cut-off (average of the combined negative control measurements + 3 standard deviations) and an absence of error bars indicates error within the marker.

5.3.3 Sensing of SARS-CoV-2 in clinical samples and evaluation of cross-reactivity

Given the highly sensitive detection of purified S trimer, we next sought to determine if TRI2-2 functionalized cantilevers could be utilized to detect SARS-CoV-2 in patient samples. Residual diagnostic nasopharyngeal swabs were collected from patients presenting to Northwestern Memorial Hospital between March of 2021 and January of 2022. Cycle threshold (Ct) values were calculated by quantitative reverse transcription and PCR (RT-qPCR) as a proxy for SARS-CoV-2 viral load. The genotype of the virus in each sample was determined by whole-genome sequencing using the ARTIC protocol. Pango lineages were assigned to the consensus sequences using pangolin software to assign variant designations. Specimens that tested negative for SARS-CoV-2 or positive for another respiratory virus (Influenza A Virus) were used as negative controls.

We tested a wide range of positive patient samples (Table 10.1), including samples from early in 2021, samples from confirmed Alpha variant infections, and samples from confirmed Omicron variant infections (Fig. 5.3A-C). A linear relationship was observed between Ct value measured by RT-qPCR and cantilever deflection (Fig. 5.3E), and thus a log-linear relationship with viral RNA copy number (Table 10.2). Consistent with the results using purified antigen, patient samples exhibited detection after approximately 5 minutes of incubation of the sample with the sensor. For the lowest Ct value sample tested (Ct = 39), we observed signal significantly different from the negative patient samples (Fig. 5.3D) after 4 minutes ($p=0.045$ or 0.0015 , 2way ANOVA with Sidak's multiple comparisons test), indicating confident detection of antigen concentration corresponding to 96 viral RNA copies per mL (Table 10.2). RT-qPCR negative samples from two separate patients exhibited little deflection (Fig. 5.3E), comparable to the buffer negative control (Fig. 1D). To probe the cross reactivity of our sensor, we tested an RT-qPCR positive Influenza A sample and recombinant purified middle east respiratory syndrome coronavirus (MERS-CoV) spike and observed little deflection (Fig. 3F), indicating our sensor is specific for SARS-CoV-2.

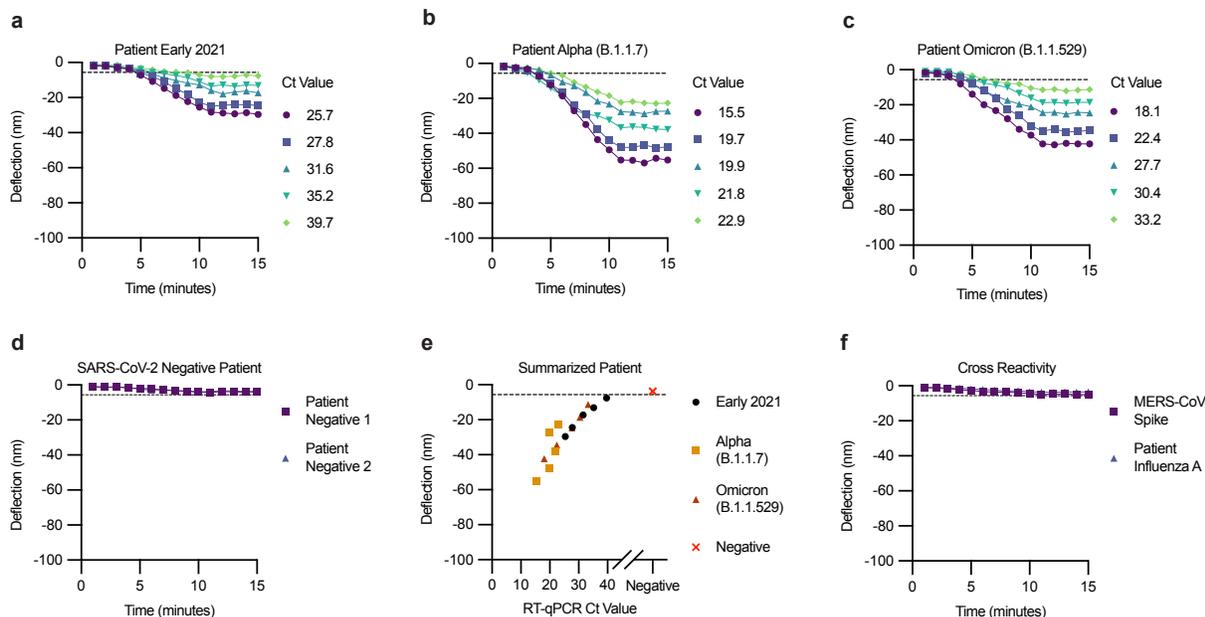


Figure 5.3. Detection of SARS-CoV-2 in nasopharyngeal swabs from infected individuals. All samples were evaluated in triplicate using three distinct cantilevers. (A-C) Deflection of cantilevers observed over time in response to samples collected early in 2021 (A), from samples confirmed to contain the Alpha variant (B), and from samples confirmed to contain the Omicron variant (C). (D) Deflection of cantilevers measured against patient nasopharyngeal swabs RT-qPCR negative for SARS-CoV-2 (mean \pm SEM, $n = 3$). (E) Comparison of the RT-qPCR Ct value and cantilever deflection for the tested patient samples (mean \pm SEM, $n = 3$). (F) Deflection of cantilevers measured against recombinant purified MERS-CoV spike protein (1,000 ng/ml) and against a patient nasopharyngeal swab RT-qPCR positive for Influenza A (Ct = 19.33) (mean \pm SEM, $n = 3$). For all plots, the dashed horizontal line indicates the deflection LOD cut-off (average of the combined negative control measurements + 3 standard deviations) and an absence of error bars indicates error within the marker.

5.4 Discussion

Here, we have demonstrated a nanomechanical sensor platform for the detection of the SARS-CoV-2 S trimer antigen using multivalent minibinders. Measurements with patient samples do not require amplification steps, show results in <5 minutes, and an LOD >2 orders of magnitude better than currently authorized LFIA antigen tests^{449,462,464}. They also exhibit an LOD that is comparable to or better than both laboratory and point-of-care nucleic acid tests either with^{6,477-479} or without amplification⁴⁸⁰. This low LOD could also enable simple pooled sample testing to conserve resources⁴⁸¹. The system can be made accessible in point-of-care settings by leveraging alternative detection modalities like a metal-oxide semiconductor field-effect transistor (MOSFET) detector for electronic readout in a handheld device⁴⁸². Further development may also enable simultaneous multiplexed detection of multiple distinct antigens⁴⁸².

The highly sensitive detection of antigens in the femtomolar range may also have applications in other relevant fields, such as cancer diagnostics^{483,484}. We anticipate that the reported sensor platform will enable fast and sensitive detection of antigens for disease diagnosis and pandemic mitigation for SARS-CoV-2 as well as for other pathogens and disease markers.

5.5 Materials and Methods

5.5.1 Chemical reagents and materials

SARS-CoV-2 Chimeric monoclonal antibody for Spike (S1) protein (40150-D003), SARS-CoV-2 Spike protein (40150-V08B2), and MERS-CoV spike (S1) protein (40069-V08H) were all purchased from Sino Biological Inc. Recombinant SARS-CoV-2 spike protein for the Wuhan-Hu-1 (SPN-C52H9), Delta (SPN-C52He), and Omicron (SPN-C52Hz) variants were purchased from Acro Biosystems. SARS-CoV-2 spike proteins for the Alpha, Beta, and Gamma variants were expressed and purified by Lauren Carter at the Institute for Protein Design at the University of Washington. Phosphate Buffer Saline (P3813) and PBS-Tween 20 sachets (P3563) were purchased from Millipore-Sigma. Reducing agent tris (2-carboxyethyl) phosphine hydrochloride (TCEP) was purchased from Sigma Aldrich (C4706). The set-up used for the deflection experiments was Bruker Bioscope Resolve liquid imaging system. The tipless silicon cantilevers used in these experiments were acquired from Nanoworld Incorporation. (ARROW-TL1Au-50). All the target protein solutions were prepared in Phosphate Buffer Saline (pH 7.4) solution.

5.5.2 Recombinant protein expression and purification

Proteins were expressed and purified as previously described³⁵. Briefly, cell free protein synthesis (CFPS) reactions were run using an extract from the BL21 StarTM (DE3) (Thermo Fisher, C601003) strain of *Escherichia coli*. Cell free extract preparation methods and cell free protein synthesis reaction compositions were the same as previously described³³. Linear DNA templates were generated using Gibson assembly of synthetic linear dsDNA DNA (Integrated DNA Technologies) with the pJL1 backbone followed by PCR amplification of the linear expression template as previously described³⁵. Unpurified linear expression templates at a concentration of 6.66% v/v were used to drive CFPS reactions. 2 mL reactions were run in 6-well plates (Costar, 3736) for 20 hours at 30°C. AHB2 proteins were purified using StrepTactinXT Superflow high capacity resin (IBA, 2-4030-010) as previously described³⁵ and extensively

dialyzed into 50 mM HEPES pH 7.4, 150 mM NaCl using a 3.5K molecular weight cut off cassette (ThermoFisher Scientific, 87724). TRI2-2 proteins were purified using Ni-NTA Agarose (Qiagen, 30230) as previously described³⁵ and extensively dialyzed into 50 mM HEPES pH 7.4, 150 mM NaCl using a 3.5K molecular weight cut off cassette (ThermoFisher Scientific, 87724). The protein and DNA sequence corresponding to the purified proteins are below.

AHB2_cys protein sequence:

MEKKIELEEQVMHVLDQVSELAHELLHKL TGEELERAAYFNWWATEMMLELIKSDDEREIREIEEEEARRIL
EHLEELARKGGASPAAPAPGGCGGSAWSHPQFEK

AHB2_cys open reading frame DNA sequence:

ATGGAGAAAAAATCGAGCTGGAAGAGCAGGTCATGCATGTGTTGGACCAGGTATCCGAATTAGCG
CACGAAC TTTTACATAAGTTGACGGGCGAAGAATTGGAACGTGCAGCCTACTTTAACTGGTGGGCCA
CCGAAATGATGCTGGAGCTGATTAATCTGATGATGAGCGTGAAATCCGTGAAATTGAAGAAGAAGC
TCGTCGTATTCTTGAGCATTGGAGGAGCTGGCACGCAAGGGTGGAGCCAGTCCAGCAGCTCCTGC
GCCTGGCGGGTGTGGTGGTTCGCGATGGTCCCACCCTCAATTTGAAAAATAA

TRI2-2_cys protein sequence:

MEKKIH HHHHHGGGSGGGELEEQVMHVLDQVSELAHELLHKL TGEELERAAYFNWWATEMMLELIKSD
DEREIREIEEEEARRILEHLEELARKGGSEALEELEKALRELKKSTDELERSTEELEKNPSEDALVENNRLIV
ENNKIIVEVLRIIAKVLKGGASPAAPAPGGC

TRI2-2_cys open reading frame DNA sequence:

ATGGAGAAAAAATCCACCATCACCATCACCATGGTGGAGGAAGTGGAGGCGGAGAGTTGGAAGAA
CAGGTGATGCACGTA CTTGACCAGGTTTCGGAATTGGCGCACGAATTATTACACAACTTACGGGTG
AAGAATTGGAACGCGCGGCGTATTTTAACTGGTGGGCTACAGAGATGATGTTAGAGTTGATTAATC
GGACGACGAGCGCGAAATCCGCGAGATCGAAGAAGAGGCACGTCGTATCCTTGAGCACTTGGAGG
AATTGGCTCGCAAGGGTGGCTCAGAAGCGTTGGAAGAGTTGGAGAAAGCCCTTCGTGAGCTTAAGA
AATCAACGGACGAGTTAGAGCGTTCCACTGAGGAACTTGAAAAAACCCGAGTGAAGATGCTTTAGT
TGAGAATAACCGTTTAAATCGTAGAGAATAACAAAATTATTGTGCGAGGTACTTCGTATTATTGCTAAAGT
CCTTAAGGGCGGGGCAAGCCCAGCAGCTCCGGCTCCTGGCGGCTGTTAA

5.5.3 Cantilever surface biofunctionalization

The gold coated cantilevers were plasma cleaned prior to surface biofunctionalization (peptide immobilization) step to remove any organic residues. The synthesized cysteine terminated minibinder TRI2-2 to SARS-CoV-2 spike proteins were covalently attached onto the gold coated cantilever surface through thiol chemistry inside a glass plate containing tiny wells. Prior to this, the peptide solution (20 µg/ml) was reacted with a reducing agent TCEP solution (prepared in DI water) in 1:1 ratio for around 1 hour at room temperature to cleave the pair of Cysteine sulfhydryl groups linked by di-sulphide bond (-S-S-). Following this, the gold-coated microcantilevers were allowed to incubate with peptide solution for two hours to facilitate a covalent immobilization. Microcantilevers were then washed with PBS-tween-20 solution, dried and then fixed in the AFM sample holder for deflection measurement.

5.5.4 Optical detection method

The tipless gold-coated microcantilevers used for this study was 500 µm in length, 95 µm in width and 1 µm thick. Because a cantilever deflection also strongly depends on geometry, all cantilevers used in this manuscript were the same geometry. The cantilevers and experimental conditions used in each measurement were similar. The absolute deflection at the free end of each cantilever (Δz) was measured using a fluidic-atomic force microscopy (f-AFM) based optical detection system. A microfluidic reaction chamber sized approximately 2 mm in diameter was used to perform the deflection measurement. The chamber temperature was remained constant over the course of experiments. The biofunctionalized microcantilevers were brought near the microfluidic chamber containing target solutions (each 10 µl approx.) through a stepper motor. The deflection experiments were performed for different concentrations of SARS-CoV-2 trimeric spike (S1) proteins, UK mutant variant along with real clinical samples. The cantilever vertical deflection was measured as cantilever start bending (downwards) due to differential surface stress. The limit of detection for cantilever deflection was calculated as the mean \pm 3x the standard deviation of all tested negative controls (buffer only, negative patient 1, negative patient 2, patient Influenza A, and recombinant MERS) after 15 minutes of equilibration of the cantilever with the sample.

5.5.5 Patient sample collection and Ct value calculation

Residual nasopharyngeal specimens from patients presenting at Northwestern Memorial Hospital with were collected per study protocol STU00212260. Nasopharyngeal swabs were stored in Viral

Transport Media, inactivated by incubation at 60°C for 1 hour, de-identified, and frozen in 1mL aliquots at -80°C prior to downstream use. All patient samples were processed prior to optical detection method for microcantilever deflection monitoring. The samples were first centrifuged at 5000 rpm for 1 minute to remove mucous and other interferants. The samples were then mixed with the extraction buffer for 10 minutes and then used for the measurements.

5.5.6 Ct Value Calculation and Variant Identification

Viral RNA was extracted from nasopharyngeal specimens utilizing the QIAamp Viral RNA Minikit (Qiagen). Cycle threshold (Ct) values were calculated by quantitative reverse transcription and PCR (qRT-PCR) with the CDC 2019-nCoV RT-PCR Diagnostic Panel utilizing N1 and RNase P probes. Standard curves for N1 and RNase P were included in all qRT-PCR experiments using known standards (IDT # 10006626, 10006625) and run alongside negative controls. For specimens with a Ct value <34, whole genome sequencing of the SARS-CoV-2 genome was used to identify variants. From isolated RNA, cDNA synthesis was performed with SuperScript IV First Strand Synthesis Kit (Thermo) using random hexamer primers according to manufacturer's specifications. Direct amplification of the viral genome cDNA was performed in multiplexed PCR reactions to generate ~400 base pair amplicons tiled across the genome. The multiplex primer set, comprised of two non-overlapping primer pools, was created using Primal Scheme and provided by the Artic Network (version 4.1). PCR amplification was carried out using Q5 Hot Start HF Taq Polymerase (NEB) with 5 µl of cDNA in a 25 µl reaction volume. A two-step PCR program was used with an initial step of 98°C for 30 seconds, then 35 cycles of 98°C for 15 seconds followed by five minutes at 65°C. Separate reactions were carried out for each primer pool and validated by agarose gel electrophoresis. Sequencing library preparation of genome amplicon pools was performed using the SeqWell plexWell 384 kit per manufacturer's instructions. Pooled libraries of up to 96 genomes were sequenced on the Illumina MiSeq using the V2 500 cycle kit. Sequencing reads were trimmed to remove adapters and low-quality sequences using Trimmomatic v0.36. Trimmed reads were aligned to the reference genome sequence of SARS-CoV-2 (accession MN908947.3) using bwa v0.7.15. Pileups were generated from the alignment using samtools v1.9 and consensus sequence determined using iVar v1.2.2 with a minimum depth of 10, a minimum base quality score of 20, and a consensus frequency threshold of 0 (*i.e.*, majority base as the consensus). Consensus sequences with ≥ 10% missing bases were discarded.

SARS-CoV-2 clades were assessed using Nextclade (clades.nextstrain.org) and Pango lineages were assigned to the consensus sequences using pangolin software. All SARS-CoV-2 genome sequences were deposited in the public GISAID database (Table 10.1).

Chapter 6 –Future Work and Perspectives

6.1 Proposed Future Work

In this dissertation, I have clearly demonstrated the utility of high-throughput cell-free methodologies for individual protein expression and evaluation. However, as is the case with all PhDs, my journey ends with more work to do than when I started. There are numerous potential extensions of my work, some already ongoing in the projects of other researchers in the Jewett lab and elsewhere.

A simple and direct extension of my would be to accelerate broadly neutralizing therapeutic protein discovery with cell-free protein synthesis. Broadly neutralizing therapeutics can neutralize many different genetic variants of a pathogen, typically by targeting highly conserved epitopes, and thus are highly sought after⁴⁸⁵. Indeed, the antibody therapeutics for the treatment of COVID-19 that have seen the most success are broadly neutralizing^{370,371,486}, and broadly neutralizing antibodies against other viruses like the human immunodeficiency virus⁴⁸⁷ and influenza⁴⁸⁸ are promising. However, these antibodies are frequently rare, and thus methods that are sufficiently high throughput to identify these candidates are of great use⁴⁸⁵. Such an effort would build off the antibody discovery pipeline presented in **Chapter 3**.

In addition to direct extensions of my work, there are opportunities to further improve the pipeline for the identification and characterization of binding proteins. Coupling the methods presented in this dissertation to an in vitro display-based selection could enable the generation and characterization of high affinity binding proteins from naïve libraries in a matter of days. I am also particularly excited about the application of high throughput protein expression for the profiling and directed evolution of enzymes. Other possible extensions of the work in this dissertation will be left as an exercise for the reader.

6.2 Perspective

The work that I present here is the culmination of more than half a decade of immersion in the field of cell-free systems and engineering biology. I see the biggest accomplishments of my PhD as ones of method development, method integration, and research acceleration. At the beginning of my PhD, I spent days cloning and expressing a single protein. By the end of my PhD, I could routinely individually express and functionally characterize hundreds of proteins in a single day, and others have already adapted this

workflow and are regularly expressing thousands of protein variants. The flexibility of screening individual variants in high-throughput also allows the methods to be highly generalizable beyond protein-protein interactions to other protein functions.

The individual chapters also contain major scientific advancements on their own. **Chapter 2** details a new method for utilizing protein-protein interactions to build logical operations, which ultimately will enable designed biological systems to integrate and respond to new information in minutes instead of hours like traditional transcription factor-based genetic circuits. The work in **Chapter 3** lays the groundwork for an accelerated antibody discovery pipeline with more than an order of magnitude improvement in throughput and an approximately 3.5x reduction in the time required to screen antibody candidates when compared to state-of-the-art methods. A major achievement in **Chapter 4** is the design of a molecule that has sufficient therapeutic potential for the treatment of COVID-19 to warrant clinical trials. At the time of this writing, the lead molecule is undergoing investigational new drug enabling studies prior to human trials. An understated achievement in **Chapter 4** is the routine nature with which we were able to screen for the biophysical effects multivalent binding. Multivalency is a phenomenon leveraged throughout biology³⁹², however in protein-protein interactions it can be difficult to assess the biophysical consequences with standard techniques, particularly for high affinity interactions. The developed multivalency screen not only was able to distinguish the consequences of multivalent binding, but it also allowed us to design proteins with apparent dissociation rates that were measured to be slower than $1 \times 10^{-7} \text{ s}^{-1}$. This is comparable to that of the biotin-streptavidin interaction⁴⁸⁹, one of the regularly cited strongest known biomolecular interactions. The work in **Chapter 5** achieves fM antigen detection in minutes, which is >100x more sensitive than state of the art lateral flow tests, with a minibinder-based detector, highlighting the potential of non-antibody binders in the field of diagnostics.

I hope that this dissertation provides additional grounds for the continued development of cell free systems and computationally designed therapeutics. Continued development of cell free systems will enable faster interrogation and design of increasingly complex biological systems, ultimately impacting research and development in the sectors of materials, energy, medicine and more. Computationally designed therapeutics, in my eyes, also have enormous potential not only to improve upon current therapeutic modalities, but also more importantly be more accessible therapeutics. Ultimately, I also hope

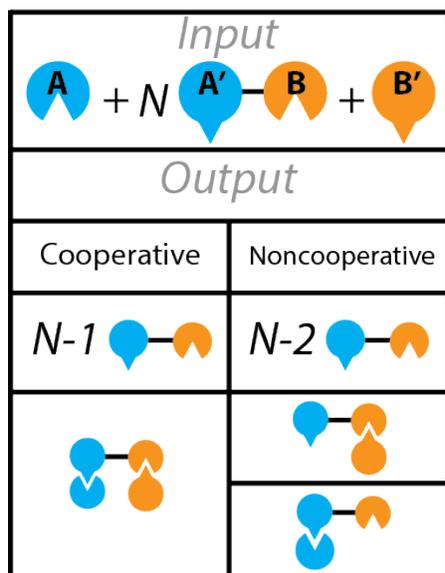
that my work serves as inspiration for others, particularly would be method developers, to continue innovating and seeking out solutions beyond the default.

“And so it goes...”

-Kurt Vonnegut

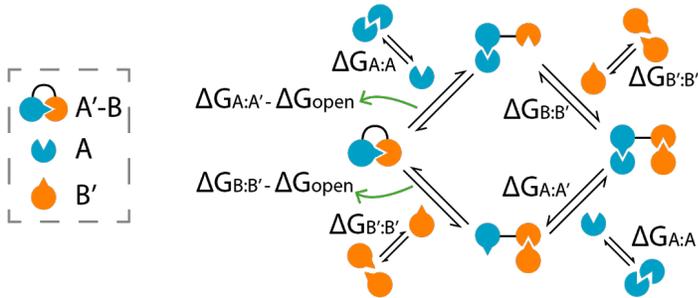
Chapter 7 - Appendix Containing Supplementary Information for Chapter 2

7.1 Thermodynamic Modeling of Cooperativity



For an induced dimerization system involving proteins A, A'-B, and B', a stoichiometric excess ($N \gg 1$) of the A'-B protein results in partially assembled dimeric complexes if the binding is non-cooperative, but fully assembled trimeric complexes if the binding is cooperative.

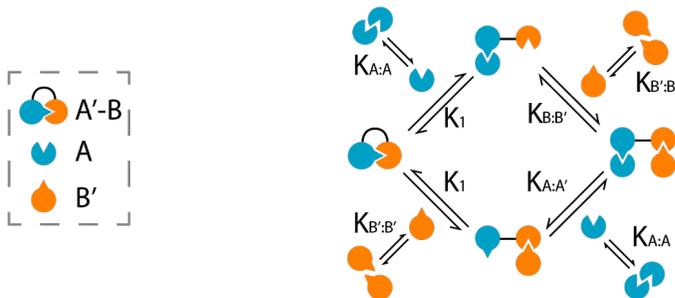
We model the cooperatively induced dimerization system at thermodynamic equilibrium. Shown below, assuming a 'closed' state for A'-B, where the binding interfaces are buried within the four-helix bundle, the binding of A'-B to either A or B' helix hairpins needs to overcome an energy barrier of transitioning from the 'closed' to 'open' state (ΔG_{open}). Therefore the free energy of binding between A'-B to A or B' can be expressed as $\Delta G_{A:A'} - \Delta G_{\text{open}}$ and $\Delta G_{B:B'} - \Delta G_{\text{open}}$, respectively, where $\Delta G_{A:A'}$ and $\Delta G_{B:B'}$ represent the free energy of binding between the cognate pairs in the absence of the fusion. Once the A:A'-B or A-B':B complexes form, subsequent binding can simply be represented by the binding between cognate heterodimers: $\Delta G_{A:A'}$ or $\Delta G_{B:B'}$. We also observed the presence of (A)₂ and (B')₂ homodimers, therefore we added free energy terms describing such processes into the model ($\Delta G_{A:A}$ or $\Delta G_{B':B'}$).



ΔG relates to equilibrium constants by $\Delta G = -RT \ln K$, and we further consider the system in terms of K . We make the simplifying assumption that the affinity of $A'-B$ to either A or B' is identical ($K_1 = [A:A'-B]/([A][A'-B]) = [A'-B:B']/([B][A'-B])$). Finally, we define the cooperativity of the system, c , as the ratio between the equilibrium constants in the presence or absence of the other partner ($c = K_{B:B'}/K_1 = K_{A:A'}/K_1$). For an entirely non-cooperative process ($c=1$), $K_{B:B'}=K_1$ and $K_{A:A'}=K_1$ *i.e.*, the first binding event does not affect the affinity of the subsequent binding event.

Since $K_1 = \exp(-(\Delta G_{A:A'} - \Delta G_{open})/RT)$, rewriting the equation for c in terms of free energies leads to $c = \exp(\Delta G_{open})/RT$. Therefore, the extent of cooperativity is solely determined by the magnitude of the free energy required to partially unfold/expose the buried binding interfaces of the dimerizer $A'-B$.

We note that explicitly incorporating the equilibrium constants for homodimerization ($K_{A:A}$ and $K_{B':B'}$) only affect the absolute position of each equilibrium, but does not affect the magnitude of the cooperativity. Indeed, taking A as an example, the binding to the closed state becomes $K_1 * K_{A:A}$, and the binding to the open state becomes $K_{A:A'} * K_{A:A}$. Because $K_{A:A}$ is present in both the numerator and the denominator, they cancel out, and c remains purely defined by the relative magnitudes of K_1 and $K_{A:A'}$.



We solved the following system of equations in *Mathematica* 11 (Wolfram Research, Inc.) to simulate the amount of A:A'-B:B' at equilibrium as a function of the initial concentration of A'-B:

$$K_{A:A} = \frac{[A_2]}{[A][A]}$$

$$K_{B':B'} = \frac{[B'_2]}{[B'][B']}$$

$$K_1 = \frac{[A:A' - B]}{[A][A' - B]}$$

$$K_1 = \frac{[B':A' - B]}{[B'][A' - B]}$$

$$K_{A:A'} = \frac{[A:A' - B:B']}{[A][A' - B:B']}$$

$$K_{B:B'} = \frac{[A:A' - B:B']}{[B'][A' - B]}$$

$$[A]_{tot} = 2 * [A_2] + [A] + [A:A' - B] + [A:A' - B:B']$$

$$[B']_{tot} = 2 * [B'_2] + [B'] + [A' - B:B'] + [A:A' - B:B']$$

$$[A' - B]_{tot} = [A' - B:B'] + [A:A' - B] + [A:A' - B:B']$$

We knew from previous native MS titration experiments that the equilibrium dissociation constants of cognate designed heterodimers (DHDs) is in the ~10 nM range²⁸³, therefore $K_{A:A'} = K_{B:B'} = 0.1 \text{ nM}^{-1}$. Varying values of K_1 (and hence the cooperativity factor, $c = K_{A:A'}/K_1$) showed different responses of the amount of A:A'-B:B' at equilibrium as a function of the initial concentration of A'-B, as shown in Fig. 1C.

We experimentally estimated K_1 using native MS experiments. Mixing 10 μM of **1** and **1'-2'** resulted in no detectable amount of the **1:1'-2'** complex, suggesting very weak binding. The sensitivity of native MS places a lower-bound on the concentration of species that can be detected (0.0375 μM , Table 7.8). Using this value, a lower-bound for the affinity of **1:1'-2'** can be estimated ($1/K_1 \geq 2.65 \text{ mM}$). This agrees with the value of 9.91 mM obtained by calculating the affinity based on the c value of 991,000 reported in Fig. 5.1H.

This thermodynamic modeling demonstrates that binding cooperativity can be achieved for an induced dimerization system through occlusion of the binding interfaces. We achieved this by fusing

hairpins via a flexible linker, rationalizing that the spontaneous folding of these constructs would bury the interaction interfaces on the inside of a four helical bundle like topology. Formation of these structures is corroborated by: *i*) SAXS profiles that are consistent with DHDs structures, *ii*) *m*-values from chemical denaturation experiments consistent with ΔS_{ASA} for the unfolding of DHD topologies (Fig. 7.2, Table 7.3), and *iii*) $\Delta G_{open} < \Delta G_{folding}$, suggesting that exposing the binding interfaces requires partial unfolding of these fused constructs, but does not exceed the folding free energy of these proteins (a physically unrealistic scenario).

7.2 Supplementary Figures

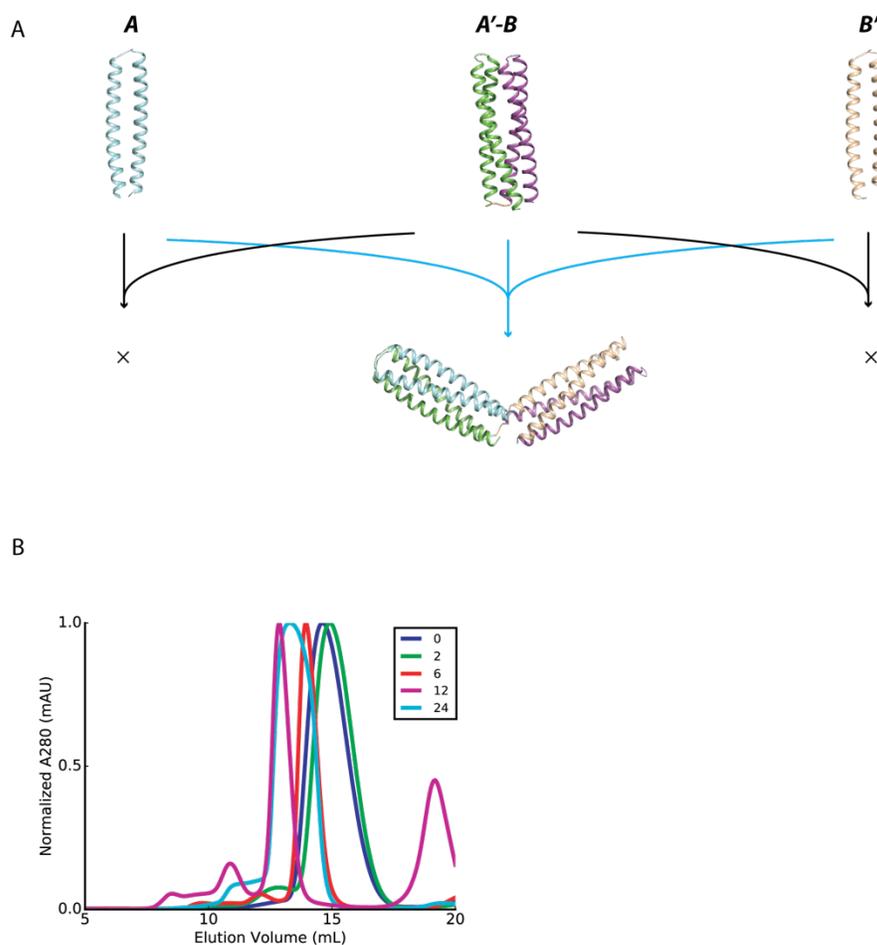


Figure 7.1. Additional information on the induced dimerization system. (A) Molecular implementation of the cooperative induced dimerization system, binding only occurs when all three components are present. (B) Size exclusion chromatography profiles of $1'$ - $2'$ variants with 0, 2, 6, 12, and 24 amino acids as the flexible linker connecting $1'$ and $2'$.

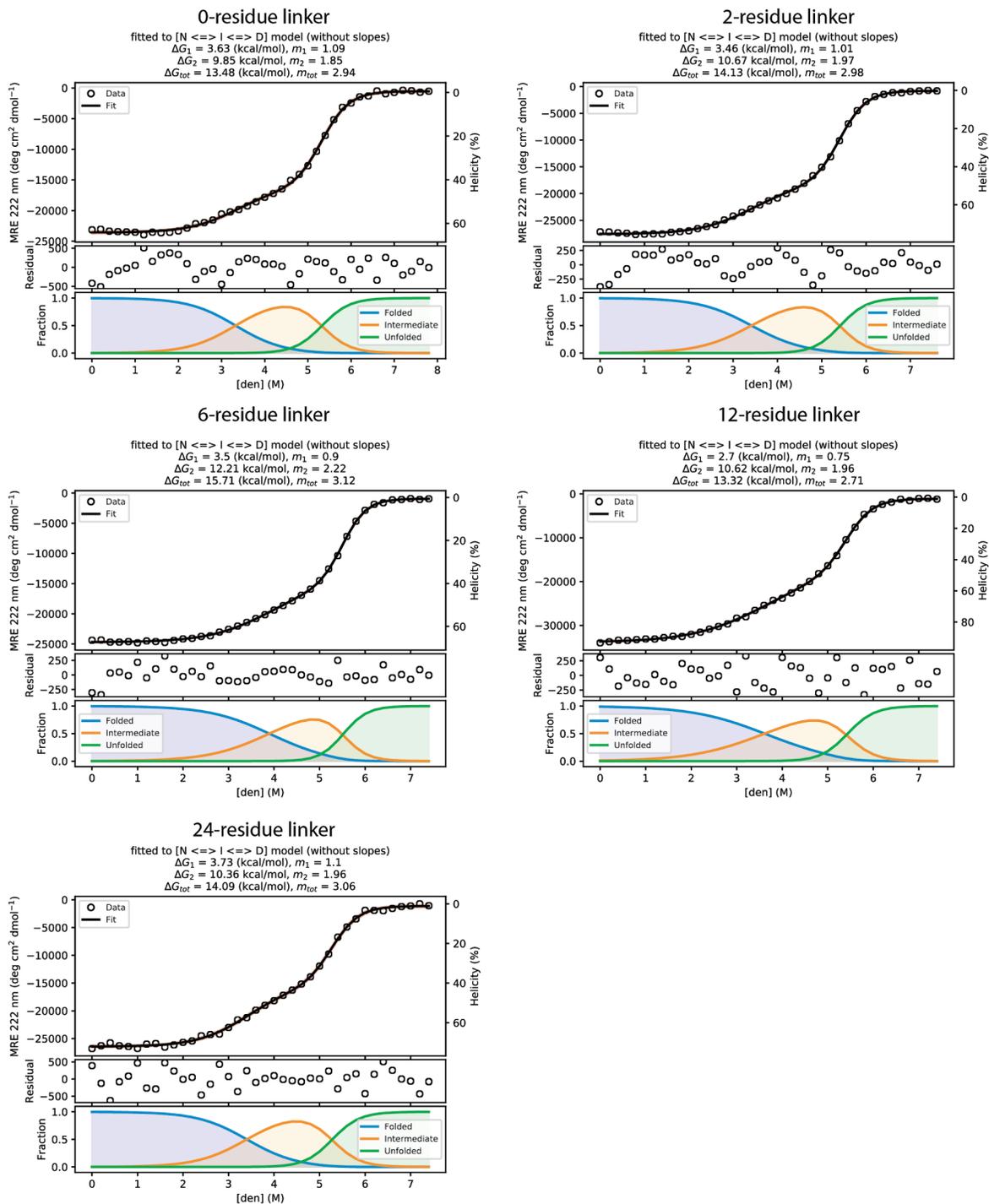


Figure 7.2. GdnHCl equilibrium denaturation experiments monitored by CD at 222 nm for 1'-2' designs with 0-, 2-, 6-, 12- and 24- residue linkers. For each plot, the top subplot shows the experimental

data (circles) fitted to the 3-state unimolecular unfolding model (black line). The middle subplots show residuals to the fit, and the bottom subplots indicate the fraction of each species as a function of denaturant concentration. An approximation of helical content is also reported (second ordinate, based on the MRE value at 222 nm using the model of Muñoz & Serrano⁴⁹⁰). All fitted parameters are reported in Table 7.3. ΔG and m -values are also reported in the title of each plot. ΔG represent the free energy change extrapolated to buffer. The corresponding m -values represent the denaturant sensitivity of each transition, and relate to the $\Delta SASA$ associated with each event.

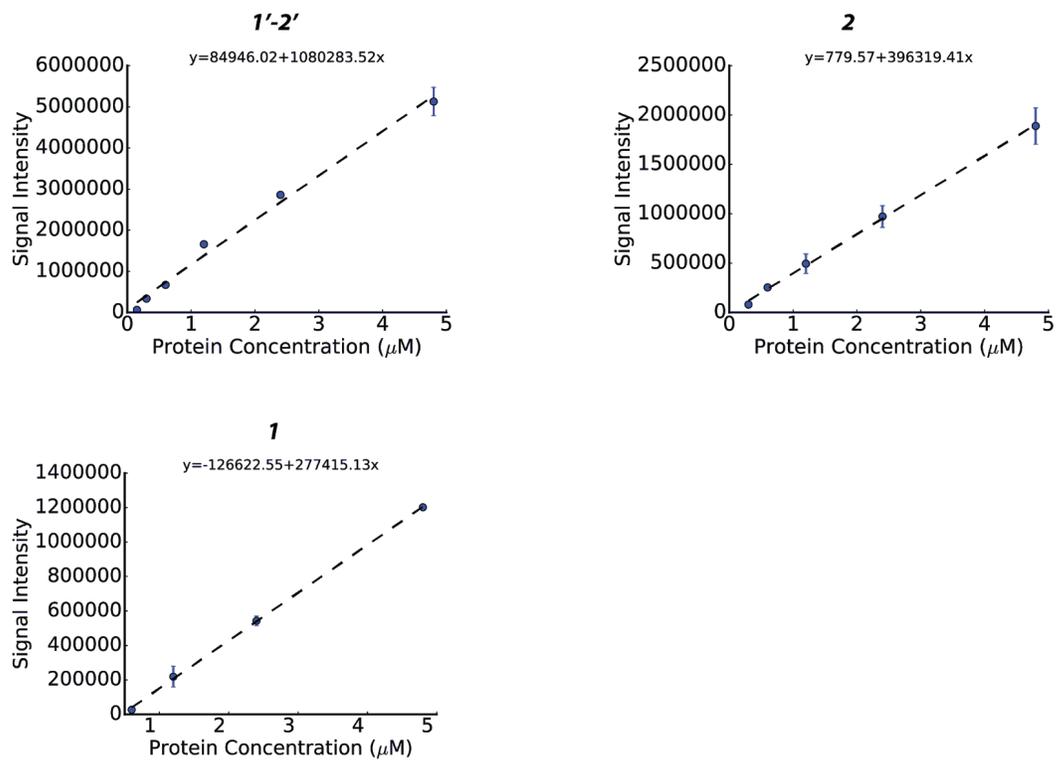


Figure 7.3. Native MS signal intensity of the dimerizer 1'-2' (6-residue linker), 1, and 2 against protein concentrations (see also Table 7.8). Equations for linear regression fits are displayed. All error bars are reported as standard deviations of $n=3$ independent replicates.

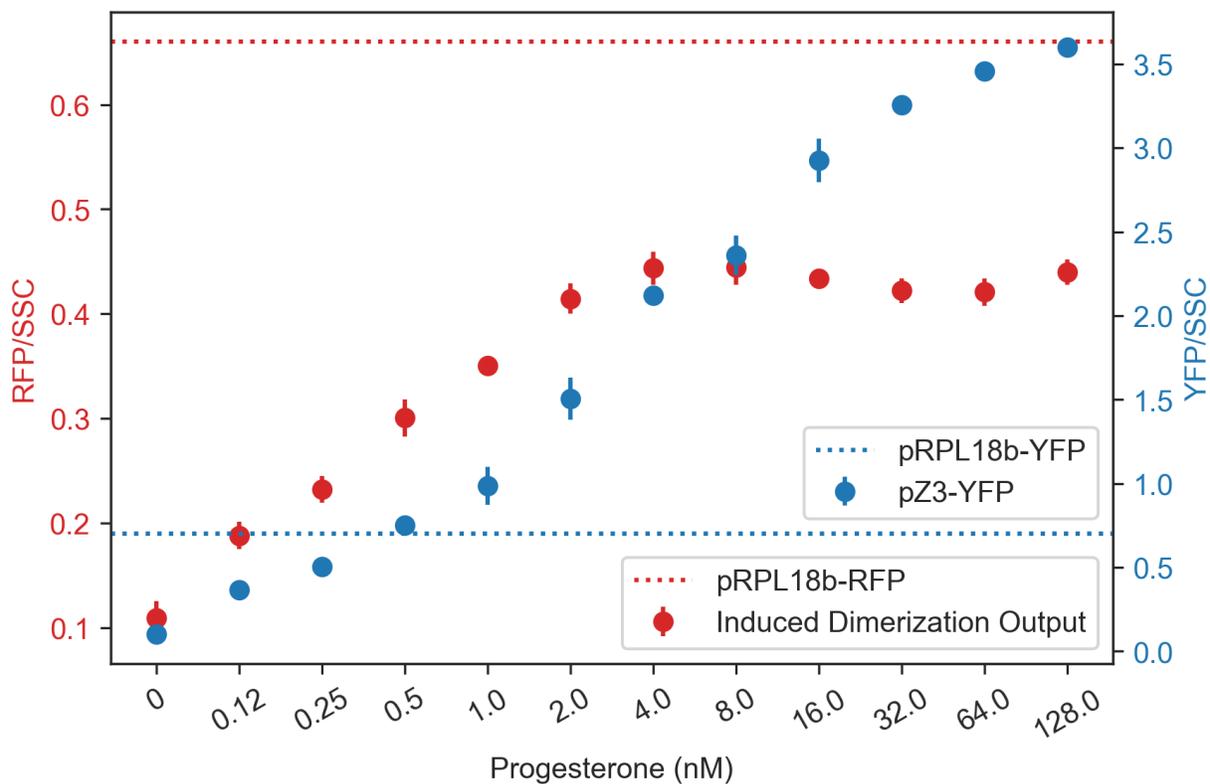


Figure 7.4. Induction of RFP (mScarlet) expression by the “Induced Dimerization” system in Fig. 2.1J, where the dimerizer species is measured by a separate, proxy “pZ3-YFP” (Venus) reporter. Dotted lines represent constitutive expression levels of YFP and RFP under the RPL18b promoter used to express the DBD and AD monomer species in Fig. 2.1J. Maximal induction of the dimerizer/YFP (128 nM Progesterone) is about 5x the value of pRPL18b-YFP, indicating that the cooperative binding functions across a wide range of concentrations. All error bars are reported as standard deviations of n=3 independent replicates.

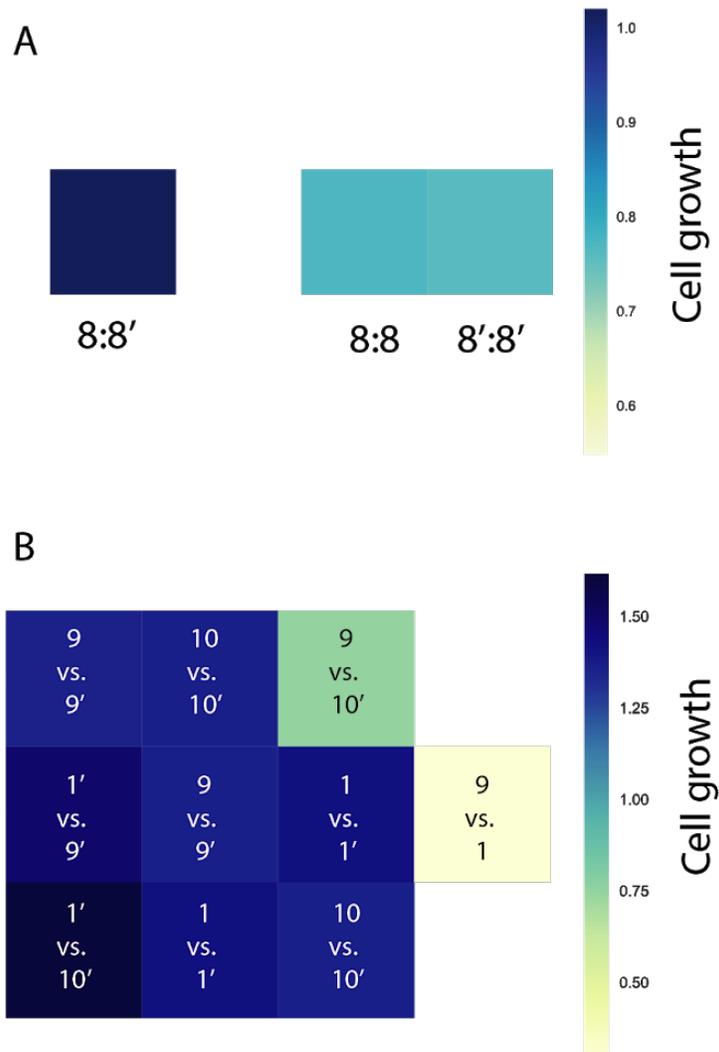


Figure 7.5. Binding affinity gradient from individual Y2H experiments. (A) The 8:8' heterodimer binds more tightly than the homodimers of its monomers. (B) Binding affinity gradient among the monomers of 1:1', 9:9', and 10:10' pairs. Cell growth was measured on different days across the rows but on the same day within each row, and not corrected against background growth.

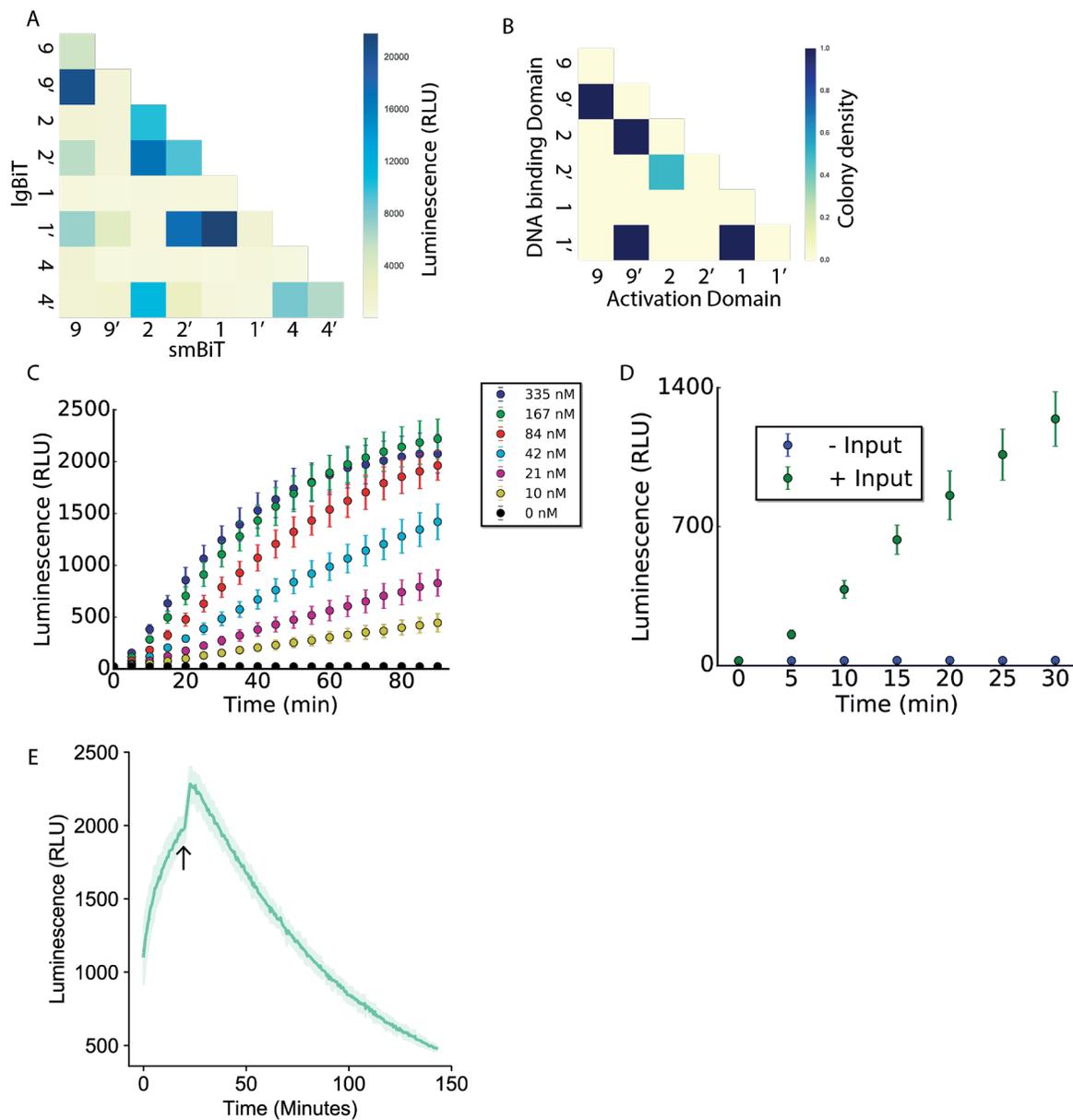


Figure 7.6. NanoBiT characterization of CIPHR logic gate components. (A) Specificity matrix for designs tested in the cell-free expression system. (B) Yeast two-hybrid specificity matrix for the same designs. (C) Kinetic measurements of induced dimerization gate at varying concentrations of the dimerizer input. (D) Kinetics of induced dimerization (with or without 335 nM dimerizer protein) in the first 30 minutes. (E) Kinetic measurements of NOR gate activation and subsequent deactivation via competitive binding upon addition of input (arrow). All error bars are reported as standard deviations of $n=3$ independent replicates.

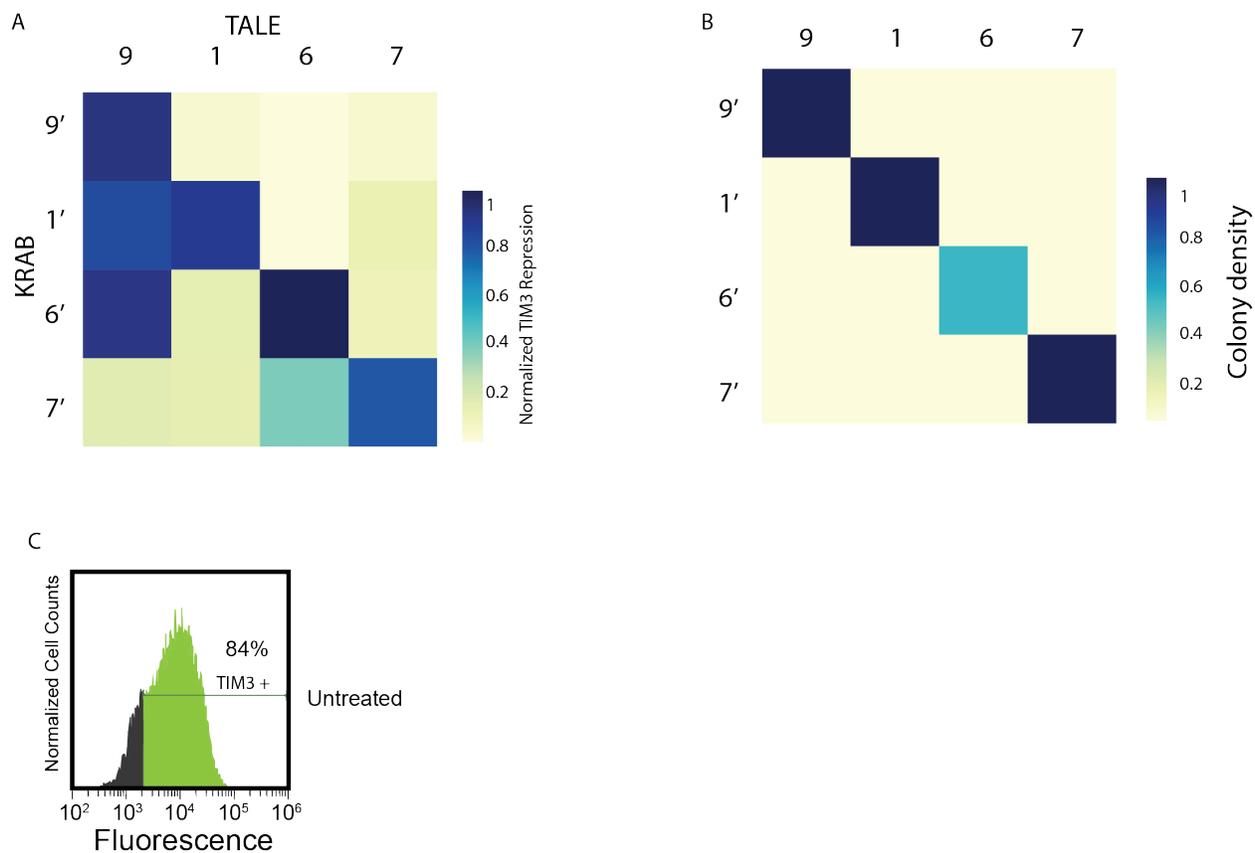


Figure 7.7. T cell KRAB-TALE characterization of CIPHR logic gate components. (A) Specificity matrix for designs tested in T cells. (B) Yeast two-hybrid specificity matrix for the same designs. (C) Expression level of TIM3 in T cells not treated with CIPHR logic gates.

Table 7.1. Design Sequences. Amino acid sequences of all designs tested. Two versions of DHD 15 heterodimers were used here, with slight differences in the loop sequence. For simplicity they are all labeled 2 and 2', differentiated with an additional asterisk (*). See attached Supplementary Table 1.xlsx for design sequences and Vectors.zip for plasmid maps. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>.

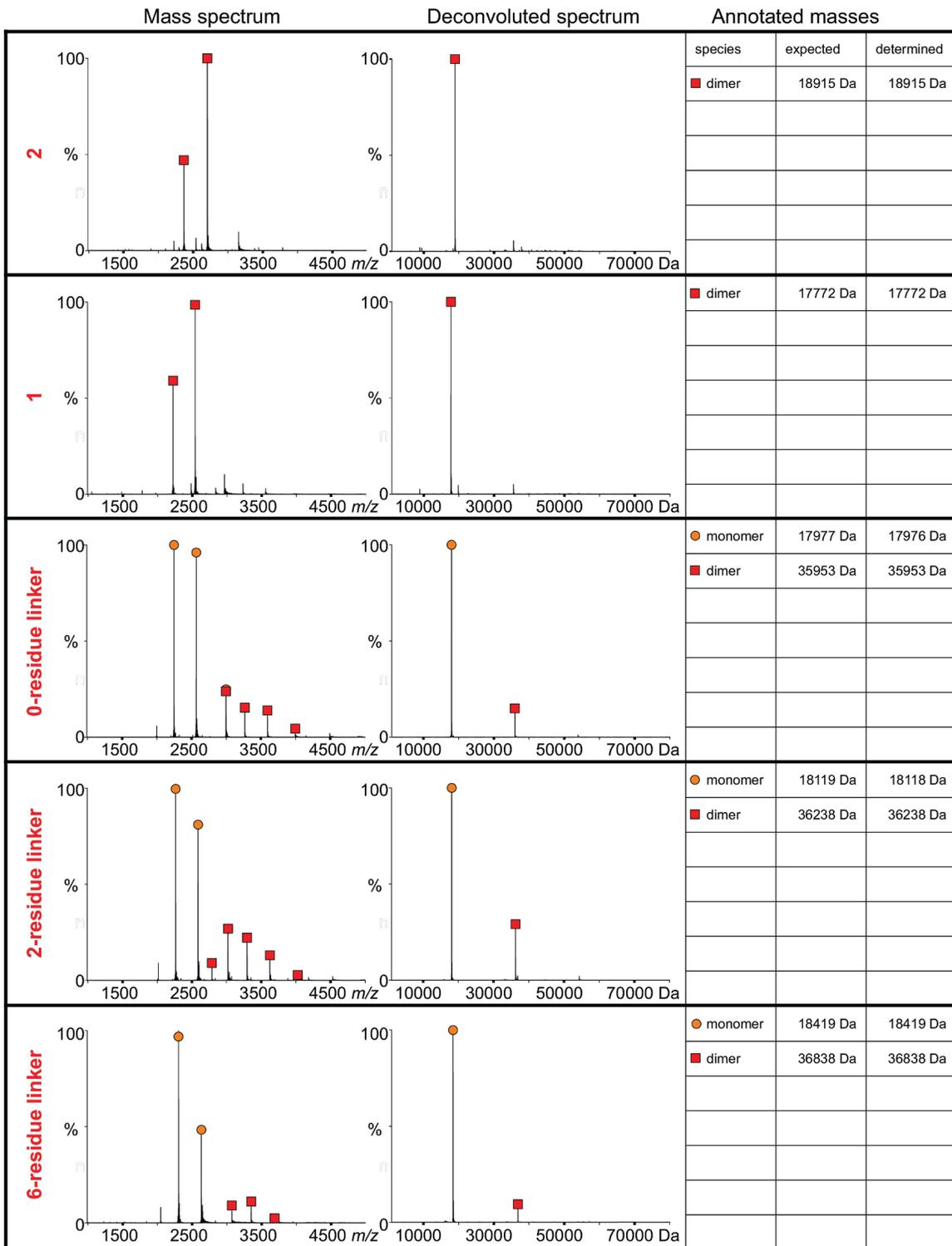
Table 7.2. SAXS analysis statistics. R_g is the radius of gyration, R_c is the cross-sectional radius of gyration determined from Guinier fitting, and P_x is the Porod exponent. The R_c value for most designs cluster around 12 Å, in a close agreement with design models.

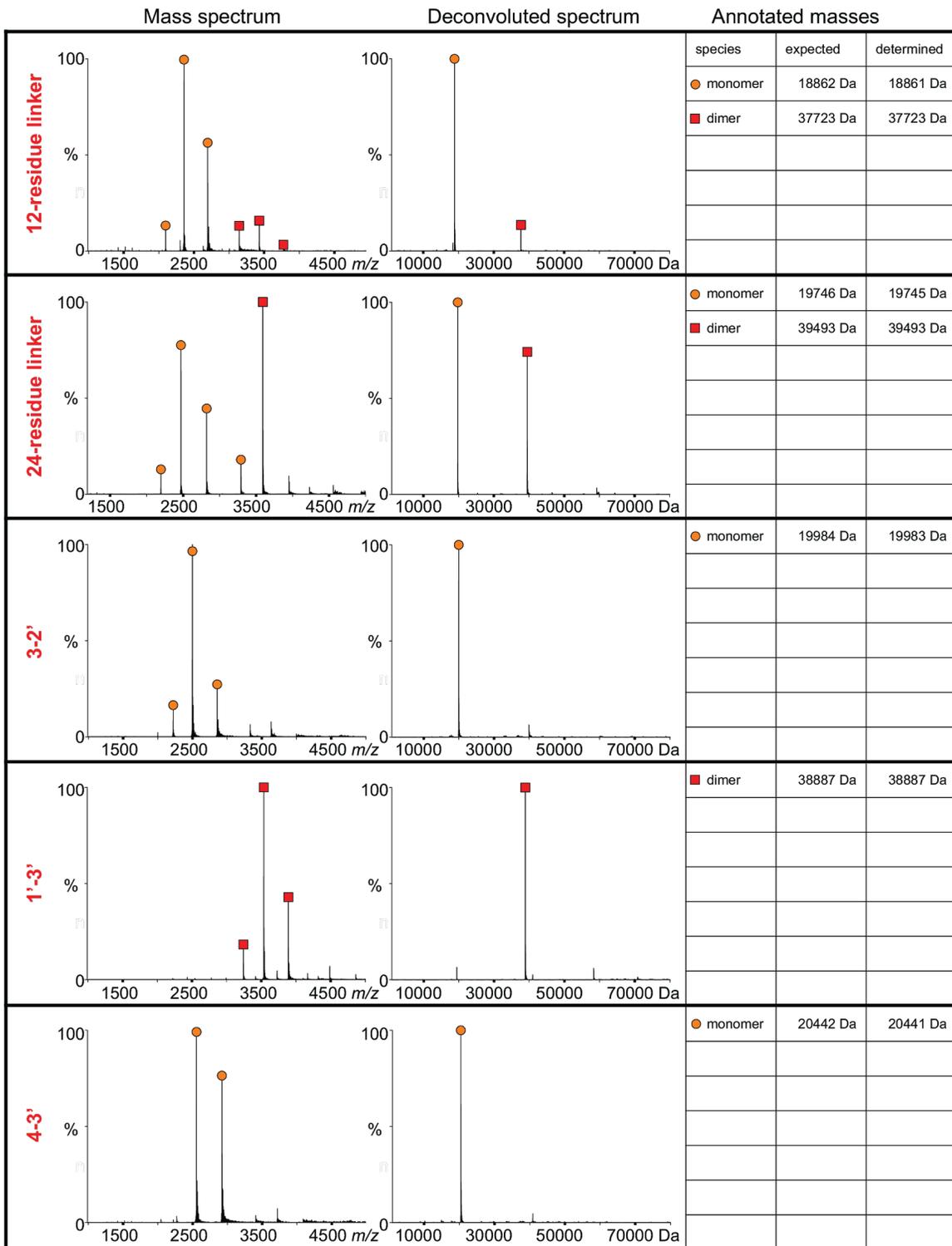
| Design name | $I(0)$ (cm ⁻¹) [Reciprocol space] | $I(0)$ (cm ⁻¹) [Real space] | R_g [Reciprocol space] | R_g [Real space] | Porod volume estimate (10,000 Å ³) | Dmax (Å) | R_c (Å) | P_x |
|------------------|--|--|--------------------------|--------------------|--|----------|-----------|-------|
| 15B-37B_linker6 | 13.00 | 12.00 | 21.54 | 21.06 | 4.2 | 63 | 12.46 | 3 |
| 15B-37B_linker12 | 13.00 | 12.00 | 20.86 | 20.66 | 3.9 | 62 | 12.04 | 3 |
| 15A | 17.00 | 17.00 | 19.22 | 19.86 | 3.5 | 61 | 12.58 | 3.4 |
| 37A | 17.00 | 14.00 | 22.51 | 21.43 | 4.5 | 65 | 13.3 | 3.1 |
| 15B-131A | 8.30 | 7.60 | 20.67 | 20.91 | 4.1 | 65 | 12.26 | 2.8 |
| 131B-37B | 12.00 | 9.90 | 28.82 | 27.31 | 7.3 | 85 | 15.14 | 2.8 |
| 131B-101A | 12.00 | 11.00 | 21.14 | 21.35 | 4.4 | 65 | 13.72 | 3.1 |

Table 7.3. CD data parameter fitting. Fitted parameters for equilibrium chemical denaturation. Errors represent fitting errors.

| | Linker 0 | Linker 2 | Linker 6 | Linker 12 | Linker 24 |
|--|----------------|----------------|----------------|------------------|----------------|
| $\Delta G_1^{(N \rightleftharpoons I)}$ (kcal mol ⁻¹) | 3.6 (±0.4) | 3.5 (±0.2) | 3.5 (±0.2) | 2.7 (±0.1) | 3.7 (±0.3) |
| $\Delta G_2^{(I \rightleftharpoons D)}$ (kcal mol ⁻¹) | 9.8 (±0.6) | 10.7 (±0.4) | 12.2 (±0.4) | 10.6 (±0.5) | 10.4 (±0.8) |
| $\Delta G_{tot}^{(N \rightleftharpoons D)}$ (kcal mol ⁻¹) | 13.5 (±0.7) | 14.1 (±0.4) | 15.7 (±0.5) | 13.3 (±0.5) | 14.1 (±0.8) |
| m_1 (kcal mol ⁻¹ M ⁻¹) | 1.1 (±0.2) | 1.0 (±0.1) | 0.9 (±0.1) | 0.75 (±0.05) | 1.1 (±0.1) |
| m_2 (kcal mol ⁻¹ M ⁻¹) | 1.8 (±0.1) | 1.97 (±0.07) | 2.22 (±0.08) | 1.96 (±0.08) | 2.0 (±0.1) |
| m_{tot} (kcal mol ⁻¹ M ⁻¹) | 2.9 (±0.2) | 3.0 (±0.1) | 3.1 (±0.1) | 2.71 (±0.09) | 3.1 (±0.2) |
| MRE_N (deg cm ² dmol ⁻¹) | -23,574 (±114) | -27,561 (±84) | -24,712 (±63) | -33,849 (±131) | -26,438 (±123) |
| MRE_I (deg cm ² dmol ⁻¹) | -16,330 (±749) | -18,139 (±540) | -14,779 (±710) | -17,362 (±1,158) | -15,567 (±914) |
| MRE_D (deg cm ² dmol ⁻¹) | -525 (±107) | -785 (±82) | -937 (±68) | -1,104 (±99) | -1,125 (±133) |

Table 7.4. Native MS spectra of individual designs. 5 μ L protein (10-100 pmol) were injected and detected by online buffer exchange coupled to an Exactive Plus EMR Orbitrap instrument. Spectra were deconvoluted by UniDec. Species and determined molecular weights are depicted. Signals corresponding to monomers are indicated by an orange circle, dimers are indicated by a red square.





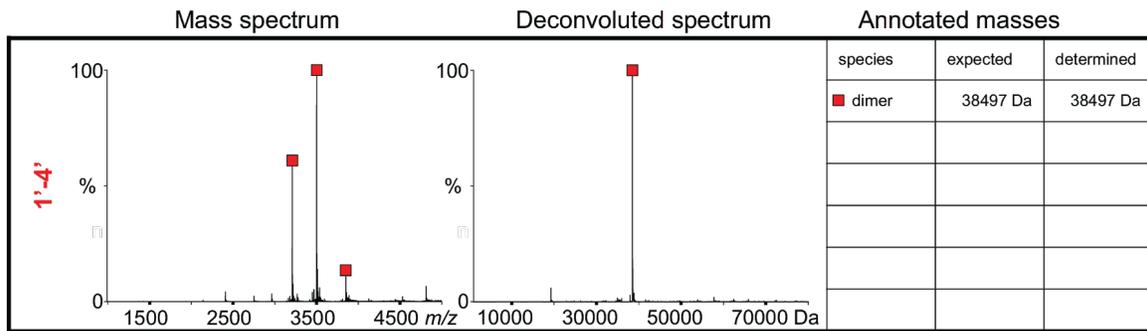


Table 7.5. Protein concentrations native MS titration experiments.

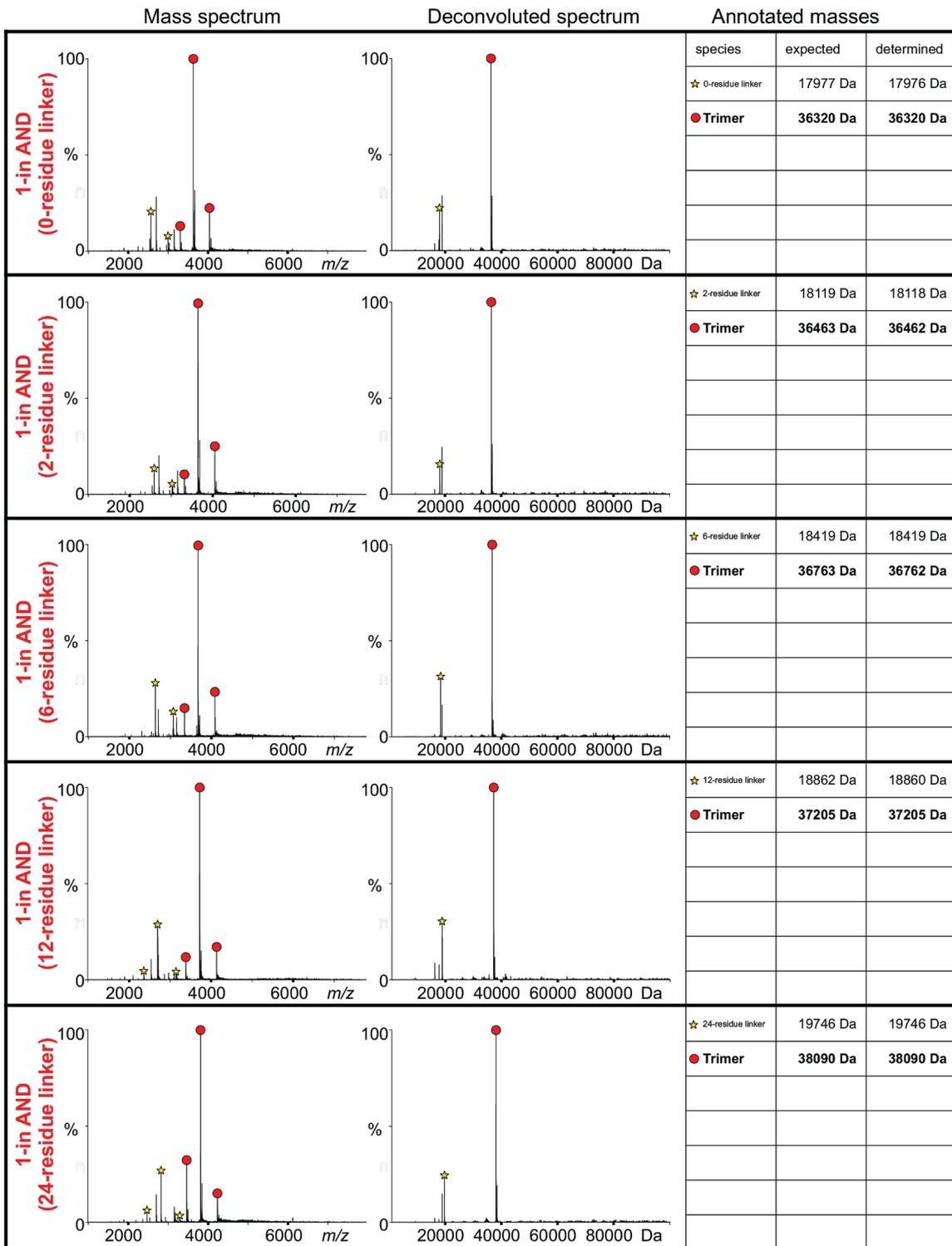
| Titration for Figure 2.1G | | | | | | |
|---------------------------|------------------------------|----|----|----|----|----|
| Design | concentration/ μM | | | | | |
| 2 | 0 | 3 | 6 | 10 | 15 | 20 |
| 1 | 10 | 10 | 10 | 10 | 10 | 10 |
| 1'-2' | 10 | 10 | 10 | 10 | 10 | 10 |

| Titration for Figure 2.1H | | | | | | | | |
|---------------------------|------------------------------|----|----|----|----|----|----|----|
| Design | concentration/ μM | | | | | | | |
| 2 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 1 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 1'-2' | 0 | 3 | 6 | 10 | 15 | 20 | 40 | 60 |

| Titration for Figure 2.1L | | | | | | | | |
|---------------------------|------------------------------|----|----|----|----|----|----|----|
| Design | concentration/ μM | | | | | | | |
| 2 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 1 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 3-2' | 0 | 3 | 6 | 10 | 15 | 20 | 40 | 60 |
| 1'-3' | 0 | 3 | 6 | 10 | 15 | 20 | 40 | 60 |

| Titration for Figure 2.1N | | | | | | | | |
|---------------------------|------------------------------|----|----|----|----|----|----|----|
| Design | concentration/ μM | | | | | | | |
| 2 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 1 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 1'-4' | 0 | 3 | 6 | 10 | 15 | 20 | 40 | 60 |
| 4-3' | 0 | 3 | 6 | 10 | 15 | 20 | 40 | 60 |
| 3-2' | 0 | 3 | 6 | 10 | 15 | 20 | 40 | 60 |

Table 7.6 Native MS spectra of complexes. Samples were prepared by mixing equimolar amounts of **1**, **2**, and the corresponding dimerizer(s) (i.e. **1'-2'** for induced dimerization, **1'-3' + 3-2'** for 2-in AND, linker **1'-4+ 4-3'+ 3-2'+** for 3-in AND). After denaturation, refolding and dialysis against 200 mM AmAc, the mixtures were diluted 12.5 fold with 200 mM AmAc and 5 μ L were flow-injected into an Exactive Plus EMR Orbitrap instrument. Spectra were deconvoluted by UniDec. Species and determined molecular weights are depicted. Signals corresponding to the 1-(dimerizer)_n-2 complex are indicated by a red circle, free dimerizer is indicated by a yellow star.



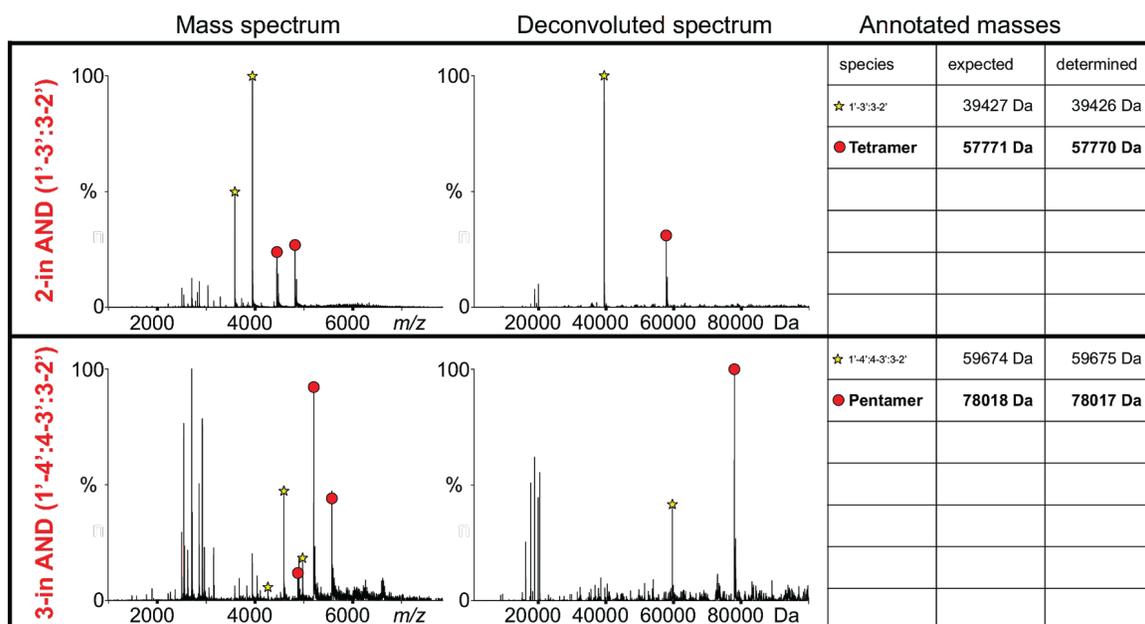
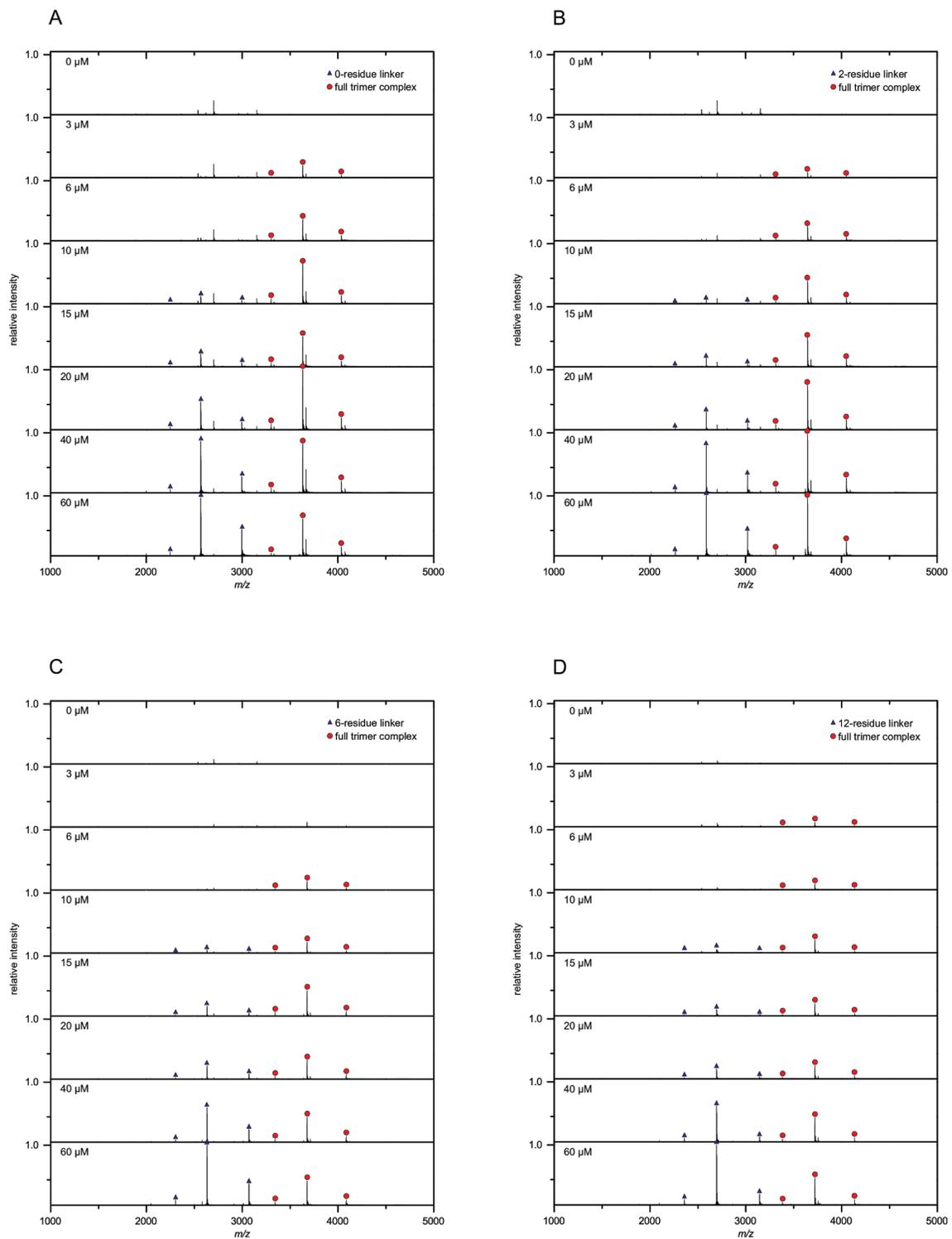


Table 7.7. Detailed native MS titration spectra of AND logic gates. Samples were prepared by mixing 10 μ M **1** and **2** with 0 – 60 μ M of (A) **1'-2'** with 0-residue linker, (B) **1'-2'** with 2-residue linker, (C)

1'-2' with 6-residue linker, (D) **1'-2'** with 12-residue linker, (E) **1'-2'** with 24-residue linker, (F) **1'-3' + 3-2'**, and (G) **1'-4+ 4-3'+ 3-2'**. After denaturation, refolding and dialysis against 200 mM AmAc, the mixtures were diluted 12.5 fold with 200 mM AmAc and 5 μ L were flow-injected into an Exactive Plus EMR Orbitrap instrument. Signals corresponding to the full trimer complexes in (A) - (E), full tetramer complex in (F) and full pentamer complex in (G) are indicated by a red circle. Free dimerizer/ dimerizer complexes are indicated by a blue triangle. The spectra for each titration were normalized and stacked to visualize the change in complex abundance in dependency of added dimerizer(s). As indicated by a red arrow and dashed line, in (G) relative intensities for signals with $m/z > 4300$ are magnified 4x, and in (E) relative intensities for signals with $m/z > 4200$ are magnified 5x for $m/z > 4200$ to help visualize formation of full complexes in the presence of the easily ionizable dimerizers. Lists of all identified and quantified species in the titration has been deposited.



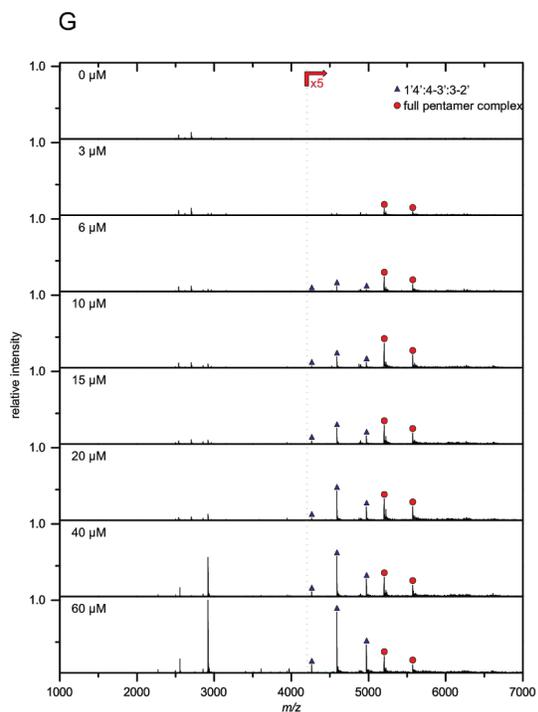
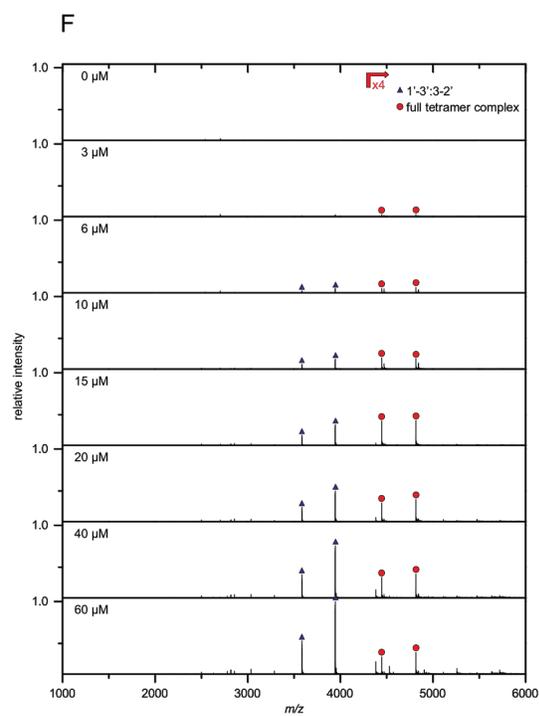
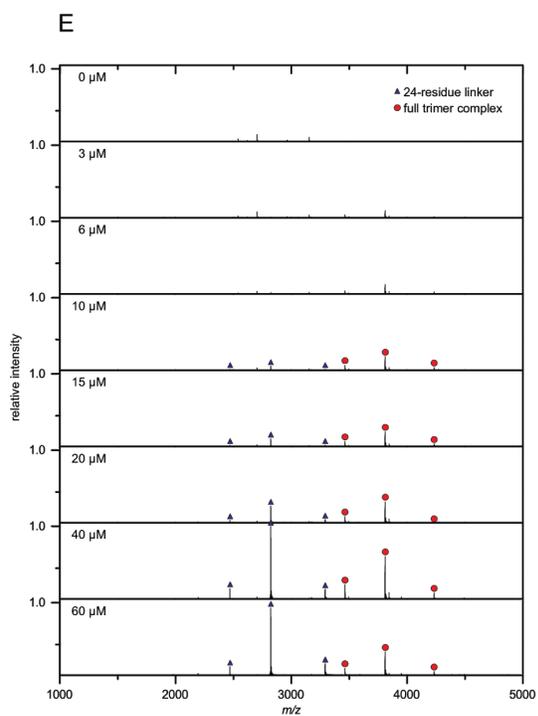
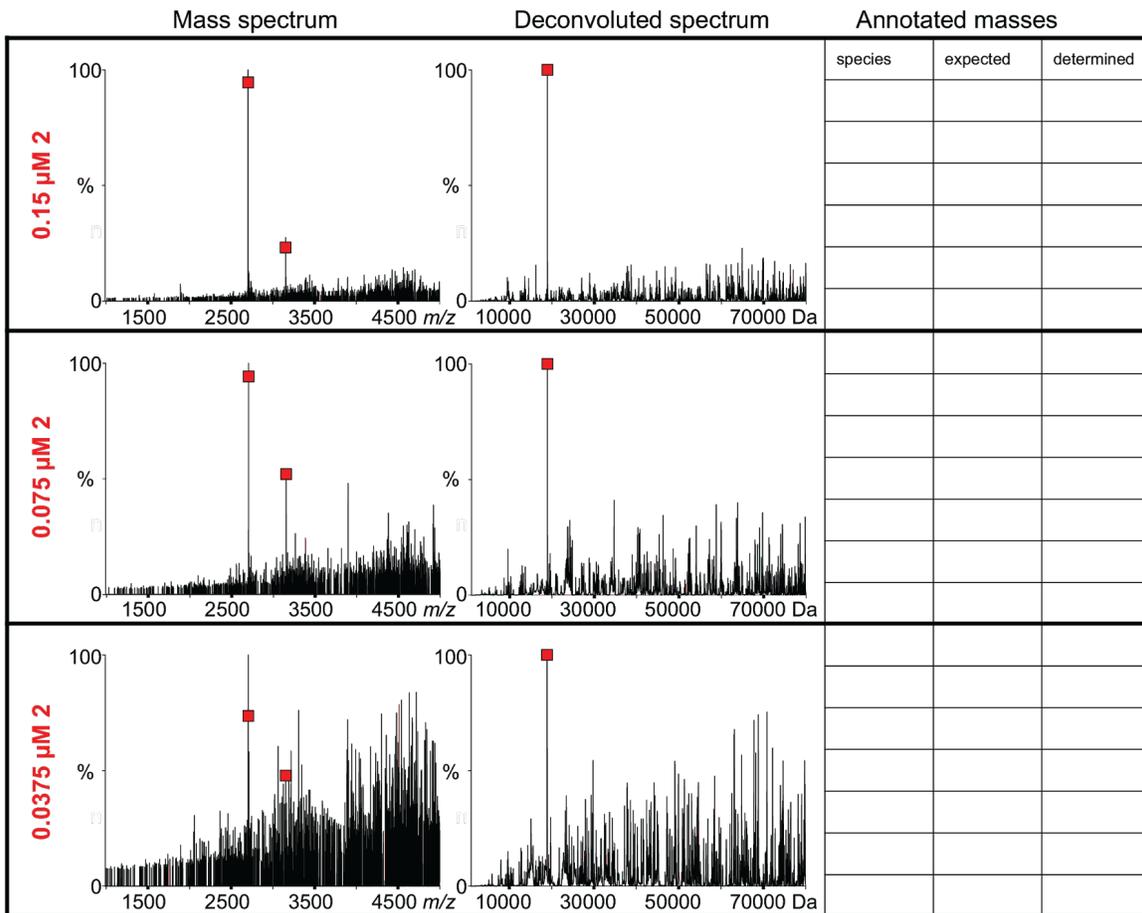
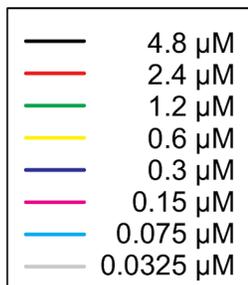
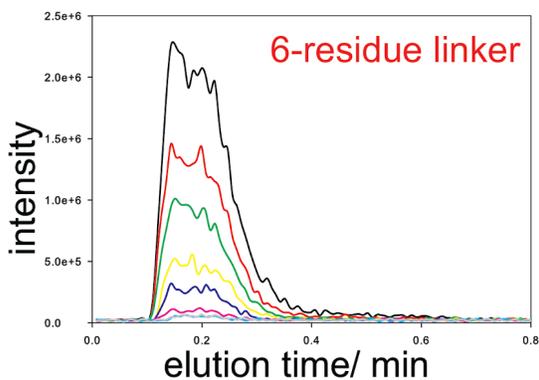
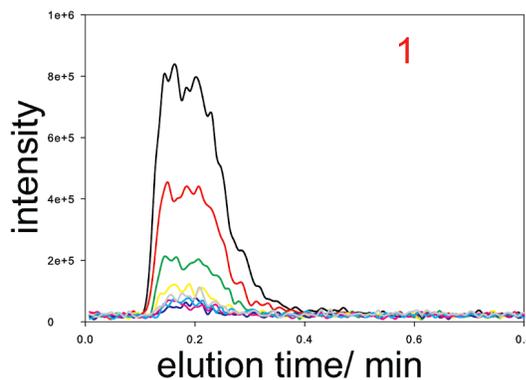
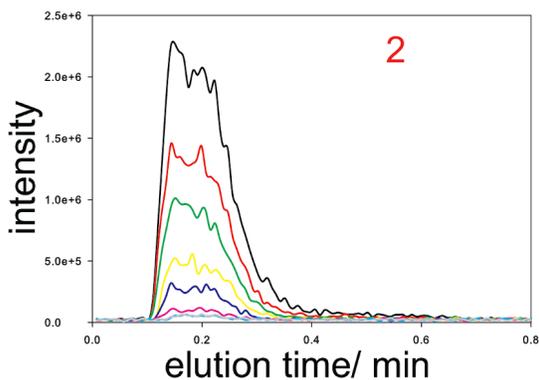
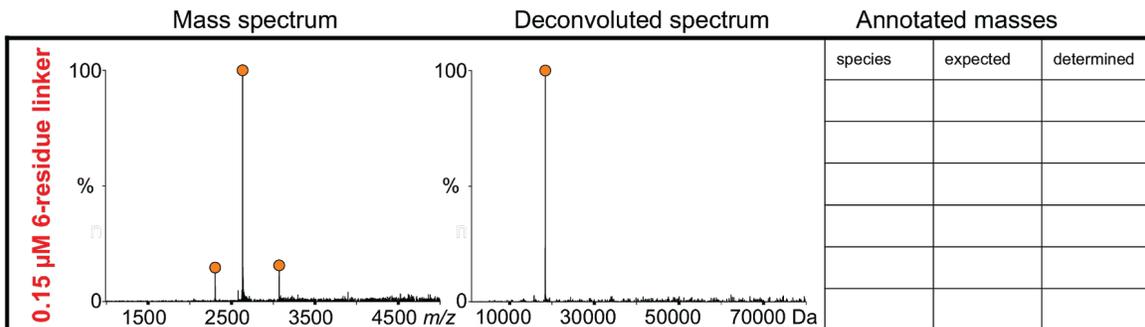


Table 7.8. Native MS spectra of serial dilution experiments. Samples at 60 μM monomer concentration were prepared by denaturation, refolding and dialysis against 200 mM AmAc. Proteins were diluted to the indicated concentrations (4.8 μM to 0.0375 μM). 5 μL (24 pmole to 0.1875 pmole) were flow-injected into an Exactive Plus EMR Orbitrap instrument. Spectra were deconvoluted by UniDec. Species and determined molecular weights are depicted. Signals corresponding to monomers are indicated by an orange circle, dimers are indicated by a red square. Total ion chromatograms (1000-8000 m/z) are depicted for flow-injection profiles.





Chapter 8 - Appendix Containing Supplementary Information for Chapter 3

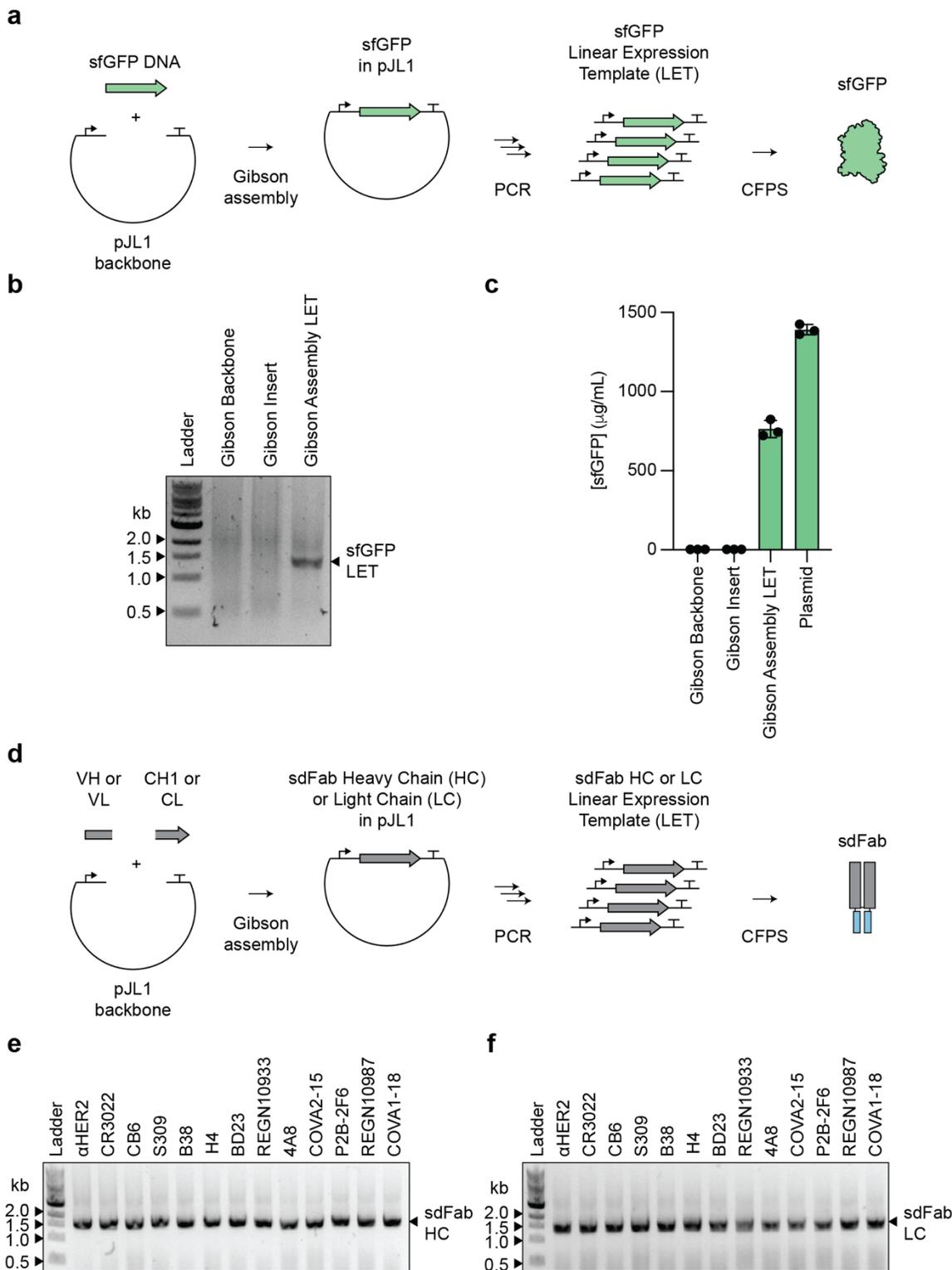


Figure 8.1. The cell-free DNA assembly and amplification workflow. (A) Schematic of the cell-free DNA assembly and amplification protocol for generating sfGFP linear expression template for CFPS. (B) Agarose gel of amplified LET PCR products of Gibson assembly reactions. Backbone only and insert only conditions included as negative controls. Labeled band indicates assembly and amplification of the correct length PCR product. (C) sfGFP yields in Origami™ B(DE3) CFPS from cell-free assembled linear expression templates and purified plasmid (n = 3 independent replicates ± SEM). (D) Schematic of the cell-free DNA assembly and amplification protocol for generating sdFab linear expression template for CFPS. (E) Agarose gel of amplified sdFab heavy chain (HC) LET PCR products. Labeled bands indicate assembly and amplification of the correct length PCR product. (F) Agarose gel of amplified sdFab light chain (LC) LET PCR products. Labeled bands indicate assembly and amplification of the correct length PCR product.

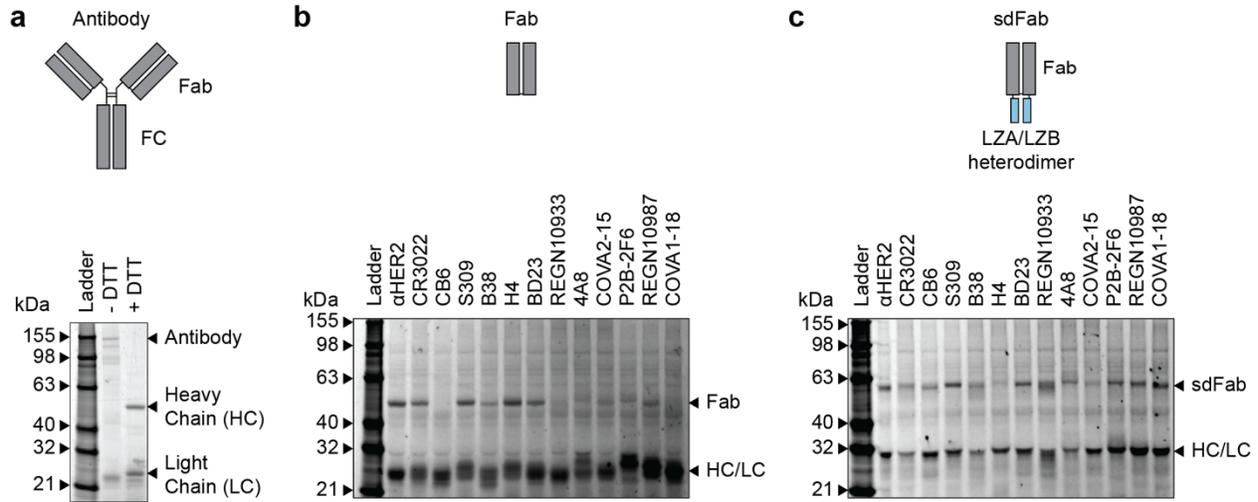


Figure 8.2. Development of an Origami™ B(DE3) CFPS system for the expression of antibodies and antibody fragments. (A-C) SDS PAGE of antibodies and antibody fragments manufactured in CFPS. Samples were fluorescently labeled with the FluoroTect™ reagent during protein synthesis. (A) Expression and assembly of full-length Trastuzumab (α HER2). The antigen-binding fragment (Fab) and fragment, crystallizable (FC) domains are labeled. (B) Expression and assembly of a panel of 13 Fabs. (C) Expression and assembly of a panel of 13 sdFabs. The leucine zipper heterodimer (LZA and LZB) assisting with Fab assembly are labeled.

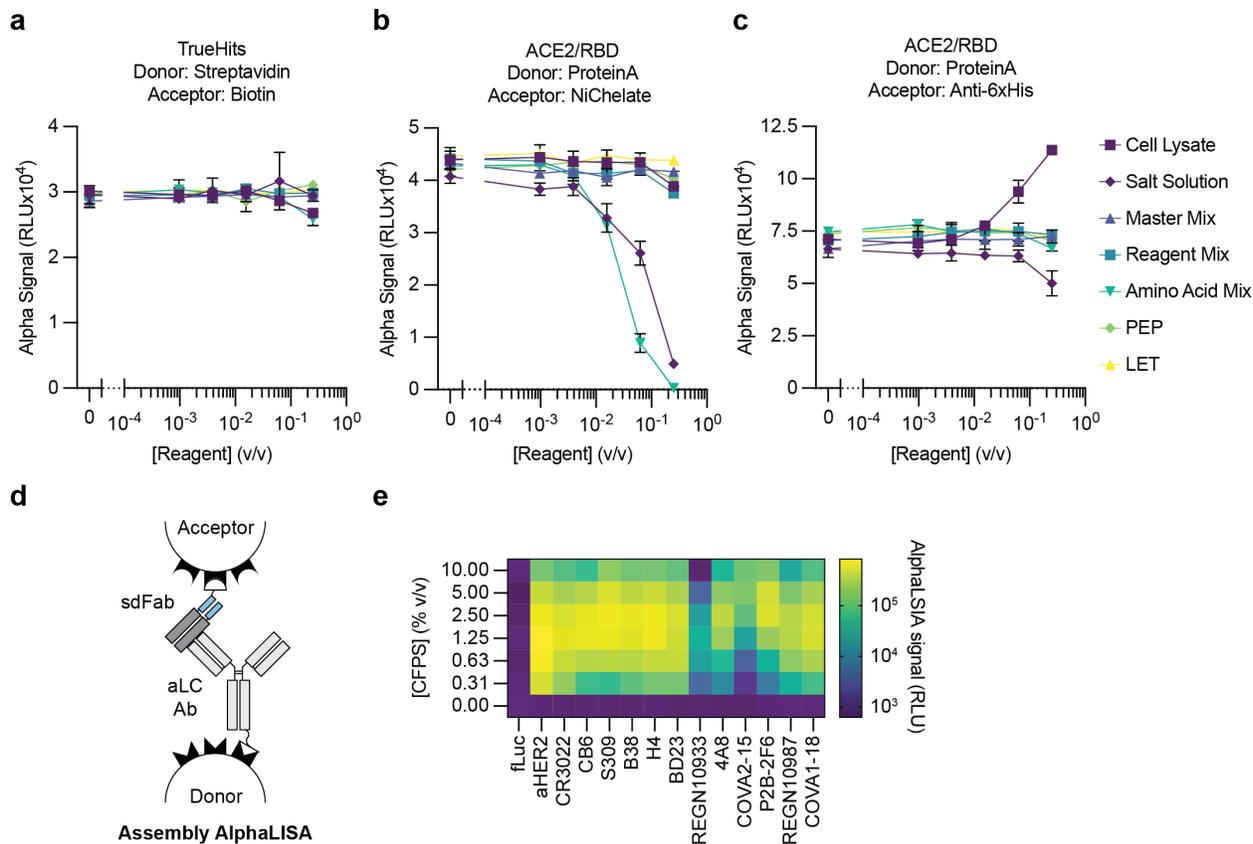


Figure 8.3. AlphaLISA for profiling antibody protein-protein interactions in CFPS. (A-C) Evaluation of the effect of CFPS reagents on AlphaLISA. Concentrations are plotted as v/v fraction of the final concentration of the reagent in a CFPS reaction. Reagents were diluted in water at the concentration they normally reside at in CFPS. Reagents were tested in mixtures that were used to assemble CFPS reactions. The salt solution contains 8 mM magnesium glutamate, 10 mM ammonium glutamate, and 130 mM potassium glutamate. Master Mix contains 1.2 mM ATP, 0.85 mM GTP, 0.85 mM UTP, 0.85 mM CTP, 0.03 mg/mL folinic acid, and 0.17 *E. coli* tRNA. Reagent Mix contains 0.4 mM NAD, 0.27 mM CoA, 4 mM oxalic acid, 1 mM putrescine, 1.5 mM spermidine, and 57 mM HEPES. Amino Acid Mix contains 2 mM of all 20 amino acids. PEP is 30 μ M phosphoenolpyruvate. LET is 0.066 v/v fraction unpurified PCR mix containing the LET for sfGFP. (A) Evaluation of the effect of CFPS reagents on AlphaLISA detection chemistry using the TrueHits kit. Biotin and Streptavidin labeled beads associate directly with one another and serve as a control for reagents impacting the AlphaLISA measurement chemistry. (B) Evaluation of the effect of CFPS reagents on AlphaLISA detection of the SARS-CoV-2 RBD and ACE2 interaction measured by the Protein A donor bead and Ni Chelate acceptor bead. The Salt Solution and Amino Acids Mix inhibit immobilization of his-tagged proteins on the NiChelate bead. (C) Evaluation of the effect of CFPS reagents on AlphaLISA detection of the SARS-CoV-2 RBD and ACE2 interaction measured by the Protein A donor bead and anti-6xhis acceptor bead. The Salt Solution and Amino Acids Mix inhibit immobilization of his-tagged proteins on the NiChelate bead. (D) Schematic of AlphaLISA setup for measuring sdFab assembly. (E) Assembly AlphaLISA measurement of a panel of sdFabs. AlphaLISA signal is indicative of sdFab assembly, though the signal is subject to the hook effect⁵¹ resulting in a lower signal at higher concentrations.

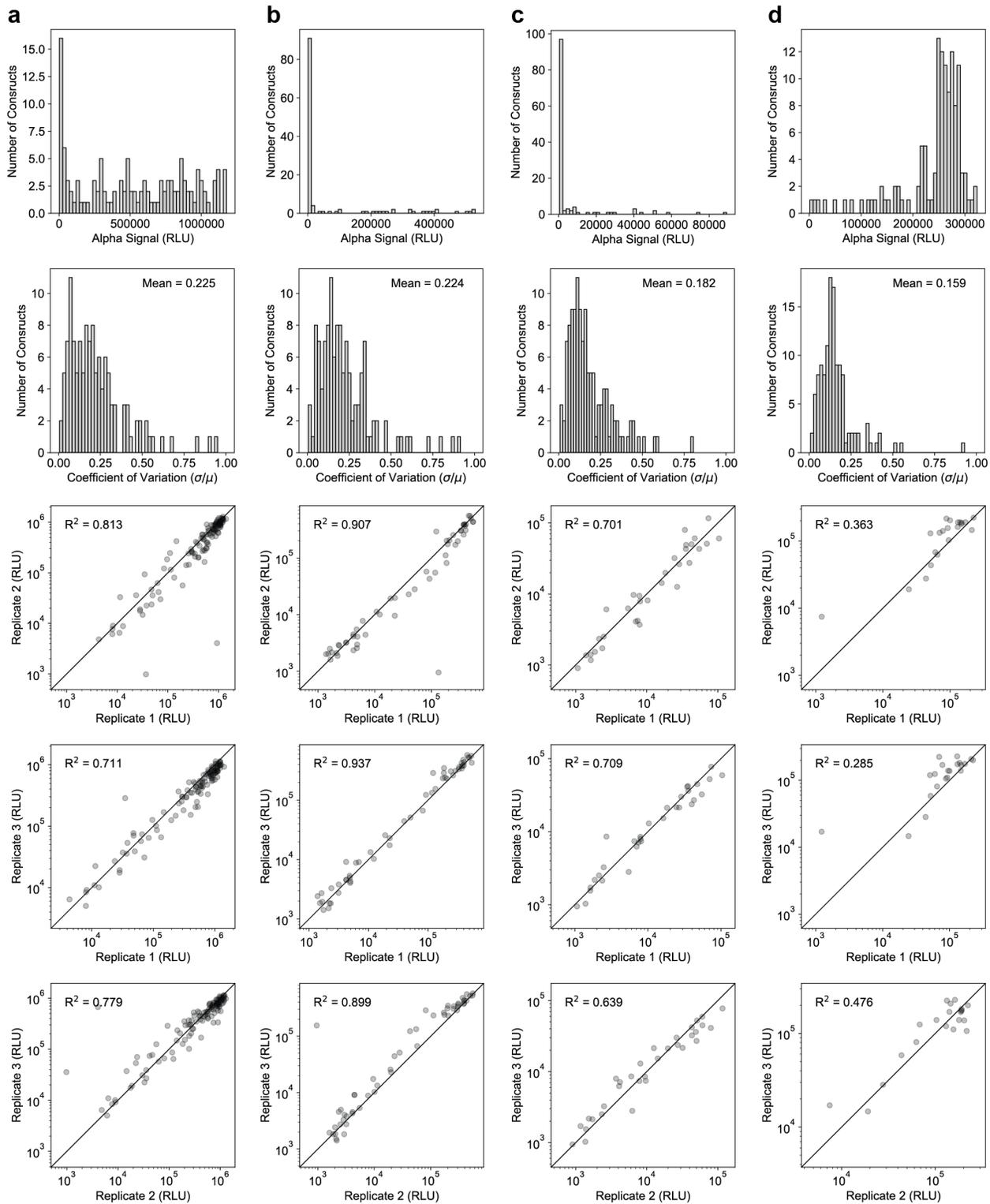


Figure 8.4. Analysis of variability in AlphaLISA replicates. (A-D) From top to bottom: Histogram of raw AlphaLISA values (mean of $n = 3$ independent replicates) to visualize the spread of the data. A histogram of coefficient of variation (standard deviation divided by the mean) to visualize the typical error within a sample with the mean coefficient of variation displayed on the plot. Parity plots of the three replicates were fit to the line $y=x$ to visualize the consistency of replicates with the corresponding R^2 value is displayed on the chart. Only values found to be significantly different from the background are plotted ($p < 0.05$, two-sided t-test adjusted for multiple comparisons using FDR with a family-wise error rate of 5%) (A) sdFab assembly AlphaLISA. (B) SARS-CoV-2 S trimer binding AlphaLISA. (C) SARS-CoV-2 RBD binding AlphaLISA. (D) sdFab competition with ACE2 for the SARS-CoV-2 RBD AlphaLISA.

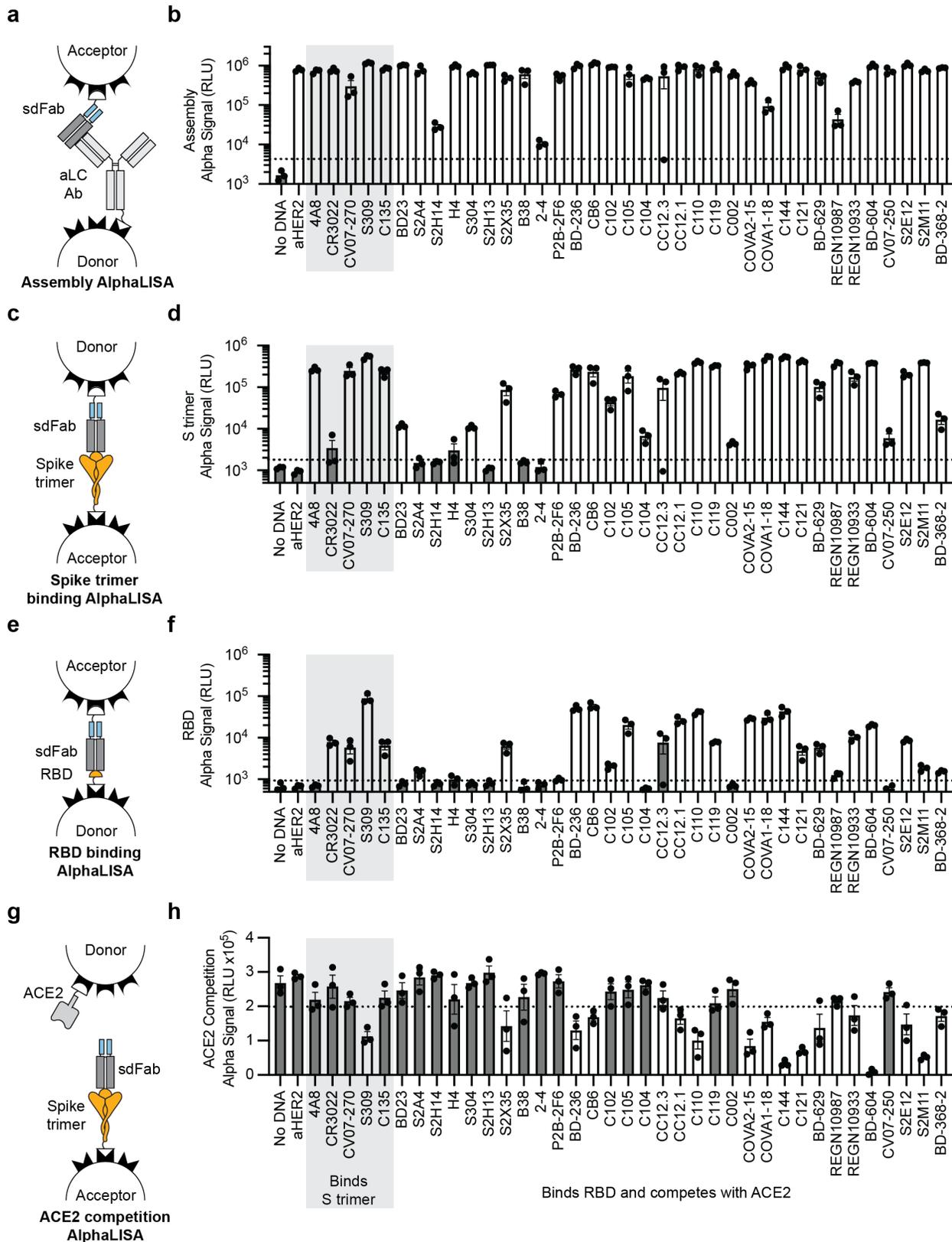


Figure 8.5. AlphaLISA profiling of 38 published antibodies. Data are a more quantitative depiction of the data for 38 of the antibodies in the heatmaps in Fig. 3.2a-b. (A) Schematic depicting AlphaLISA setup for measuring sdFab assembly. (B) AlphaLISA measurement of sdFab assembly. (C) Schematic depicting AlphaLISA setup for measuring S trimer binding. (D) AlphaLISA measurement of sdFab binding to the SARS-CoV-2 S trimer. (E) Schematic depicting AlphaLISA setup for measuring RBD binding. (F) AlphaLISA measurement of sdFab binding to the SARS-CoV-2 RBD. (G) Schematic depicting AlphaLISA setup for measuring sdFab competition with ACE2 for the RBD. (H) AlphaLISA measurement of sdFab competition with ACE2 for the RBD. All AlphaLISA data are the mean of three independent replicates \pm the SEM. The dashed line indicates three standard deviations away from the background. Samples determined not to be significantly distinguished from the background ($p > 0.05$ two-sided t-test corrected using the FDR procedure) have bars that are filled dark grey. The samples are ranked within each category from worst (left) to best (right) neutralizers by their minimum neutralization IC_{50} value.

Table 8.1. Summary of antibody screening studies designed to identify SARS-CoV-2 neutralizing antibodies. Studies were evaluated for their efficiency at identifying potent neutralizing antibodies. For each study, the total number of antibodies evaluated as well as the number of neutralizing antibodies with a neutralization IC_{50} less than 0.01 $\mu\text{g/mL}$ and 0.25 $\mu\text{g/mL}$ were summarized. Either authentic- or pseudovirus neutralization IC_{50} was considered based on the breadth of the reported data. The 0.01 $\mu\text{g/mL}$ cutoff was chosen because this value is approximately within an order of magnitude of the most potent reported neutralizing antibodies³³⁸. 0.25 $\mu\text{g/mL}$ was chosen as a practical cutoff for moderately potent neutralizing antibodies³³⁵. For Hansen *et al.*³⁴⁴ the neutralization potencies were reported in M and were converted to $\mu\text{g/mL}$ assuming an antibody molecular weight of 150 kDa. For this analysis, only the studies containing antibodies expressed in this manuscript were used. Furthermore, only studies whose purpose was to identify neutralizing antibodies from a large set of candidates were considered. Studies that didn't describe the results of their antibody discovery process in sufficient detail to collect the desired information were omitted. NR indicates not reported in sufficient detail to determine.

| Reference | DOI | Antibodies screened | Antibodies $IC_{50} < 0.01$ $\mu\text{g/mL}$ | Antibodies $IC_{50} < 0.25$ $\mu\text{g/mL}$ |
|--------------------------------------|----------------------------|---------------------|--|--|
| Kreye <i>et al.</i> ³³⁵ | 10.1016/j.cell.2020.09.049 | 598 | 5 | 40 |
| Cao <i>et al.</i> ³⁴² | 10.1016/j.cell.2020.05.025 | 216 | 1 | 22 |
| Wu <i>et al.</i> ³⁴¹ | 10.1126/science.abc2241 | 17 | 0 | 1 |
| Liu <i>et al.</i> ³³⁸ | 10.1038/s41586-020-2571-7 | 252 | 9 | NR |
| Ju <i>et al.</i> ³³⁶ | 10.1038/s41586-020-2380-z | 206 | 0 | 3 |
| Brouwer <i>et al.</i> ³³¹ | 10.1126/science.abc5902 | 84 | 2 | 9 |
| Hansen <i>et al.</i> ³⁴⁴ | 10.1126/science.abd0827 | "Thousands" | 6 | NR |
| Robbani <i>et al.</i> ⁴⁹¹ | 10.1038/s41586-020-2456-9 | 94 | 9 | 43 |

Table 8.2. Reported parameters and expected behaviors for tested antibodies. Information about antibody target epitope, pseudo- or authentic virus neutralization IC_{50} , and equilibrium dissociation constant from literature about the antibodies used in this study are summarized. Data are presented in two tabs, one for antibodies from diverse sources and one for the Brouwer *et al.* data set, in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>.

Table 8.3. Variable heavy and light chain sequences of tested antibodies. Sequences are classified by their heavy (VH) or light chain (VL) as well as the light chain class (kappa or lambda). The variable domain protein sequence, the variable domain *E. coli* codon-optimized DNA sequence, and the ordered DNA sequence containing all additional (Gibson assembly homology, n-terminal expression tag, etc.) sequences are listed. Note that the antibodies COVA2-15 and COVA1-18 were ordered and evaluated twice in this data set, and there are thus two separate entries for the two sets of sequences. Sequences are listed in a single tab in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>.

Table 8.4. Summary of observed antibody binding significance statistics and comparison to literature. Results from the AlphaLISA screen are compared against literature data and a qualitative match with the reported values are summarized. Comments on the possible origin of observed inconsistent results are listed for all data with reported neutralization IC₅₀ values less than 10 µg/mL. Data are presented in two tabs, one for antibodies from diverse sources and one for the Brouwer *et al.* data set, in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>.

Table 8.5. Raw and processed data for AlphaLISA measurements for the 120 antibodies evaluated in this manuscript. The individual replicates, average, standard deviation, coefficient of variation, p-value from a two-sided t-test against the background, and the FDC corrected p-value are reported. Note that the antibodies COVA2-15 and COVA1-18 were ordered and evaluated twice in this data set, and there are thus two separate entries for the two sets of sequences. Data are presented in four tabs, one for each AlphaLISA measurement modality, in a separate .xlsx file. These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>.

Chapter 9 - Appendix Containing Supplementary Information for Chapter 4

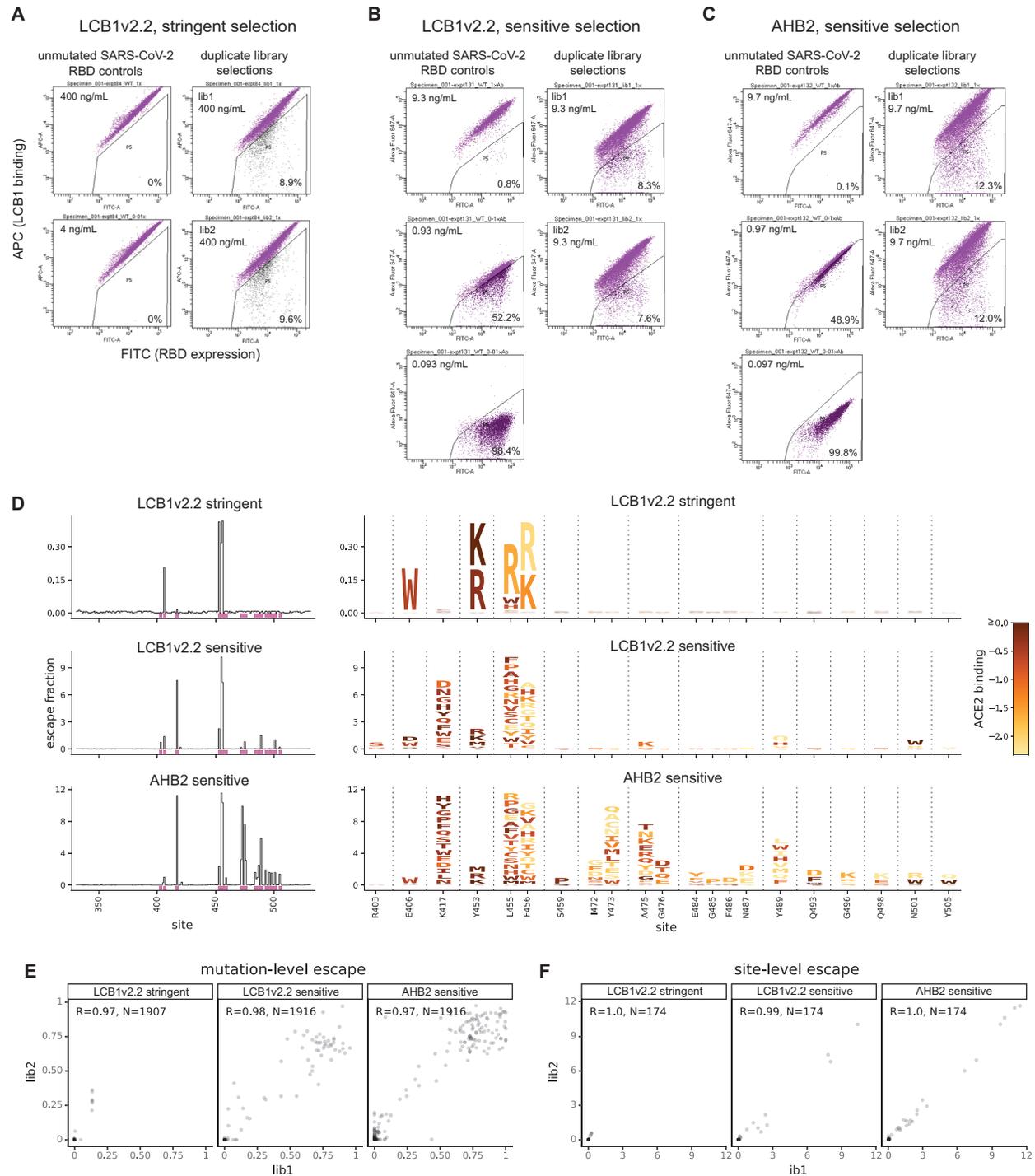


Figure 9.1. Site saturation mutagenesis of RBD to predict MON1 (LCB1v2.2) and MON2 (AHB2) escape mutants. (A-C) FACS gates for deep mutational scanning selections. Duplicate yeast display libraries expressing all mutations tolerated in the SARS-CoV-2 RBD were labeled with LCB1_v2.2 (A-B) or

AHB2 (C). The population of cells indicated escaping minibinder binding was sorted and deep sequenced, enabling calculation of the “escape fraction” of each RBD mutation (the fraction of yeast expressing that mutant that were found in the escape bin). (D) Miniprotein escape profiles averaged across duplicate deep mutational scanning library selections. Line plots, left, show the total escape (sum of per-mutation escape) at all sites in the RBD. Sites marked in pink are shown in logoplots, right. Logoplots illustrate per-mutation escape by the height of letters reflecting individual amino acid mutations. Letters are colored according to their previously measured effects on ACE2-binding affinity³⁸⁹, according to the scale bar at right. (E-F) Correlation in escape between duplicate libraries, at the level of individual mutations (E) or total escape per site (F).

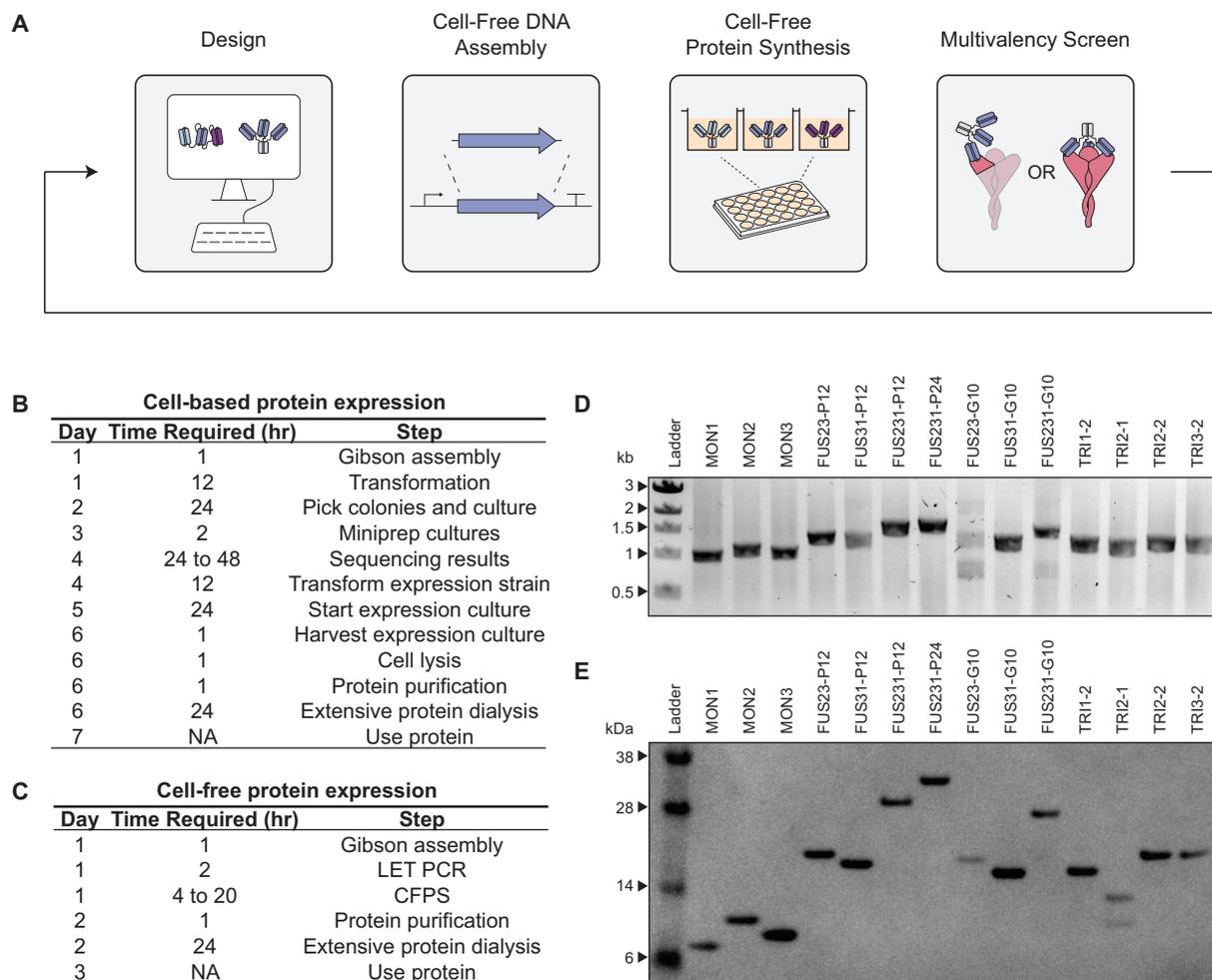


Figure 9.2. Cell-free DNA assembly and protein synthesis of multivalent minibinders. (A) A cell-free workflow for the expression and evaluation of multivalent minibinders. (B) Step-by-step workflow for cell-based DNA assembly and protein expression. (C) Standard step-by-step workflow for cell-free DNA assembly and protein expression. (D) Agarose gel of linear DNA templates for CFPS assembled via Gibson assembly and amplified via PCR. (E) SDS-PAGE of purified proteins expressed via CFPS.

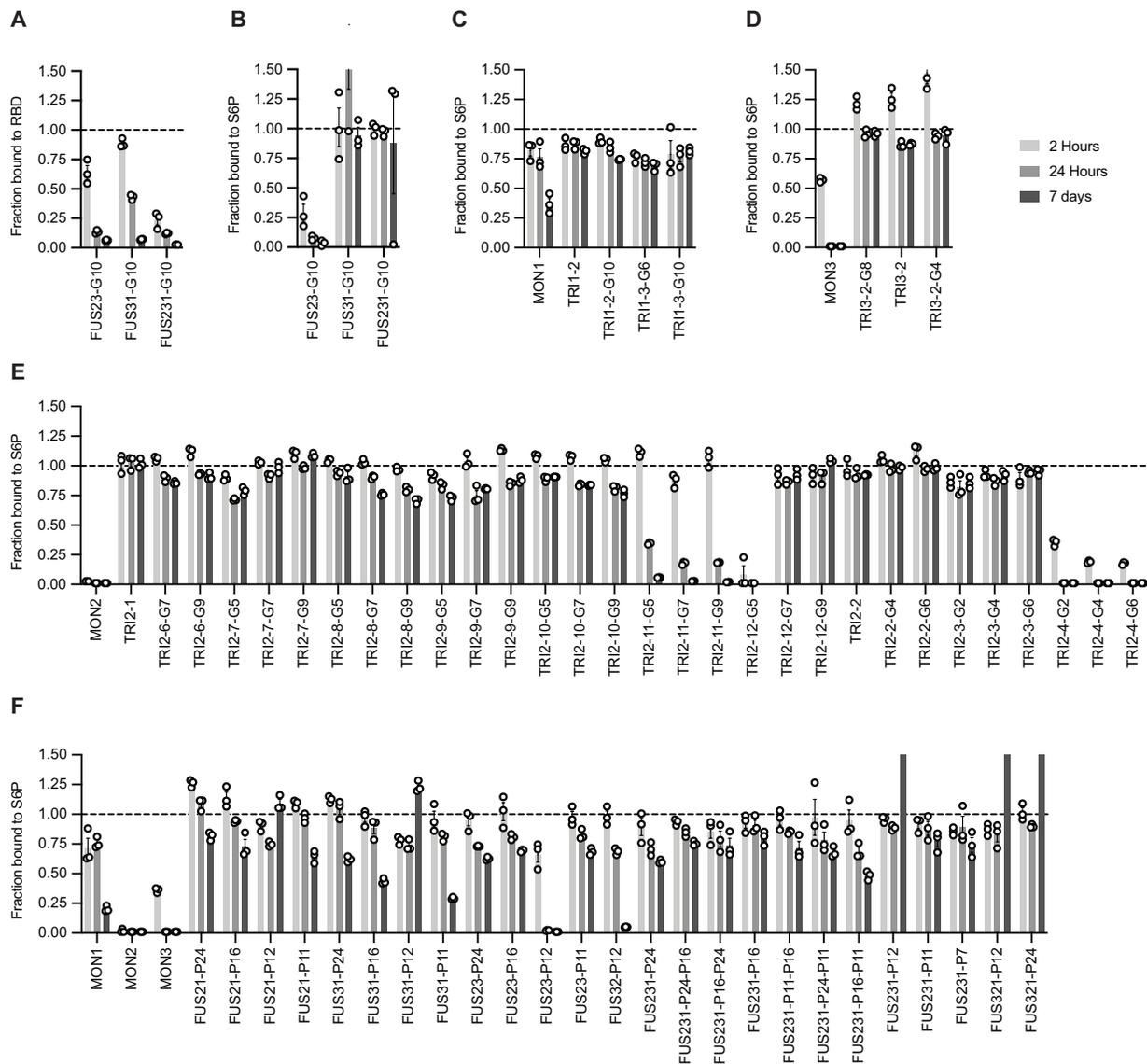


Figure 9.3. Off-rate screen of multivalent minibinders. Dissociation of minibinder constructs from RBD or S6P over time was measured via competition AlphaLISA (mean \pm SEM, $n=3$ technical replicates from a single independent experiment). (A) Dissociation of FUS proteins with G10 linkers from RBD measured via competition with MON1. (B) Dissociation of FUS proteins with G10 linkers from S6P measured via competition with MON1. (C) Dissociation of TRI1 proteins from S6P measured via competition with TRI2-1. (D) Dissociation of TRI3 proteins from S6P measured via competition with TRI2-1. (E) Dissociation of TRI2 proteins from S6P measured via competition with TRI2-1. (F) Dissociation of FUS proteins from S6P measured via competition with FUS231-P24.

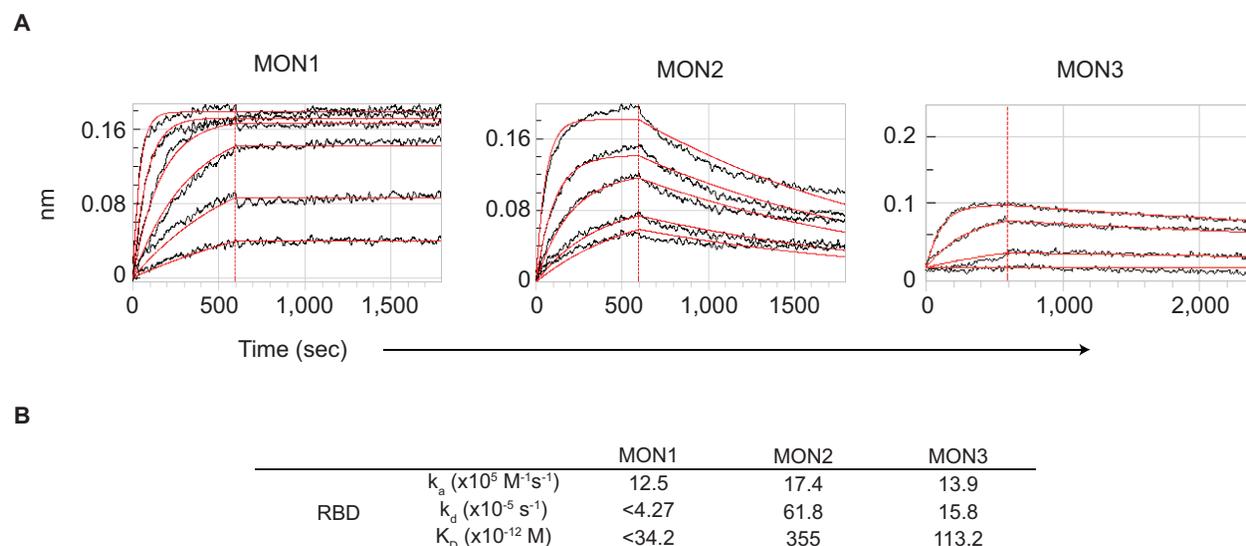


Figure 9.4. Kinetics of MON1, MON2, and MON3 binding to the SARS-CoV-2 RBD. Kinetic curves of Minibinders MON1 at 2-fold dilutions from 20 – 0.625 nM, MON2 at 2-fold dilutions from 10 – 0.625 nM, and MON3 at 3-fold dilutions from 20 – 0.25 nM were analyzed against monovalent soluble RBD covalently linked to AR2G biosensors. (A) Sensorgrams of minibinders binding to RBD ($n = 1$ technical replicate from a single experiment). Red curves are fits obtained from 1:1 Langmuir binding model. Sensorgram images were copied directly from export files from ForteBio Data Analysis v11.0 software without modification. (B) Summary of determined kinetic and equilibrium binding parameters for each minibinder. The < symbol indicates a cut-off of less than 5% dissociation observed during the dissociation phase, indicating insufficient time to accurately quantify the dissociation rate constant. The reported rate constants and equilibrium binding constants are estimates due to the very tight binding of the minibinders studied.

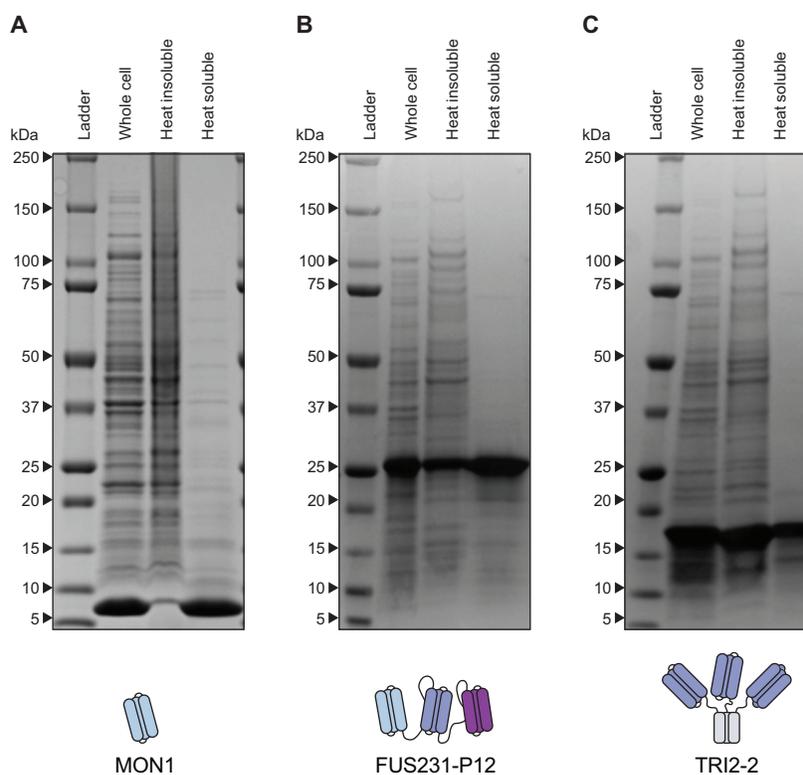


Figure 9.5. Heat-based purification of mono- and multivalent minibinders. Proteins were expressed in *E. coli* using standard protein overexpression procedures. Harvested cells were heat treated at 70 °C for 5 min to lyse cells and precipitate cellular proteins. Bands observed in the Heat insoluble lane are indicative of protein precipitated during the heat treatment step. Bands observed in the Heat soluble step are indicative of proteins that did not precipitate during the heat treatment step. Data are from a single experiment. (A) SDS PAGE of MON1. (B) SDS PAGE of FUS231-P12. (C) SDS PAGE of TRI2-2.

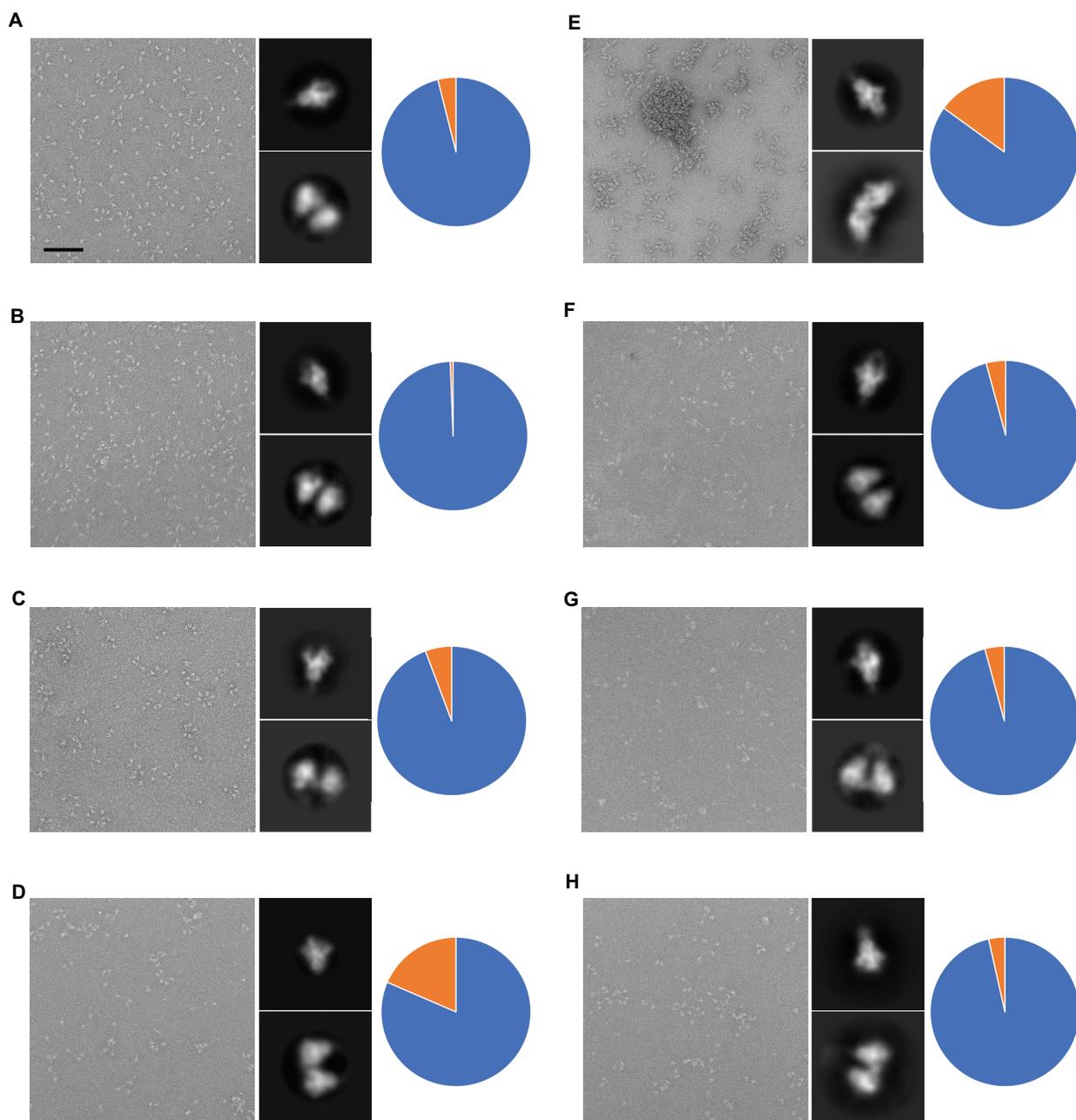


Figure 9.6. Negative stain EM analysis of minibinder-mediated crosslinking/aggregation of S trimers. Representative electron micrograph and 2D class averages of SARS-CoV-2 S in complex with TRI2-2 (A), FUS231-G10 (B), FUS231-P24 (C), FUS31-G8 (D), TRI1-5-G2 (E), TRI1-5-G4 (F), TRI1-5-G6 (G) or FUS31-G10 (H). Scale bar: 100 nm. After two-dimensional classification, the number of particles assigned to classes with well-defined single (blue) or two neighboring cross-linked (orange) S trimers are presented as pie charts. The fraction of total cross-linked S trimers is underestimated since higher-order cross-linked trimers did not yield well-defined 2D averages.

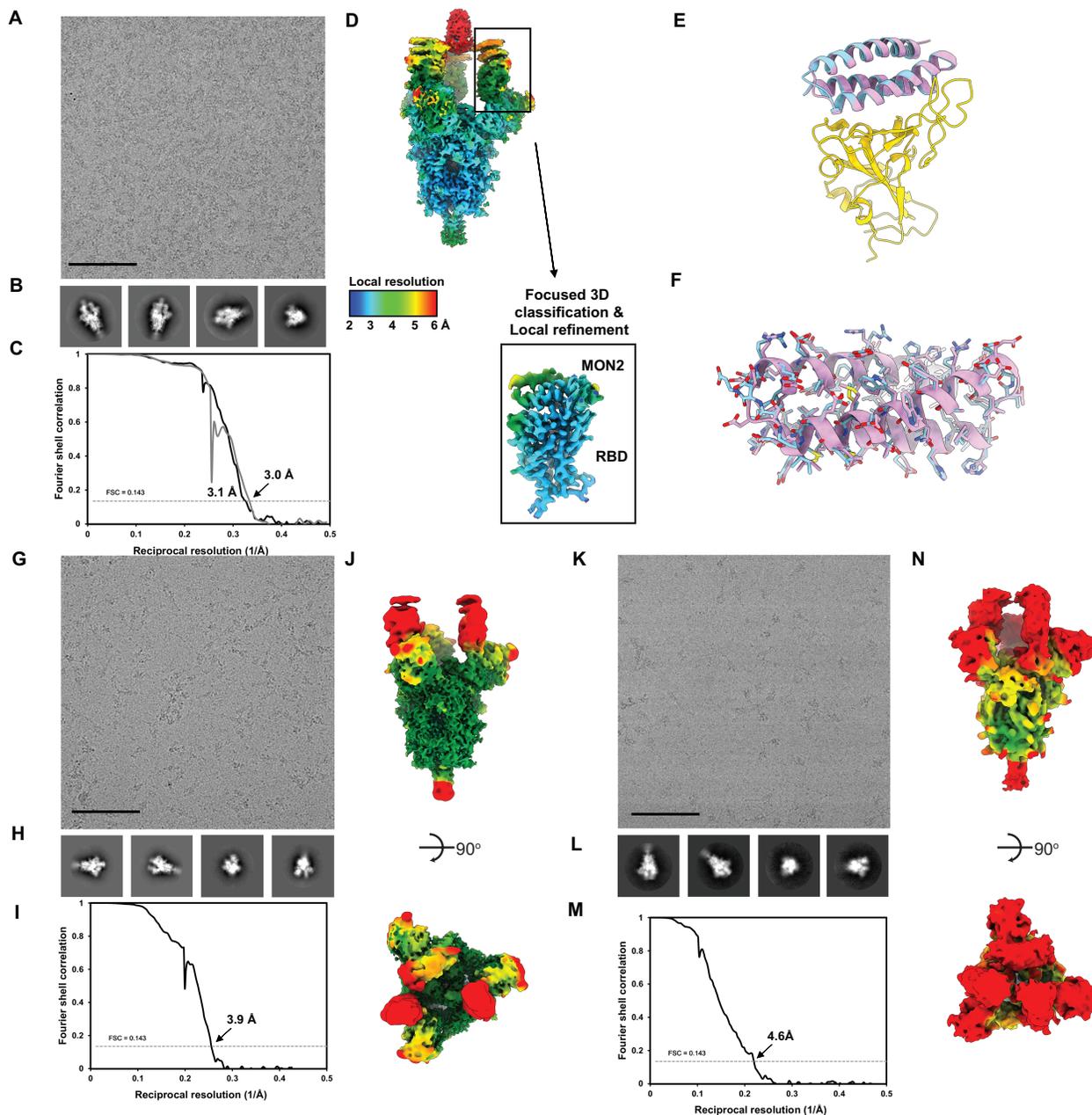


Figure 9.7. Cryo-EM data collection and processing of the S6P/TRI2-2, S6P/FUS31-G10, and S6P/FUS231-P24 datasets. (A-D) Representative electron micrograph and 2D class averages of SARS-CoV-2 S in complex with TRI2-2 embedded in vitreous ice. Scale bar: 100 nm. (C) Gold-standard Fourier shell correlation curves for TRI2-2-bound SARS-CoV-2 S trimer (black line) and locally refined RBD/TRI2-2 region (gray line). The 0.143 cutoff is indicated by a horizontal dashed line. (D) Local resolution maps calculated using cryoSPARC for the whole reconstruction as well as for the locally refined RBD/MON2 region. (E) Ribbon diagram of the RBD/MON2 designed model (pink) superimposed with the MON2 cryoEM structure (cyan). (F) MON2 designed model (pink) superimposed with the S6P/TRI2-2 cryoEM structure (cyan) with side chains displayed as sticks. (G-J) Representative electron micrograph (G), 2D class averages (H), gold-standard Fourier shell correlation curve (I) and local resolution map (J) of SARS-CoV-2 S in complex with FUS231-P24 embedded in vitreous ice. (K-N) Representative electron micrograph (K), 2D class averages (L), gold-standard Fourier shell correlation curve (M), and local resolution map (N) of SARS-CoV-2 S in complex with FUS31-G10 embedded in vitreous ice.

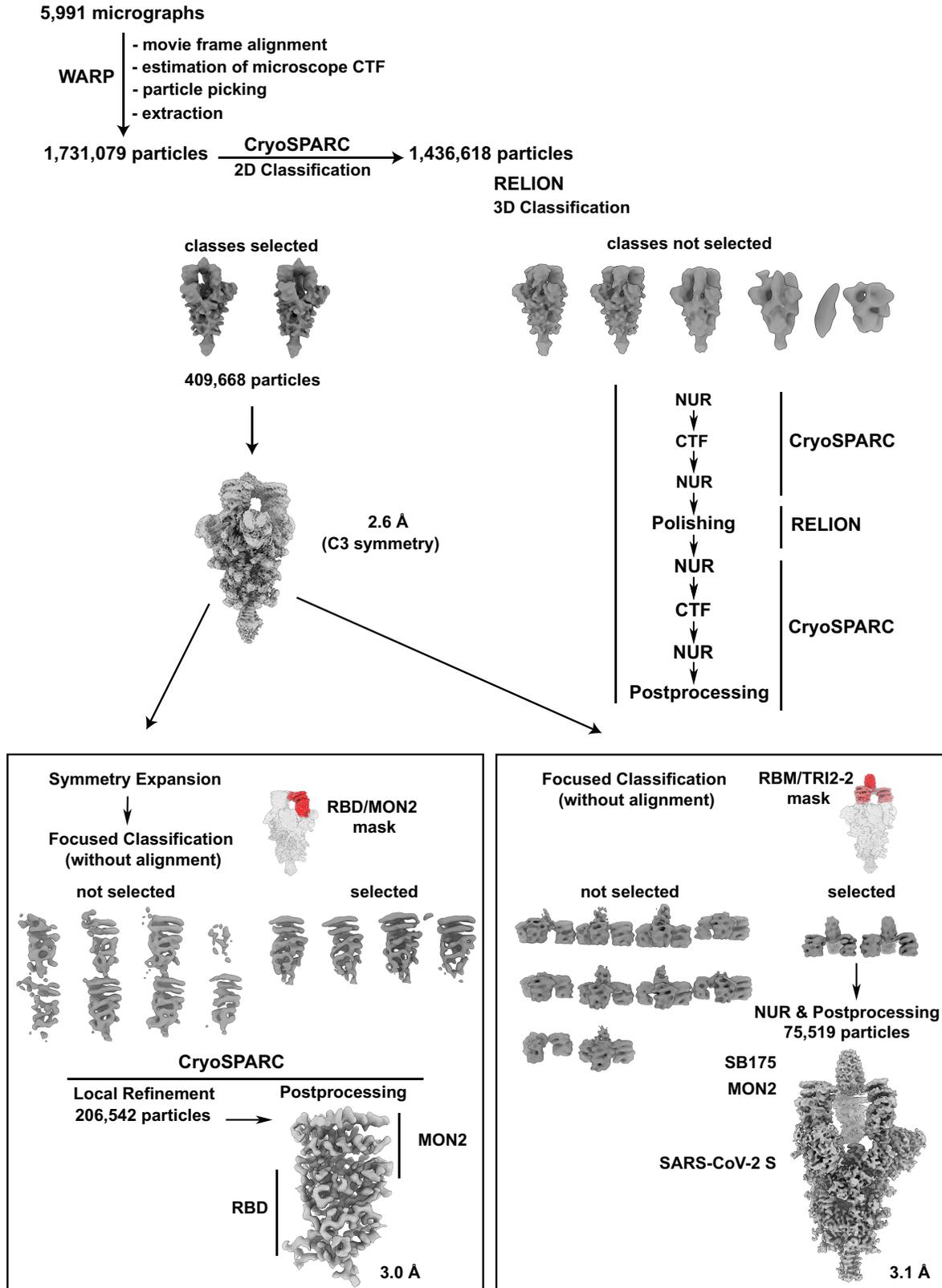


Figure 9.8. Cryo-EM processing scheme of SARS-CoV-2 S/TRI2-2 complex.

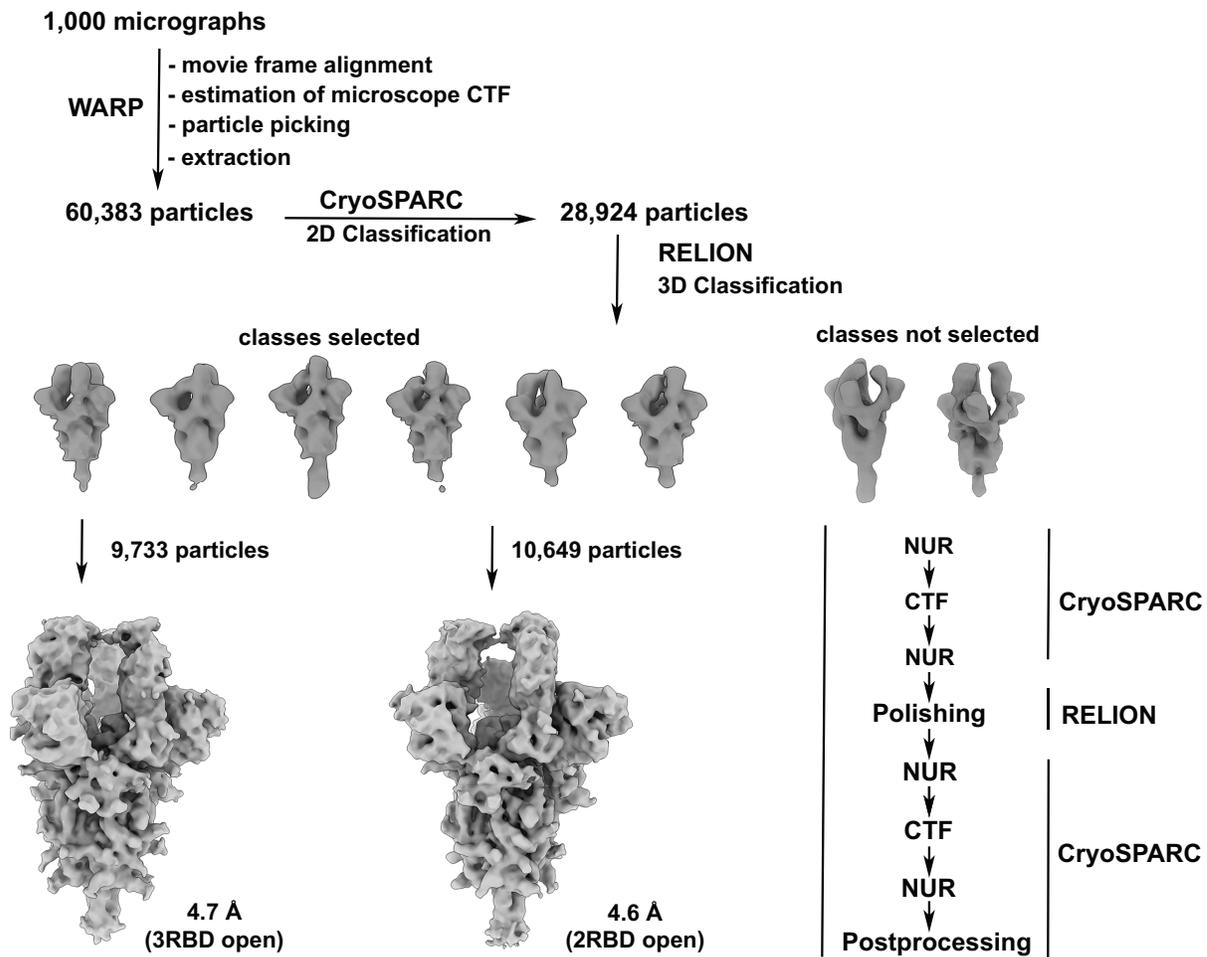


Figure 9.9. Cryo-EM processing scheme of SARS-CoV-2 S/FUS31-G10 complex.

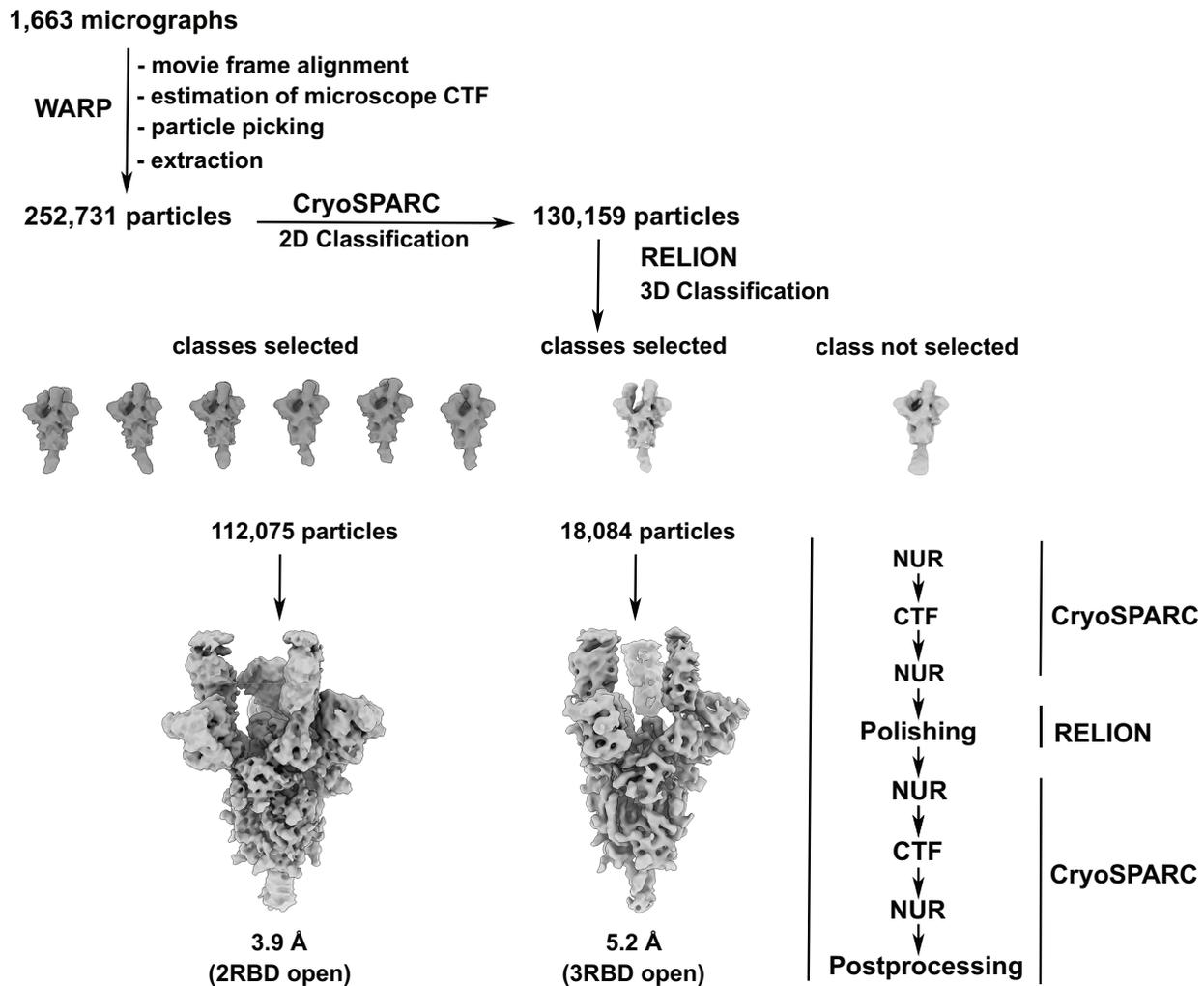


Figure 9.10. Cryo-EM processing scheme of SARS-CoV-2 S/FUS231-P24 complex.

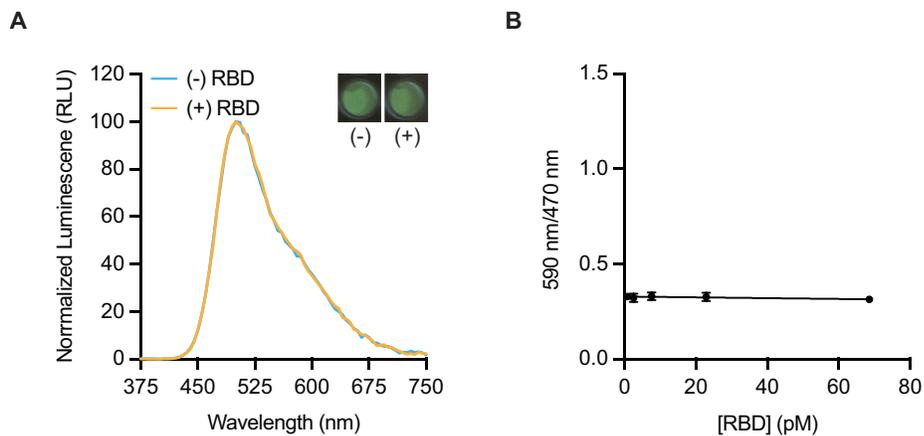
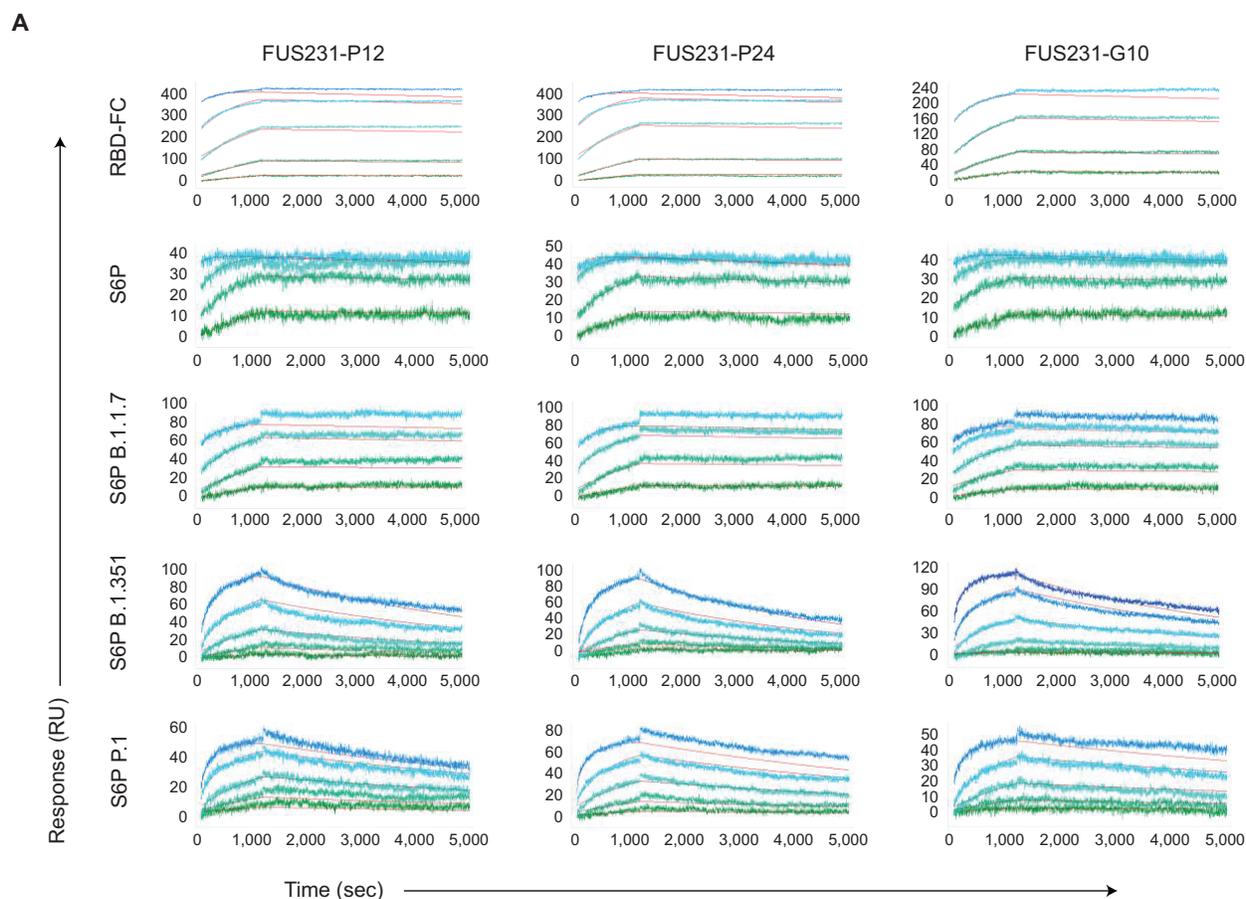


Figure 9.11. FUS231-P12 BRET sensor does not detect monomeric RBD. (A) Luminescence emission spectra and image of the BRET sensor (100 pM) in the presence (yellow trace, 100 pM) and absence (blue trace) of RBD. (B) Titration of RBD at 100 pM sensor (Mean \pm SEM, n=3 technical replicates, error bars that aren't visible are obscured by the marker).



B

| | | FUS231-P12 | FUS231-P24 | FUS231-G10 |
|-------------|---|------------|------------|------------|
| RBD-FC | k_a ($\times 10^5$ M $^{-1}$ s $^{-1}$) | 6.5 | 7.5 | 9.9 |
| | k_d ($\times 10^{-5}$ s $^{-1}$) | <1.4 | <1.4 | <1.4 |
| | K_D^* ($\times 10^{-12}$ M) | <21.9 | <18.8 | <14.4 |
| S6P | k_a ($\times 10^5$ M $^{-1}$ s $^{-1}$) | 38.2 | 38.9 | 29.2 |
| | k_d ($\times 10^{-5}$ s $^{-1}$) | 1.7 | 2.7 | 1.9 |
| | K_D^* ($\times 10^{-12}$ M) | 4.5 | 7.1 | 6.5 |
| S6P B.1.1.7 | k_a ($\times 10^5$ M $^{-1}$ s $^{-1}$) | 14.8 | 17.4 | 10.8 |
| | k_d ($\times 10^{-5}$ s $^{-1}$) | <1.4 | <1.4 | <1.4 |
| | K_D^* ($\times 10^{-12}$ M) | <9.6 | <82.1 | <13.2 |
| S6P B.1.351 | k_a ($\times 10^5$ M $^{-1}$ s $^{-1}$) | 4.2 | 4.3 | 2.2 |
| | k_d ($\times 10^{-5}$ s $^{-1}$) | 17.6 | 25.8 | 19.9 |
| | K_D^* ($\times 10^{-12}$ M) | 411.0 | 597.4 | 899.8 |
| S6P P.1 | k_a ($\times 10^5$ M $^{-1}$ s $^{-1}$) | 8.1 | 6.2 | 5.3 |
| | k_d ($\times 10^{-5}$ s $^{-1}$) | 10.9 | 12.2 | 8.6 |
| | K_D^* ($\times 10^{-12}$ M) | 133.5 | 195.3 | 163.1 |

Figure 9.12. Kinetic analysis of interactions between multivalent minibinders and SARS-CoV-2 S protein variants using surface plasmon resonance. Multivalent minibinders were injected at concentrations ranging from 20 nM to 0.08 nM in three-fold serial dilutions against S protein variants covalently coupled to the chip via amine coupling. The double-referenced data were fit globally to a simple 1:1 Langmuir binding model in Carterra's Kinetic tool. (A) Sensograms of FUS231 proteins binding to the RBD, S6P, and S6P mutants ($n = 1$ technical replicate from a single independent experiment). (B) Summary of kinetic and equilibrium parameters of the measured interactions. *All listed K_D values are apparent K_D s due to the potential for multivalency. The < symbol indicates a cut-off of less than 5% dissociation observed

during the dissociation phase, indicating insufficient time to accurately quantify the dissociation rate constant.

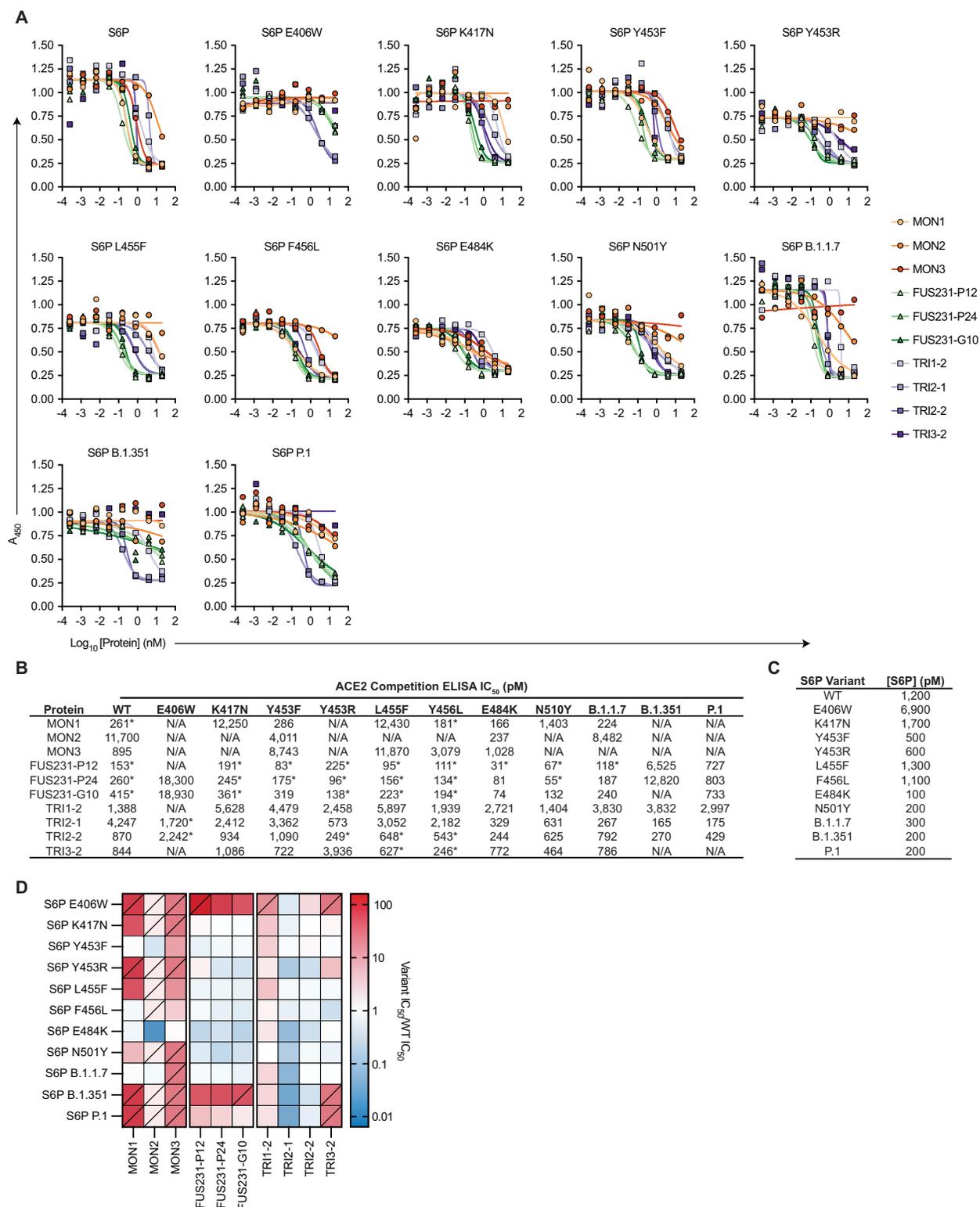


Figure 9.13. Competition of ACE2 and mini binder constructs for S6P. (A) Competition ELISA curves (mean, n=2 technical replicates representative of two independent experiments). (B) Summary of

minibinder construct competition IC_{50} values. An * denotes an IC_{50} value less than 2-fold lower than the limiting concentration of S6P variant present in the well. N/A indicates an IC_{50} value higher than the tested concentration range and greater than 20,000 pM. (C) Summary of EC_{50} values of S6P variants binding to ACE2 used as the concentration for minibinder construct competition. (D) Ratio of mutant to wild type (WT) IC_{50} values of minibinder constructs. The ratio in cells containing a slash was determined using the highest tested minibinder construct concentration which is indicative of a measurement at the upper limit of detection.

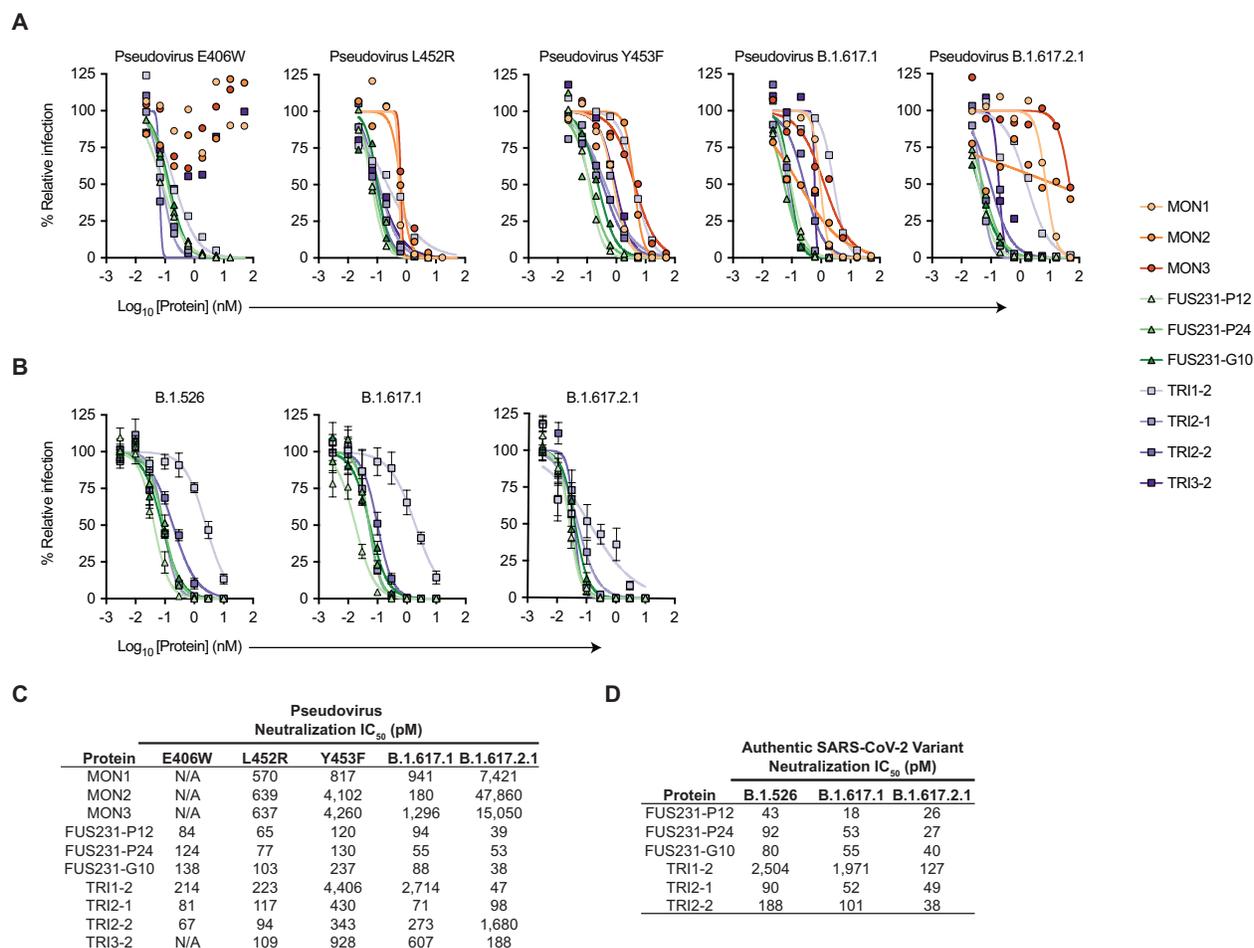


Figure 9.14. Neutralization of additional SARS-CoV-2 variants. (A) Neutralization of SARS-CoV-2 pseudovirus variants by minibinder constructs (mean, $n=2$ technical replicates representative of two independent experiments). (B) Neutralization of SARS-CoV-2 variants by minibinder constructs (mean \pm SEM, $n=4$ technical replicates from two independent experiments). (C) Table summarizing neutralization potencies of multivalent minibinder constructs against SARS-CoV-2 pseudovirus variants. N/A indicates an IC₅₀ value above the tested concentration range and an IC₅₀ greater than 50,000 pM. (D) Table summarizing neutralization potencies of multivalent minibinder constructs against SARS-CoV-2 variants.

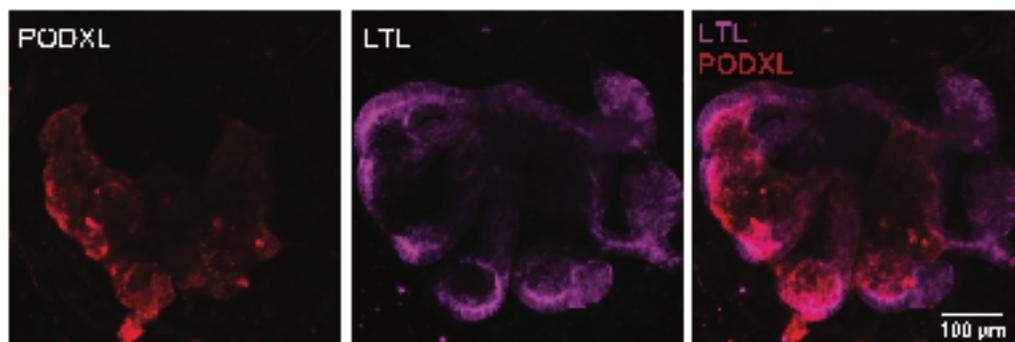


Figure 9.15. Representative confocal images of human kidney organoids. Organoids were derived from H9 human embryonic stem cells (LTL, *Lotus tetragonolobus* lectin, proximal tubule marker in magenta; PODXL, podocalyxin, podocyte marker in red).

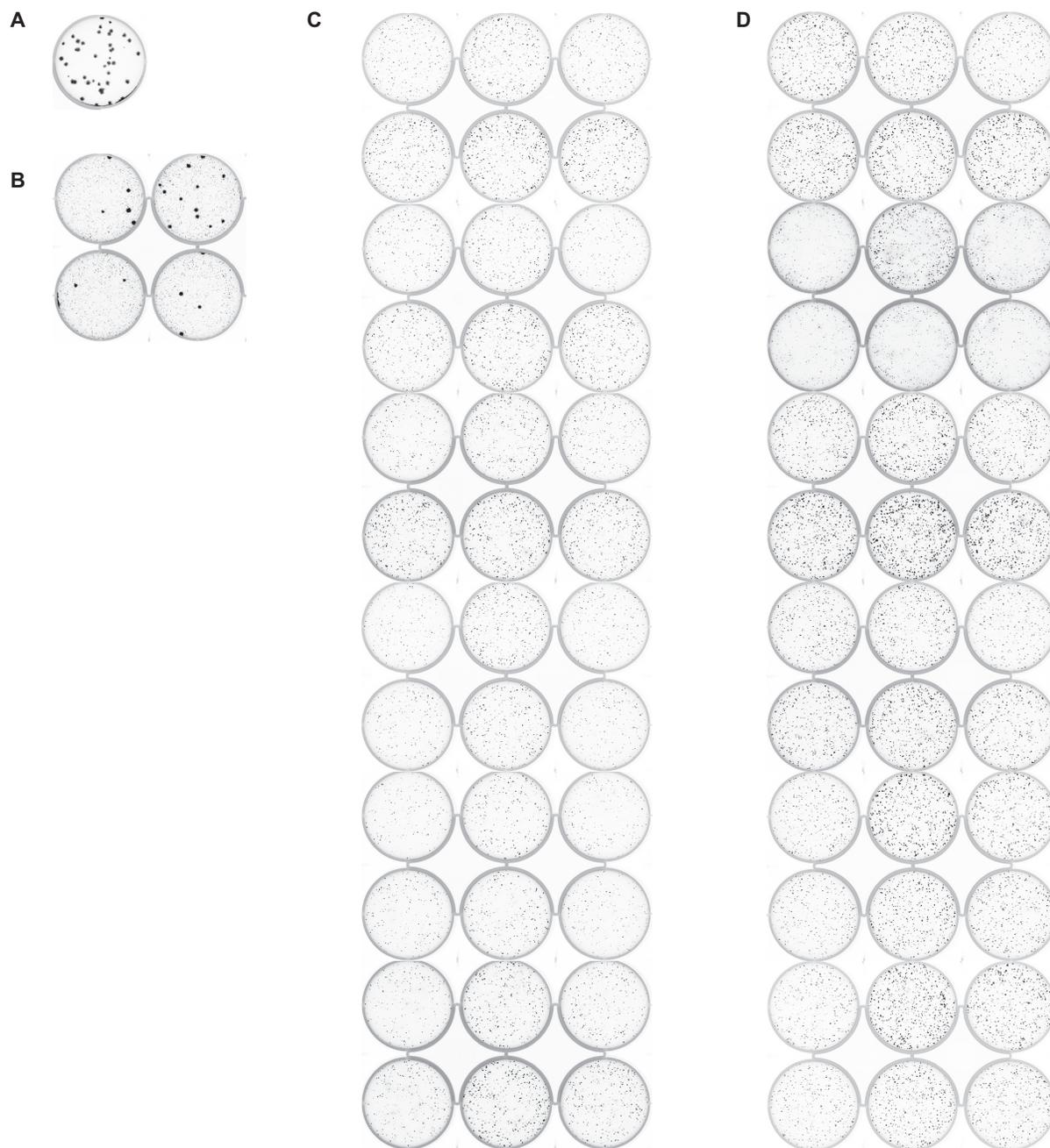


Figure 9.16. Replicate plates for VSV-SARS-CoV-2 escape studies. Plaque assays were performed to isolate VSV-SARS-CoV-2 chimera virus escape mutants against a control neutralizing antibody (2B04) and the FUS231-P12 and TRI2-2 multivalent minibinders. Large plaques are indicative of escape. FUS231-P12 and TRI2-2 replicates were performed in three separate experiments consisting of two plates each. (A) No inhibitor in overlay. (B) 2B04 neutralizing mAb in overlay. (C) FUS231-P12 in overlay. (D) TRI2-2 in overlay.

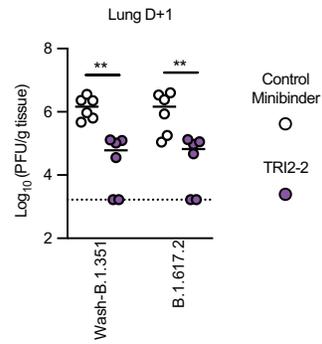


Figure 9.17. Measurement of SARS-CoV-2 viral titers in the lung of K18-hACE2-transgenic mice post intranasal delivery of TRI2-2. At 6 dpi (B.1.351) or 7 dpi (B.1.617.2), animals (n = 6 from two independent experiments) were sacrificed and analyzed for SARS-CoV-2 viral titers in the lung by plaque assays (line at median, two-tailed Mann-Whitney test: ** P < 0.01).

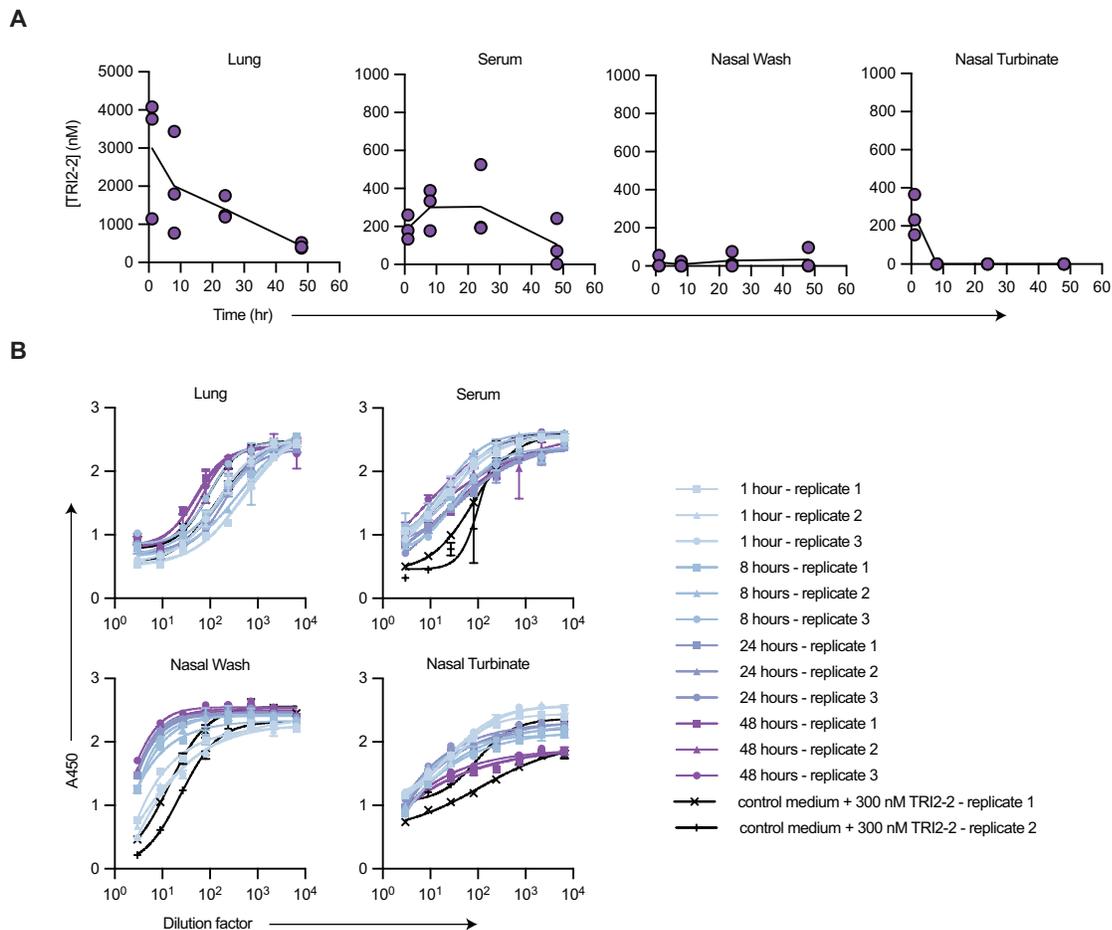


Figure 9.18. Pharmacokinetics of TRI2-2 delivered intranasally in C57BL/6J mice. (A) Serum and tissues were collected at 1, 8, 24, and 48 hours post-treatment with 200 μ g of TRI2-2 and quantitative competition ELISAs were used to determine concentrations of TRI2-2 present in the samples ($n = 3$ from a single experiment). The drawn line connects the mean values at each time point. (B) Raw data from competition ELISAs. Each curve represents a biological replicate with each data point consisting of $n = 2$ technical replicates.

```
ACE2 STIEEQAKTFLDKFNHEAEDLFY  
AHB2 -ELEEQVMHVLDQVSELAHELLH
```

Figure 9.19. Sequence alignment of AHB2 and ACE2. The main interacting helix of AHB2 and ACE2 were aligned using the Clustal Omega tool ⁴⁹². Conserved residues between the two sequences are highlighted in cyan.

Table 9.1. List of abbreviations used to describe multivalent minibinders in this manuscript.

| Abbreviation | Protein (N to C) | Abbreviation | Protein (N to C) |
|--------------|----------------------------------|--------------|-----------------------|
| MON1 | LCB1v2.2 | TRI2-3-G4 | AHB2-4G-SB175.1 |
| MON2 | AHB2 | TRI2-3-G6 | AHB2-6G-SB175.1 |
| MON3 | LCB3v2.2 | TRI2-4-G2 | AHB2-2G-SB175.2 |
| FUS23-P12 | AHB2v2-P12-LCB3v2.2 | TRI2-4-G4 | AHB2-4G-SB175.2 |
| FUS31-P12 | LCB3v2.2-P12-LCB1v2.2 | TRI2-4-G6 | AHB2-6G-SB175.2 |
| FUS231-P12 | AHB2v2-P12-LCB3v2.2-P12-LCB1v2.2 | TRI1-5-G2 | 36729-2G-LCB1v2.2 |
| FUS231-P24 | AHB2v2-P24-LCB3v2.2-P24-LCB1v2.2 | TRI1-5-G4 | 36729-4G-LCB1v2.2 |
| FUS23-G10 | AHB2v2-G10-LCB3v2.2 | TRI1-5-G6 | 36729-6G-LCB1v2.2 |
| FUS31-G10 | LCB3v2.2-G10-LCB1v2.2 | TRI1-2-G10 | SB175-10G-LCB1v2.2 |
| FUS231-G10 | AHB2v2-G10-LCB3v2.2-G10-LCB1v2.2 | TRI1-3-G6 | SB175.1-6G-LCB1v2.2 |
| TRI1-2 | SB175-6G-LCB1v2.2 | TRI1-3-G10 | SB175.1-10G-LCB1v2.2 |
| TRI2-1 | AHB2-4G-1rfo | TRI3-2-G8 | LCB3v2.2-8G-SB175 |
| TRI2-2 | AHB2-2G-SB175 | TRI3-2-G4 | LCB3v2.2-4G-SB175 |
| TRI3-2 | LCB3v2.2-6G-SB175 | FUS21-P24 | AHB2-P24-LCB1v2.1 |
| TRI2-6-G7 | AHB2-7G-1na0_int2-R3 | FUS21-P16 | AHB2-P16-LCB1v2.1 |
| TRI2-6-G9 | AHB2-9G-1na0_int2-R3 | FUS21-P12 | AHB2-P12-LCB1v2.2 |
| TRI2-7-G5 | AHB2-5G-1na0_int2 | FUS21-P11 | AHB2-P11-LCB1v2.1 |
| TRI2-7-G7 | AHB2-7G-1na0_int2 | FUS31-P24 | LCB3v2.2-P24-LCB1v2.1 |
| TRI2-7-G9 | AHB2-9G-1na0_int2 | FUS31-P16 | LCB3v2.2-P16-LCB1v2.1 |
| TRI2-8-G5 | AHB2-5G-6msr | FUS31-P11 | LCB3v2.2-P11-LCB1v2.1 |
| TRI2-8-G7 | AHB2-7G-6msr | FUS23-P24 | AHB2-P24-LCB3v2.3 |

| | |
|------------|---------------------|
| TRI2-8-G9 | AHB2-9G-6msr |
| TRI2-9-G5 | AHB2-5G-1gcm |
| TRI2-9-G7 | AHB2-7G-1gcm |
| TRI2-9-G9 | AHB2-9G-1gcm |
| TRI2-10-G5 | AHB2-5G-pRO-2-noHis |
| TRI2-10-G7 | AHB2-7G-pRO-2-noHis |
| TRI2-10-G9 | AHB2-9G-pRO-2-noHis |
| TRI2-11-G5 | AHB2-5G-1na0_3 |
| TRI2-11-G7 | AHB2-7G-1na0_3 |
| TRI2-11-G9 | AHB2-9G-1na0_3 |
| TRI2-12-G5 | AHB2-5G-4pn9 |
| TRI2-12-G7 | AHB2-7G-4pn9 |
| TRI2-12-G9 | AHB2-9G-4pn9 |
| TRI2-2-G4 | AHB2-4G-SB175 |
| TRI2-2-G6 | AHB2-6G-SB175 |
| TRI2-3-G2 | AHB2-2G-SB175.1 |

| | |
|----------------|----------------------------------|
| FUS23-P16 | AHB2-P16-LCB3v2.3 |
| FUS23-P12 | AHB2-P12-LCB3v2.2 |
| FUS23-P11 | AHB2-P11-LCB3v2.3 |
| FUS32-P12 | LCB3v2.2-P12-AHB2v2 |
| FUS231-P24-P16 | AHB2v2-P24-LCB3v2.2-P16-LCB1v2.2 |
| FUS231-P16-P24 | AHB2v2-P16-LCB3v2.2-P24-LCB1v2.2 |
| FUS231-P16 | AHB2v2-P16-LCB3v2.2-P16-LCB1v2.2 |
| FUS231-P11-P16 | AHB2v2-P11-LCB3v2.2-P16-LCB1v2.2 |
| FUS231-P24-P11 | AHB2v2-P24-LCB3v2.2-P11-LCB1v2.2 |
| FUS231-P16-P11 | AHB2v2-P16-LCB3v2.2-P11-LCB1v2.2 |
| FUS231-P11 | AHB2v2-P11-LCB3v2.2-P11-LCB1v2.2 |
| FUS231-P7 | AHB2v2-P7-LCB3v2.2-P7-LCB1v2.2 |
| FUS321-P12 | LCB3v2.2-P12-AHB2v2-P12-LCB1v2.2 |
| FUS321-P24 | LCB3v2.2-P24-AHB2v2-P24-LCB1v2.2 |
| FUS31-G8 | LCB3v2.2-G8-LCB1v2.2 |

Table 9.2. Oligomerization domains tested in this work.

| # | Name | Reference | Notes | Protein Sequence |
|----|--------------|-----------|--|--|
| 1 | 1rfo | 394 | | GYIPEAPRDGQAYVRKDGEWV LLSTFL |
| 2 | SB175 | This work | Modified from SB13 (2L6HC3_13) 302 | SEALEELEKALRELKKSTDELER STEELEKNPSEDALVENNRLIVE NNKIIVEVLRRIIAKVLK |
| 3 | SB175.1 | This work | Medium truncation of SB175 helices | SPELEKALRELKKSTDELERST EELEKNGSPEALVENNRLIVEN NKIIVEVLRRIIAK |
| 4 | SB175.2 | This work | Large truncation of SB175 helices | SEKALRELKKSTDELERSTEEL EKNGSPEALVENNRLIVENNKII VEVLR |
| 5 | 36729.2 | This work | Modified from 1na0_int2 | EEAELAYLLGELAYKLGEYRIAI RAYRIALKRDPNNAEAWYNLGN AYYKQGDYDEAIEYYQKALELD PNNAEAWYNLGNAYYKQGDYD EAIEYYQKALELDPNNAEAWYN LGNAYYKQGDYDEAIEYYQKAL EL |
| 6 | 1na0_int2-R3 | This work | Modified from 1na0_int2 | EEAELAYLLGELAYKLGEYRIAI RAYRIALKRDPNNAEAWYNLGN AYYKQGDYDEAIEYYQKALELD PNNAEAKQNLGNAKQKQG |
| 7 | 1na0_int2 | This work | Modified from 493 | EEAELAYLLGELAYKLGEYRIAI RAYRIALKRDPNNAEAWYNLGN AYYKQGDYDEAIEYYQKALELD PNNAEAWYNLGNAYYKQGDYD EAIEYYQKALELDPNNAEAKQN LGNKQKQG |
| 8 | 6msr | 494 | | GSEYEIRKALEELKASTAELKRA TASLRASTEELKKNPSEDALVE NNRLIVEHNNAIIVENNRRIIAVLEL IVRAIK |
| 9 | 1gcm | 495 | | RMKQIEDKIEEILSKIYHIENEIAR IKKLIGER |
| 10 | pRO-2-noHis | 494 | | GSEYEIRKALEELKASTAELKRS TASLRASTEELKKNPSEDALVE NNRLIVENNAIIVENNRRIIAVLEL IVRAIK |
| 11 | 1na0_3 | This work | | NLAEKMYKAGNAMYRKGQYTIA IIAYTLALLKDPNNAEAWYNLGN AAYKKGEYDEAIEAYQKALELD PNNAEAWYNLGNAYYKQGDYD EAIEYYQKALELDPNNAEAKQN LGNKQKQG |

| | | | | |
|----|------|-----|--------------------|------------------------------------|
| 12 | 4pn9 | 496 | Hexamer not trimer | GEIAKSLKEIAKSLKEIAWSLKEI AKSLKG |
|----|------|-----|--------------------|------------------------------------|

Table 9.3. CryoEM data collection and refinement statistics.

| | SARS-CoV-2 S/TRI2-2 PDB EMD | SARS-CoV-2 S/TRI2-2 (local refinement) PDB EMD | SARS-CoV-2 S/FUS31-G10 (2RBD-open) EMD | SARS-CoV-2 S/FUS31-G10 (3RBD-open) EMD | SARS-CoV-2 S/FUS231-P24 (2RBD-open) EMD | SARS-CoV-2 S/FUS231-P24 (3RBD-open) EMD |
|--|--------------------------------------|---|---|---|--|--|
| Data collection and processing | | | | | | |
| Magnification | 105,000 | 105,000 | 130,000 | 130,000 | 36,000 | 36,000 |
| Voltage (kV) | 300 | 300 | 300 | 300 | 200 | 200 |
| Electron exposure (e ⁻ /Å ²) | 60 | 60 | 70 | 70 | 60 | 60 |
| Defocus range (µm) | -0.5 – -2.5 | -0.5 – -2.5 | -0.5 – -2.5 | -0.5 – -2.5 | -0.5 – -2.5 | -0.5 – -2.5 |
| Pixel size (Å) | 0.4215 | 0.4215 | 0.525 | 0.525 | 1.16 | 1.16 |
| Symmetry imposed | C3 | C1 | C1 | C1 | C1 | C1 |
| Final particle images (no.) | 75,519 | 206,541 | 9,733 | 10,649 | 112,075 | 18,084 |
| Map resolution (Å) | 2.8 | 2.9 | 4.57 | 4.65 | 3.9 | 5.2 |
| FSC threshold | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 |
| Map sharpening <i>B</i> factor (Å ²) | -63 | -31 | -63 | -71 | -99 | -167 |
| | | | | | | |
| Validation | | | | | | |
| MolProbity score | 1.05 | 1.03 | | | | |
| Clashscore | 1.19 | 1.62 | | | | |

| | | | | | | |
|-------------------|-------|-------|--|--|--|--|
| Poor rotamers (%) | 0 | 0 | | | | |
| Ramachandran plot | | | | | | |
| Favored (%) | 96.62 | 97.37 | | | | |
| Allowed (%) | 2.99 | 2.25 | | | | |
| Disallowed (%) | 0.39 | 0.38 | | | | |

Table 9.4. List of DNA and protein sequences for multivalent minibinders used in this manuscript (Separate csv file). DNA sequences are the open reading frame coding for the expressed protein. Protein sequences are annotated as follows. Minibinder and homotrimer sequences are denoted by square brackets []. Secondary sequences (e.g., expression tag, purification tag, etc.) are annotated by parenthesis (). Non-minibinder or non-homotrimer sequences are annotated by curly brackets {} (captures linkers and other secondary sequences). These materials have been deposited to zenodo and can be accessed at the following link: <https://doi.org/10.5281/zenodo.6377529>.

Table 9.5. Estimates of the diversity of mutants in the VSV-SARS-CoV-2 chimera virus pool used in the multivalent minibinder escape selections. Estimates were calculated using the Pedel-AA tool ⁴⁵² for analyzing diversity in randomly mutated protein libraries. Upper and lower estimates for the VSV RNA polymerase L error rate ⁴⁰⁸ were used to calculate the range of plausible library compositions.

| # Amino Acid Mutations | # Possible Mutants | Polymerase Error Rate (error/nucleotide): 10^{-3} | | Polymerase Error Rate (error/nucleotide): 10^{-4} | |
|------------------------|--------------------|--|------------|--|------------|
| | | # Expected Distinct Mutants | % Coverage | # Expected Distinct Mutants | % Coverage |
| 1 | 2,071 | 1,700 | 82.09 | 703.9 | 33.99 |
| 2 | 2.13×10^6 | 2.04×10^5 | 9.62 | 9,420 | 0.44 |

Table 9.6. Comparison of multivalent minibinder and FDA authorized neutralizing antibody potencies. For mAb data, pseudovirus neutralization IC₅₀ values (columns 3, 4, and 5) were taken from ³⁷⁰ and authentic virus neutralization IC₅₀ values (columns 6, 7, and 8) were taken from ³⁷¹. *Value for WA1/2020 strain as opposed to WA1/2020 D614G.

| Neutralizing Protein | Commercial Name | Wu-Hu-1 Pseudovirus IC50 (ng/mL) | B.1.1.529 Pseudovirus IC50 (ng/mL) | Fold-Change | WA1/2020 | B.1.1.529 | Fold-Change |
|----------------------|-----------------|----------------------------------|------------------------------------|--------------------------------------|------------------------------------|------------------------------|---|
| | | | | Pseudovirus IC50 (B.1.1.529/Wu-Hu-1) | D614G Authentic Virus IC50 (ng/mL) | Authentic Virus IC50 (ng/mL) | Authentic Virus IC50 (B.1.1.529/WA1/2020 D614G) |
| S309 (VIR-7831) | Sotrovimab | 90.6 | 260 | 2.9 | 202 | 373 | 1.8 |
| REGN10933 | Casirivimab | 8.9 | >10,000 | >1123.6 | 11 | >10,000 | >909.1 |
| REGN10987 | Imdevimab | 25.1 | >10,000 | >398.4 | 31 | >10,000 | >322.6 |
| LY-CoV555 | Etesevimab | 21.3 | >10,000 | >469.5 | 10 | >10,000 | >1000 |
| LY-CoV016 (CB6) | Bamlanivimab | 59.2 | >10,000 | >168.9 | 72 | >10,000 | >138.9 |
| COV2-2196 (AZD8895) | Tixagevimab | 8.1 | 2,772 | 342.2 | 6 | 913 | 152.2 |
| COV2-2130 (AZD1061) | Cilgavimab | 4.3 | >10,000 | >2325.6 | 32 | 381 | 11.9 |
| TRI2-1 | n/a | 8.2 | n/a | n/a | 0.4* | 0.7 | 1.8 |
| TRI2-2 | n/a | 8.4 | n/a | n/a | 1.2* | 6.2 | 5.2 |

Chapter 10 - Appendix Containing Supplementary Information for
Chapter 5

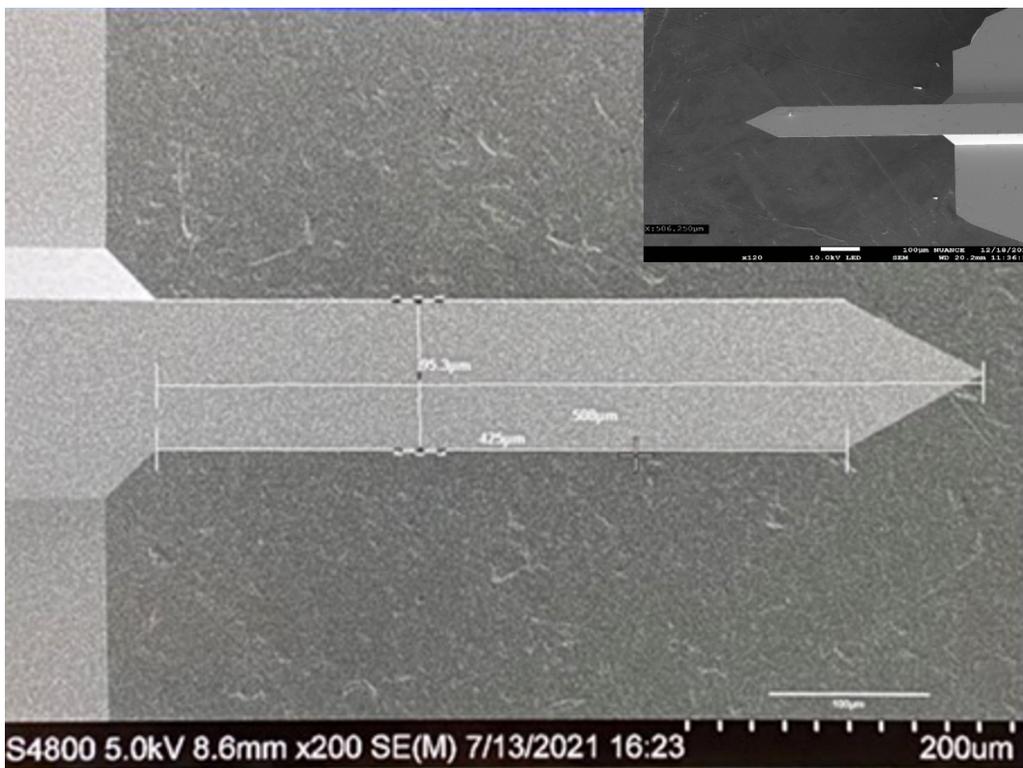


Figure 10.1. SEM micrograph of a microcantilever device.

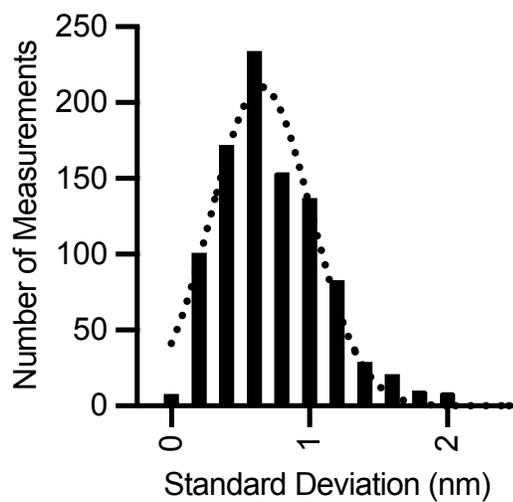


Figure 10.2. Histogram of measured error for all cantilever deflection measurements in this manuscript. Values are binned ± 0.1 nm of the center value. Dashed line indicates a gaussian fit (mean = 0.65, standard deviation = 0.36).

Table 10.1. Summary of the measured RT-qPCR Ct values, lineage designations, and GISAID identifiers for all reported patient samples (as available).

| Sample | Ct Value | GISAID ID | GISAID Accession Number | Clade |
|---------------------|----------|-------------------------------------|-------------------------|-----------------|
| Early 2021 | 25.7 | 3/11/21 hCoV-19/USA/IL-NM-4786/2020 | EPI_ISL_936661 | 20A (B.1.240) |
| Early 2021 | 27.9 | NOT SEQUENCED | n/a | n/a |
| Early 2021 | 31.6 | NOT SEQUENCED | n/a | n/a |
| Early 2021 | 35.2 | NOT SEQUENCED | n/a | n/a |
| Early 2021 | 39.7 | NOT SEQUENCED | n/a | n/a |
| Alpha (B.1.1.7) | 15.4 | hCoV-19/USA/IL-NM-7773/2021 | EPI_ISL_2009334 | 20I (Alpha, V1) |
| Alpha (B.1.1.7) | 19.7 | hCoV-19/USA/IL-NM-7764/2021 | EPI_ISL_2009331 | 20I (Alpha, V1) |
| Alpha (B.1.1.7) | 19.9 | hCoV-19/USA/IL-NM-7747/2021 | EPI_ISL_2009329 | 20I (Alpha, V1) |
| Alpha (B.1.1.7) | 21.8 | hCoV-19/USA/IL-NM-7745/2021 | EPI_ISL_2009328 | 20I (Alpha, V1) |
| Alpha (B.1.1.7) | 22.9 | hCoV-19/USA/IL-NM-7787/2021 | EPI_ISL_2009340 | 20I (Alpha, V1) |
| Omicron (B.1.1.529) | 18.1 | hCoV-19/USA/IL-NM-17432/2022 | EPI_ISL_8766922 | 21K (Omicron) |
| Omicron (B.1.1.529) | 22.5 | hCoV-19/USA/IL-NM-17427/2022 | EPI_ISL_8766918 | 21K (Omicron) |
| Omicron (B.1.1.529) | 27.8 | hCoV-19/USA/IL-NM-17430/2022 | EPI_ISL_8766921 | 21K (Omicron) |
| Omicron (B.1.1.529) | 30.4 | hCoV-19/USA/IL-NM-17435/2022 | EPI_ISL_8766924 | 21K (Omicron) |
| Omicron (B.1.1.529) | 33.3 | NOT SEQUENCED | n/a | n/a |
| Influenza A | 19.3 | NOT SEQUENCED | n/a | n/a |

Table 10.2. Summary of the measured RT-qPCR Ct values, their corresponding concentration of RNA (copies/mL), and their measured deflection (\pm standard deviation) for all reported patient samples.

| Sample | Ct Value | [RNA] copies/mL | Average Deflection at 15 min \pm sd (nm) |
|---------------------|----------|-----------------|--|
| Early 2021 | 25.7 | 0.92E+06 | -29.5 \pm 0.28 |
| Early 2021 | 27.9 | 1.89E+05 | -24.4 \pm 0.45 |
| Early 2021 | 31.6 | 1.18E+04 | -17.2 \pm 0.56 |
| Early 2021 | 35.2 | 1.48E+03 | -13.0 \pm 1.00 |
| Early 2021 | 39.7 | 9.63E+01 | -7.6 \pm 0.32 |
| Alpha (B.1.1.7) | 15.4 | 7.53E+08 | -55.2 \pm 1.06 |
| Alpha (B.1.1.7) | 19.7 | 4.84E+07 | -47.8 \pm 1.24 |
| Alpha (B.1.1.7) | 19.9 | 4.12E+07 | -37.7 \pm 0.9 |
| Alpha (B.1.1.7) | 21.8 | 1.21E+07 | -27.3 \pm 1.00 |
| Alpha (B.1.1.7) | 22.9 | 6.05E+06 | -22.6 \pm 1.15 |
| Omicron (B.1.1.529) | 18.1 | 1.93E+08 | -42.2 \pm 0.95 |
| Omicron (B.1.1.529) | 22.5 | 1.23E+07 | -34.4 \pm 1.05 |
| Omicron (B.1.1.529) | 27.8 | 1.89E+05 | -24.5 \pm 0.61 |
| Omicron (B.1.1.529) | 30.4 | 4.43E+04 | -18.6 \pm 1.27 |
| Omicron (B.1.1.529) | 33.3 | 4.61E+03 | -11.3 \pm 0.85 |
| Influenza A | 19.3 | 4.83E+07 | -3.45 \pm 0.29 |

Chapter 11 References

1. *Molecular biology of the cell.* (CRC Press, 2002).
2. Dougan, M. *et al.* Bamlanivimab plus Etesevimab in Mild or Moderate Covid-19. *N. Engl. J. Med.* **385**, 1382–1392 (2021).
3. Gupta, A. *et al.* Early Treatment for Covid-19 with SARS-CoV-2 Neutralizing Antibody Sotrovimab. *N. Engl. J. Med.* **385**, 1941–1950 (2021).
4. Weinreich, D. M. *et al.* REGEN-COV Antibody Combination and Outcomes in Outpatients with Covid-19. *N. Engl. J. Med.* **385**, e81 (2021).
5. Corti, D., Purcell, L. A., Snell, G. & Veessler, D. Tackling COVID-19 with neutralizing monoclonal antibodies. *Cell* **184**, 4593–4595 (2021).
6. Ackerman, C. M. *et al.* Massively multiplexed nucleic acid detection with Cas13. *Nature* **582**, 277–282 (2020).
7. Yoshida, S. *et al.* A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **351**, 1196–1199 (2016).
8. Jung, J. K. *et al.* Cell-free biosensors for rapid detection of water contaminants. *Nat. Biotechnol.* **38**, 1451–1459 (2020).
9. Thavarajah, W. *et al.* Point-of-use detection of environmental fluoride via a cell-free riboswitch-based biosensor. *ACS Synth. Biol.* **9**, 10–18 (2020).
10. Silverman, A. D., Akova, U., Alam, K. K., Jewett, M. C. & Lucks, J. B. Design and optimization of a cell-free atrazine biosensor. *ACS Synth. Biol.* **9**, 671–677 (2020).
11. Breslauer, D. N. Recombinant protein polymers: A coming wave of personal care ingredients. *ACS Biomater. Sci. Eng.* **6**, 5980–5986 (2020).
12. DeFrancesco, L. Hanging on a thread. *Nat. Biotechnol.* **35**, 496–499 (2017).
13. Bar-Even, A., Noor, E., Lewis, N. E. & Milo, R. Design and analysis of synthetic carbon fixation pathways. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8889–8894 (2010).
14. Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S. & Erb, T. J. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**, 900–904 (2016).

15. Liew, F. E. *et al.* Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale. *Nat. Biotechnol.* **40**, 335–344 (2022).
16. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010–2020. *Nat. Commun.* **11**, 5174 (2020).
17. Voigt, C. A. Synthetic biology 2020–2030: six commercially-available products that are changing our world. *Nat. Commun.* **11**, 6379 (2020).
18. El Karoui, M., Hoyos-Flight, M. & Fletcher, L. Future trends in synthetic biology-A report. *Front. Bioeng. Biotechnol.* **7**, 175 (2019).
19. Wintle, B. C. *et al.* A transatlantic perspective on 20 emerging issues in biological engineering. *Elife* **6**, e30247 (2017).
20. Simon, A. J., d'Oelsnitz, S. & Ellington, A. D. Synthetic evolution. *Nat. Biotechnol.* **37**, (2019).
21. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
22. Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **38**, 338–351 (2009).
23. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology* vol. 5 (2014).
24. Porro, D., Sauer, M., Branduardi, P. & Mattanovich, D. Recombinant protein production in yeasts. *Mol. Biotechnol.* **31**, 245–259 (2005).
25. Contreras-Gómez, A., Sánchez-Mirón, A., García-Camacho, F., Molina-Grima, E. & Chisti, Y. Protein production using the baculovirus-insect cell expression system. *Biotechnol. Prog.* **30**, 1–18 (2014).
26. Wurm, F. M. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.* **22**, 1393–1398 (2004).
27. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).
28. Jia, B. & Jeon, C. O. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biol.* **6**, (2016).
29. Gieselmann, L. *et al.* Effective high-throughput isolation of fully human antibodies targeting infectious pathogens. *Nat. Protoc.* (2021) doi:10.1038/s41596-021-00554-w.

30. Dodevski, I., Markou, G. C. & Sarkar, C. A. Conceptual and methodological advances in cell-free directed evolution. *Curr. Opin. Struct. Biol.* **33**, 1–7 (2015).
31. Mayr, L. M. & Bojanic, D. Novel trends in high-throughput screening. *Curr. Opin. Pharmacol.* **9**, 580–588 (2009).
32. Schneider, G. Automating drug discovery. *Nat. Rev. Drug Discov.* **17**, 97–113 (2018).
33. Chen, Z. *et al.* De novo design of protein logic gates. *Science* **368**, 78–84 (2020).
34. Hunt, A. C. *et al.* A high-throughput, automated, cell-free expression and screening platform for antibody discovery. *bioRxiv* (2021) doi:10.1101/2021.11.04.467378.
35. Hunt, A. C. *et al.* Multivalent designed proteins protect against SARS-CoV-2 variants of concern. *bioRxiv* (2021) doi:10.1101/2021.07.07.451375.
36. Boles, K. S. *et al.* Digital-to-biological converter for on-demand production of biologics. *Nat. Biotechnol.* **35**, 672–675 (2017).
37. Garenne, D., Thompson, S., Brisson, A., Khakimzhan, A. & Noireaux, V. The all-E. coliTXTL toolbox 3.0: new capabilities of a cell-free synthetic biology platform. *Synth. Biol.* **6**, ysab017 (2021).
38. Silverman, A. D., Karim, A. S. & Jewett, M. C. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* **21**, 151–170 (2020).
39. Gregorio, N. E., Levine, M. Z. & Oza, J. P. A user's guide to cell-free protein synthesis. *Methods Protoc.* **2**, 24 (2019).
40. Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).
41. Dopp, J. L., Tamiev, D. D. & Reuel, N. F. Cell-free supplement mixtures: Elucidating the history and biochemical utility of additives used to support in vitro protein synthesis in E. coli extract. *Biotechnol. Adv.* **37**, 246–258 (2019).
42. Martin, R. W. *et al.* Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nature Communications* 2018 9:1 **9**, 1203 (2018).
43. Sawasaki, T., Ogasawara, T., Morishita, R. & Endo, Y. A cell-free protein synthesis system for high-throughput proteomics. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14652–14657 (2002).

44. Karig, D. K., Jung, S.-Y., Srijanto, B., Collier, C. P. & Simpson, M. L. Probing cell-free gene expression noise in femtoliter volumes. *ACS Synth. Biol.* **2**, 497–505 (2013).
45. Shojaeian, M., Lehr, F.-X., Göringer, H. U. & Hardt, S. On-demand production of femtoliter drops in microchannels and their use as biological reaction compartments. *Anal. Chem.* **91**, 3484–3491 (2019).
46. Zawada, J. F. *et al.* Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* **108**, 1570–1578 (2011).
47. Asahara, H. *et al.* Guidelines for nucleic acid template design for optimal cell-free protein synthesis using an *Escherichia coli* reconstituted system or a lysate-based system. in *Recombinant Protein Expression: Prokaryotic Hosts and Cell-Free Systems* 351–369 (Elsevier, 2021). doi:10.1016/bs.mie.2021.07.005.
48. Garenne, D., Bowden, S. & Noireaux, V. Cell-free expression and synthesis of viruses and bacteriophages: applications to medicine and nanotechnology. *Curr. Opin. Syst. Biol.* **28**, 100373 (2021).
49. McSweeney, M. A. & Styczynski, M. P. Effective use of linear DNA in cell-free expression systems. *Front. Bioeng. Biotechnol.* **9**, 715328 (2021).
50. Schinn, S. M., Broadbent, A., Bradley, W. T. & Bundy, B. C. Protein synthesis directly from PCR: Progress and applications of cell-free protein synthesis with linear DNA. *N. Biotechnol.* **33**, 480–487 (2016).
51. Lee, K.-H. & Kim, D.-M. Recent advances in development of cell-free protein synthesis systems for fast and efficient production of recombinant proteins. *FEMS Microbiol. Lett.* (2018) doi:10.1093/femsle/fny174.
52. Dopp, J. L., Rothstein, S. M., Mansell, T. J. & Reuel, N. F. Rapid prototyping of proteins: Mail order gene fragments to assayable proteins within 24 hours. *Biotechnol. Bioeng.* **116**, 667–676 (2019).
53. Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V. & Murray, R. M. Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* based TX-TL cell-free system. *ACS Synth. Biol.* **3**, 387–397 (2014).

54. Gagoski, D. *et al.* Gateway-compatible vectors for high-throughput protein expression in pro- and eukaryotic cell-free systems. *J. Biotechnol.* **195**, 1–7 (2015).
55. Ojima-Kato, T., Nagai, S. & Nakano, H. Ecobody technology: rapid monoclonal antibody screening method from single B cells using cell-free protein synthesis for antigen-binding fragment formation. *Sci. Rep.* **7**, 13979 (2017).
56. Ahn, J.-H., Lee, K.-H., Shim, J.-W., Lee, E. Y. & Kim, D.-M. Streamlined cell-free protein synthesis from sequence information. *Biotechnol. Bioprocess Eng.* **18**, 1101–1108 (2013).
57. Chiocchini, C. *et al.* From Electronic Sequence to Purified Protein Using Automated Gene Synthesis and In Vitro Transcription/Translation. *ACS Synth. Biol.* acssynbio.0c00060 (2020) doi:10.1021/acssynbio.0c00060.
58. Michel-Reydellet, N., Woodrow, K. & Swartz, J. Increasing PCR fragment stability and protein yields in a cell-free system with genetically modified *Escherichia coli* extracts. *J. Mol. Microbiol. Biotechnol.* **9**, 26–34 (2005).
59. Ahn, J.-H. *et al.* Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem. Biophys. Res. Commun.* **338**, 1346–1352 (2005).
60. Lesley, S. A., Brow, M. A. D. & Burgess, R. R. Use of in vitro protein synthesis from polymerase chain reaction-generated templates to study interaction of *Escherichia coli* transcription factors with core RNA polymerase and for epitope mapping of monoclonal antibodies. *J. Biol. Chem.* **266**, 2632–2638 (1991).
61. Hadi, T. *et al.* Rolling circle amplification of synthetic DNA accelerates biocatalytic determination of enzyme activity relative to conventional methods. *Sci. Rep.* **10**, 10279 (2020).
62. Kumar, G. & Chernaya, G. Cell-free protein synthesis using multiply-primed rolling circle amplification products. *Biotechniques* **47**, 637–639 (2009).
63. Kuzminov, A. & Stahl, F. W. Stability of linear DNA in *recA* mutant *Escherichia coli* cells reflects ongoing chromosomal DNA degradation. *J. Bacteriol.* **179**, 880–888 (1997).
64. Seki, E., Matsuda, N. & Kigawa, T. Multiple inhibitory factor removal from an *Escherichia coli* cell extract improves cell-free protein synthesis. *J. Biosci. Bioeng.* **108**, 30–35 (2009).

65. Batista, A. C. *et al.* Differentially optimized cell-free buffer enables robust expression from unprotected linear DNA in exonuclease-deficient extracts. *ACS Synth. Biol.* (2022) doi:10.1021/acssynbio.1c00448.
66. Marshall, R., Maxwell, C. S., Collins, S. P., Beisel, C. L. & Noireaux, V. Short DNA containing χ sites enhances DNA stability and gene expression in *E. coli* cell-free transcription–translation systems. *Biotechnol. Bioeng.* **114**, 2137–2141 (2017).
67. Shrestha, P., Smith, M. T. & Bundy, B. C. Cell-free unnatural amino acid incorporation with alternative energy systems and linear expression templates. *N. Biotechnol.* **31**, 28–34 (2014).
68. Zhu, B. *et al.* Increasing cell-free gene expression yields from linear templates in *Escherichia coli* and *Vibrio natriegens* extracts by using DNA-binding proteins.
69. Norouzi, M., Panfilov, S. & Pardee, K. High-Efficiency Protection of Linear DNA in Cell-Free Extracts from *Escherichia coli* and *Vibrio natriegens*. *ACS Synth. Biol.* **10**, 1615–1624 (2021).
70. Yim, S. S., Johns, N. I., Noireaux, V. & Wang, H. H. Protecting linear DNA templates in cell-free expression systems from diverse bacteria. *ACS Synth. Biol.* acssynbio.0c00277 (2020) doi:10.1021/acssynbio.0c00277.
71. Wu, P. S. C. *et al.* Cell-free transcription/translation from PCR-amplified DNA for high-throughput NMR studies. *Angew. Chem. Int. Ed Engl.* **46**, 3356–3358 (2007).
72. Cole, S. D. *et al.* Quantification of Interlaboratory Cell-Free Protein Synthesis Variability. *ACS Synth. Biol.* **8**, 2080–2091 (2019).
73. Romantseva, E. F., Tack, D. S., Alperovich, N., Ross, D. & Strychalski, E. A. Best practices for toward improved reproducibility in. in *Cell-Free Gene Expression* 3–50 (Springer US, 2022). doi:10.1007/978-1-0716-1998-8_1.
74. Chappell, J., Jensen, K. & Freemont, P. S. Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res.* **41**, 3471–3481 (2013).
75. Su’etsugu, M., Takada, H., Katayama, T. & Tsujimoto, H. Exponential propagation of large circular DNA by reconstitution of a chromosome-replication cycle. *Nucleic Acids Res.* **45**, 11525–11534 (2017).
76. Schoborg, J. A., Hodgman, C. E., Anderson, M. J. & Jewett, M. C. Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. *Biotechnol. J.* **9**, 630–640 (2014).

77. Swartz, J. R., Jewett, M. C. & Woodrow, K. A. Cell-Free Protein Synthesis With Prokaryotic Combined Transcription-Translation. in *Recombinant Gene Expression: Reviews and Protocols* (eds. Balbás, P. & Lorence, A.) 169–182 (Humana Press, 2004). doi:10.1385/1-59259-774-2:169.
78. Niederholtmeyer, H., Xu, L. & Maerkl, S. J. Real-time mRNA measurement during an in vitro transcription and translation reaction using binary probes. *ACS Synth. Biol.* **2**, 411–417 (2013).
79. Hook, B. & Schagat, T. Non-Radioactive Detection of Proteins Expressed in Cell-Free Expression Systems. https://www.promega.com/resources/pubhub/tpub_049-nonradioactive-detection-of-proteins-expressed-in-cell-free-expression-systems (2011).
80. Kightlinger, W. *et al.* A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nat. Commun.* **10**, 5404 (2019).
81. Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V. & Murray, R. M. Gene circuit performance characterization and resource usage in a cell-free “breadboard.” *ACS Synth. Biol.* **3**, 416–425 (2014).
82. Moore, S. J. *et al.* Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. *Proceedings of the National Academy of Sciences* 201715806 (2018) doi:10.1073/pnas.1715806115.
83. Li, J., Wang, H., Kwon, Y.-C. & Jewett, M. C. Establishing a high yielding streptomyces-based cell-free protein synthesis system. *Biotechnol. Bioeng.* **114**, 1343–1353 (2017).
84. Dudley, Q. M. *et al.* Biofoundry-assisted expression and characterization of plant proteins. *Synth. Biol.* **6**, ysab029 (2021).
85. Wick, S. *et al.* PERSIA for direct fluorescence measurements of transcription, translation, and enzyme activity in cell-free systems. *ACS Synth. Biol.* **8**, 1010–1025 (2019).
86. Wick, S. & Carr, P. A. Measurement of transcription, translation, and other enzymatic processes during cell-free expression using PERSIA. *Methods Mol. Biol.* **2433**, 169–181 (2022).
87. Stark, J. C. *et al.* On-demand biomanufacturing of protective conjugate vaccines. *Science Advances* **7**, eabe9444 (2021).
88. Jaroentomeechai, T. *et al.* Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* **9**, (2018).

89. Schoborg, J. A. *et al.* A cell-free platform for rapid synthesis and testing of active oligosaccharyltransferases. *Biotechnol. Bioeng.* **115**, 739–750 (2018).
90. Hershewe, J. M. *et al.* Improving cell-free glycoprotein synthesis by characterizing and enriching native membrane vesicles. *Nat. Commun.* **12**, 2363 (2021).
91. Rasor, B. J., Vögeli, B., Jewett, M. C. & Karim, A. S. Cell-free protein synthesis for high-throughput biosynthetic prototyping. in *Cell-Free Gene Expression* 199–215 (Springer US, 2022). doi:10.1007/978-1-0716-1998-8_12.
92. Cabantous, S. & Waldo, G. S. In vivo and in vitro protein solubility assays using split GFP. *Nat. Methods* **3**, 845–854 (2006).
93. Liu, H. & Naismith, J. H. A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr. Purif.* **63**, 102–111 (2009).
94. Boute, N. *et al.* NanoLuc Luciferase – A Multifunctional Tool for High Throughput Antibody Screening. *Front. Pharmacol.* **7**, 1–11 (2016).
95. Ryabova, L. A., Desplancq, D., Spirin, A. S. & Plückthun, A. Functional antibody production using cell-free translation: Effects of protein disulfide isomerase and chaperones. *Nat. Biotechnol.* **15**, 79–84 (1997).
96. Kawasaki, T., Gouda, M. D., Sawasaki, T., Takai, K. & Endo, Y. Efficient synthesis of a disulfide-containing protein through a batch cell-free system from wheat germ. *Eur. J. Biochem.* **270**, 4780–4786 (2003).
97. Kim, D.-M. & Swartz, J. R. Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of *Escherichia coli*. *Biotechnol. Bioeng.* **85**, 122–129 (2004).
98. Yin, G. & Swartz, J. R. Enhancing Multiple Disulfide Bonded Protein Folding in a Cell-Free System. *Biotechnol. Bioeng.* **86**, 188–195 (2004).
99. Yang, J. *et al.* Rapid expression of vaccine proteins for B-cell lymphoma in a cell-free system. *Biotechnol. Bioeng.* **89**, 503–511 (2005).
100. Ishihara, G. *et al.* Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr. Purif.* **41**, 27–37 (2005).

101. Swartz, J. Developing cell-free biology for industrial applications. *J. Ind. Microbiol. Biotechnol.* **33**, 476–485 (2006).
102. Boyer, M. E., Stapleton, J. A., Kuchenreuther, J. M., Wang, C.-W. & Swartz, J. R. Cell-free synthesis and maturation of [FeFe] hydrogenases. *Biotechnol. Bioeng.* **99**, 59–67 (2008).
103. El-Baky, N. A., Omar, S. H. & Redwan, E. M. The anti-cancer activity of human consensus interferon-alpha synthesized in cell-free system. *Protein Expr. Purif.* **80**, 61–67 (2011).
104. Matthies, D. *et al.* Cell-free expression and assembly of ATP synthase. *J. Mol. Biol.* **413**, 593–603 (2011).
105. Groff, D. *et al.* Engineering toward a bacterial “endoplasmic reticulum” for the rapid expression of immunoglobulin proteins. *MAbs* **6**, 671–678 (2014).
106. He, W. *et al.* Cell-free expression of functional receptor tyrosine kinases. *Sci. Rep.* **5**, 12896 (2015).
107. Martin, R. W. *et al.* Development of a CHO-Based Cell-Free Platform for Synthesis of Active Monoclonal Antibodies. *ACS Synth. Biol.* **6**, 1370–1379 (2017).
108. Stech, M. *et al.* Cell-free synthesis of functional antibodies using a coupled in vitro transcription-translation system based on CHO cell lysates. *Sci. Rep.* **7**, 12030 (2017).
109. Murakami, S., Matsumoto, R. & Kanamori, T. Constructive approach for synthesis of a functional IgG using a reconstituted cell-free protein synthesis system. *Sci. Rep.* **9**, 671 (2019).
110. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
111. Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature* **381**, 571–579 (1996).
112. Waudby, C. A., Dobson, C. M. & Christodoulou, J. Nature and Regulation of Protein Folding on the Ribosome. *Trends Biochem. Sci.* **xx**, 1–13 (2019).
113. Underwood, K. A., Swartz, J. R. & Puglisi, J. D. Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnol. Bioeng.* **91**, 425–435 (2005).
114. Yan, X., Hoek, T. A., Vale, R. D. & Tanenbaum, M. E. Dynamics of translation of single mRNA molecules in vivo. *Cell* **165**, 976–989 (2016).
115. Kang, S.-H. *et al.* Cell-free production of aggregation-prone proteins in soluble and active forms. *Biotechnol. Prog.* **21**, 1412–1419 (2005).

116. Welsh, J. P., Bonomo, J. & Swartz, J. R. Localization of BiP to translating ribosomes increases soluble accumulation of secreted eukaryotic proteins in an Escherichia coli cell-free system. *Biotechnol. Bioeng.* **108**, 1739–1748 (2011).
117. Sasaki, Y. *et al.* Amphiphilic polysaccharide nanogels as artificial chaperones in cell-free protein synthesis. *Macromol. Biosci.* **11**, 814–820 (2011).
118. Lim, H. J., Lee, K.-H. & Kim, D.-M. Rapid determination of effective folding agents by sequential cell-free protein synthesis. *Biochem. Eng. J.* **138**, 106–110 (2018).
119. Niwa, T., Kanamori, T., Ueda, T. & Taguchi, H. Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8937–8942 (2012).
120. Sevier, C. S. & Kaiser, C. A. Formation and transfer of disulphide bonds in living cells. *Nat. Rev. Mol. Cell Biol.* **3**, 836–847 (2002).
121. Stech, M. & Kubick, S. Cell-Free Synthesis Meets Antibody Production: A Review. *Antibodies* **4**, 12–33 (2015).
122. Cai, Q. *et al.* A simplified and robust protocol for immunoglobulin expression in Escherichia coli cell-free protein synthesis systems. *Biotechnol. Prog.* **31**, 823–831 (2015).
123. Beebe, E. T. *et al.* Robotic large-scale application of wheat cell-free translation to structural studies including membrane proteins. *N. Biotechnol.* **28**, 239–249 (2011).
124. Vinarov, D. A., Loushin Newman, C. L. & Markley, J. L. Wheat germ cell-free platform for eukaryotic protein production. *FEBS J.* **273**, 4160–4169 (2006).
125. Makino, S.-I., Goren, M. A., Fox, B. G. & Markley, J. L. Cell-free protein synthesis technology in NMR high-throughput structure determination. *Methods Mol. Biol.* **607**, 127–147 (2010).
126. Aoki, M. *et al.* Automated system for high-throughput protein production using the dialysis cell-free method. *Protein Expr. Purif.* **68**, 128–136 (2009).
127. Adiga, R. *et al.* Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat. Biomed. Eng.* **2**, 675–686 (2018).
128. Hughes, R. A. & Ellington, A. D. Synthetic DNA synthesis and assembly: Putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* **9**, a023812 (2017).

129. Caschera, F. *et al.* High-Throughput Optimization Cycle of a Cell-Free Ribosome Assembly and Protein Synthesis System. *ACS Synth. Biol.* **7**, 2841–2853 (2018).
130. Karim, A. S., Heggestad, J. T., Crowe, S. A. & Jewett, M. C. Controlling cell-free metabolism through physiochemical perturbations. *Metab. Eng.* **45**, 86–94 (2018).
131. Caschera, F. *et al.* Coping with complexity: machine learning optimization of cell-free protein synthesis. *Biotechnol. Bioeng.* **108**, 2218–2228 (2011).
132. Marshall, R., Garamella, J., Noireaux, V. & Pierson, A. High-throughput microliter-sized cell-free transcription-translation reactions for synthetic biology applications using the Echo® 550 liquid handler. *Labcyte App. Note* (2018).
133. Borkowski, O. *et al.* Large scale active-learning-guided exploration for in vitro protein production optimization. *Nat. Commun.* **11**, 1872 (2020).
134. Hammerling, M. J. *et al.* In vitro ribosome synthesis and evolution through ribosome display. *Nat. Commun.* **11**, 1108 (2020).
135. Marshall, R. & Noireaux, V. Quantitative modeling of transcription and translation of an all-E. coli cell-free system. *Sci. Rep.* **9**, 11980 (2019).
136. Kopniczky, M. B. *et al.* Cell-Free Protein Synthesis as a Prototyping Platform for Mammalian Synthetic Biology. *ACS Synth. Biol.* **9**, 144–156 (2020).
137. McManus, J. B., Emanuel, P. A., Murray, R. M. & Lux, M. W. A method for cost-effective and rapid characterization of engineered T7-based transcription factors by cell-free protein synthesis reveals insights into the regulation of T7 RNA polymerase-driven expression. *Arch. Biochem. Biophys.* **674**, 108045 (2019).
138. Angenendt, P. *et al.* Cell-free protein expression and functional assay in nanowell chip format. *Anal. Chem.* **76**, 1844–1849 (2004).
139. Contreras-Llano, L. E. & Tan, C. High-throughput screening of biomolecules using cell-free gene expression systems. *Synth. Biol.* **3**, 1–13 (2018).
140. He, M. Single step generation of protein arrays from DNA by cell-free expression and in situ immobilisation (PISA method). *Nucleic Acids Res.* **29**, 73e–773 (2001).

141. He, M. & Taussig, M. J. DiscernArray technology: a cell-free method for the generation of protein arrays from PCR DNA. *J. Immunol. Methods* **274**, 265–270 (2003).
142. Angenendt, P., Kreuzberger, J., Glökler, J. & Hoheisel, J. D. Generation of high density protein microarrays by cell-free in situ expression of unpurified PCR products. *Mol. Cell. Proteomics* **5**, 1658–1666 (2006).
143. Ramachandran, N. Self-Assembling Protein Microarrays. *Science* **305**, 86–90 (2004).
144. Ramachandran, N. *et al.* Next-generation high-density self-assembling functional protein arrays. *Nat. Methods* **5**, 535–538 (2008).
145. Zárate, X., Henderson, D. C., Phillips, K. C., Lake, A. D. & Galbraith, D. W. Development of high-yield autofluorescent protein microarrays using hybrid cell-free expression with combined *Escherichia coli* S30 and wheat germ extracts. *Proteome Sci.* **8**, 32 (2010).
146. He, M. *et al.* Printing protein arrays from DNA arrays. *Nat. Methods* **5**, 175–177 (2008).
147. Biyani, M., Osawa, T., Nemoto, N. & Ichiki, T. Microintaglio printing of biomolecules and its application to in situ production of messenger ribonucleic acid display microarray. *Appl. Phys. Express* **4**, 047001 (2011).
148. Biyani, M. *et al.* Microintaglio printing of in situ synthesized proteins enables rapid printing of high-density protein microarrays directly from DNA microarrays. *Appl. Phys. Express* **6**, 087001 (2013).
149. Georgi, V. *et al.* On-chip automation of cell-free protein synthesis: new opportunities due to a novel reaction mode. *Lab Chip* **16**, 269–281 (2016).
150. Karzbrun, E., Tayar, A. M., Noireaux, V. & Bar-Ziv, R. H. Programmable on-chip DNA compartments as artificial cells. *Science* **345**, 829–832 (2014).
151. Tayar, A. M., Karzbrun, E., Noireaux, V. & Bar-Ziv, R. H. Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proceedings of the National Academy of Sciences* **114**, 11609–11614 (2017).
152. Maerkl, S. J. & Quake, S. R. A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315**, 233–237 (2007).
153. Maerkl, S. J. & Quake, S. R. Experimental determination of the evolvability of a transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 18650–18655 (2009).

154. Swank, Z., Laohakunakorn, N. & Maerkl, S. J. Cell-free gene-regulatory network engineering with synthetic transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 5892–5901 (2019).
155. Gerber, D., Maerkl, S. J. & Quake, S. R. An in vitro microfluidic approach to generating protein-interaction networks. *Nat. Methods* **6**, 71–74 (2009).
156. Markin, C. J. *et al.* Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics. *Science* **373**, eabf8761 (2021).
157. Zhang, Y., Minagawa, Y., Kizoe, H., Miyazaki, K. & Iino, R. Accurate high-throughput screening based on digital protein synthesis in a massively parallel femtoliter droplet array. (2019).
158. Heyman, Y., Buxboim, A., Wolf, S. G., Daube, S. S. & Bar-Ziv, R. H. Cell-free protein synthesis and assembly on a biochip. *Nat. Nanotechnol.* **7**, 374–378 (2012).
159. Levy, M., Falkovich, R., Daube, S. S. & Bar-Ziv, R. H. Autonomous synthesis and assembly of a ribosomal subunit on a chip. *Sci. Adv.* **6**, eaaz6020 (2020).
160. Laohakunakorn, N. *et al.* Bottom-up construction of complex biomolecular systems with cell-free synthetic biology. *Front. Bioeng. Biotechnol.* **8**, 213 (2020).
161. Courtois, F. *et al.* An integrated device for monitoring time-dependent in vitro expression from single genes in picolitre droplets. *Chembiochem* **9**, 439–446 (2008).
162. Tawfik, D. S. & Griffiths, A. D. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* **16**, 652–656 (1998).
163. Griffiths, A. D. & Tawfik, D. S. Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *EMBO J.* **22**, 24–35 (2003).
164. Agresti, J. J. *et al.* Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4004–4009 (2010).
165. Fallah-Araghi, A., Baret, J.-C., Ryckelynck, M. & Griffiths, A. D. A completely in vitro ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution. *Lab Chip* **12**, 882–891 (2012).
166. Stapleton, J. A. & Swartz, J. R. Development of an in vitro compartmentalization screen for high-throughput directed evolution of [FeFe] hydrogenases. *PLoS One* **5**, e15275 (2010).

167. Holstein, J. M., Gylstorff, C. & Hollfelder, F. Cell-free Directed Evolution of a Protease in Microdroplets at Ultrahigh Throughput. *ACS Synth. Biol.* acssynbio.0c00538 (2021) doi:10.1021/acssynbio.0c00538.
168. Hori, Y., Kantak, C., Murray, R. M. & Abate, A. R. Cell-free extract based optimization of biomolecular circuits with droplet microfluidics. *Lab Chip* **17**, 3037–3042 (2017).
169. Sakamoto, R., Noireaux, V. & Maeda, Y. T. Anomalous scaling of gene expression in confined cell-free reactions. *Sci. Rep.* **8**, (2018).
170. Matsuura, T. *et al.* Effects of compartment size on the kinetics of intracompartamental multimeric protein synthesis. *ACS Synth. Biol.* **1**, 431–437 (2012).
171. Kato, A., Yanagisawa, M., Sato, Y. T., Fujiwara, K. & Yoshikawa, K. Cell-Sized confinement in microspheres accelerates the reaction of gene expression. *Sci. Rep.* **2**, 283 (2012).
172. Hansen, M. M. K. *et al.* Macromolecular crowding creates heterogeneous environments of gene expression in picolitre droplets. *Nat. Nanotechnol.* **11**, 191–197 (2016).
173. Olivi, L. *et al.* Towards a synthetic cell cycle. *Nat. Commun.* **12**, 4531 (2021).
174. Robinson, A. O., Venero, O. M. & Adamala, K. P. Toward synthetic life: Biomimetic synthetic cell communication. *Curr. Opin. Chem. Biol.* **64**, 165–173 (2021).
175. Hanes, J. & Plückthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 4937–4942 (1997).
176. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12297–12302 (1997).
177. Amstutz, P., Forrer, P., Zahnd, C. & Plückthun, A. In vitro display technologies: novel developments and applications. *Curr. Opin. Biotechnol.* **12**, 400–405 (2001).
178. Newton, M., Cabezas-Perusse, Y., Tong, C. L. & Seelig, B. In vitro selection of peptides and proteins - advantages of mRNA display. *ACS Synth. Biol.* (2019) doi:10.1021/acssynbio.9b00419.
179. Cotten, S. W., Zou, J., Valencia, C. A. & Liu, R. Selection of proteins with desired properties from natural proteome libraries using mRNA display. *Nat. Protoc.* **6**, 1163–1182 (2011).
180. Ju, W. *et al.* Proteome-wide identification of family member-specific natural substrate repertoire of caspases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14294–14299 (2007).

181. Du, Y. *et al.* mRNA display with library of even-distribution reveals cellular interactors of influenza virus NS1. *Nat. Commun.* **11**, 2449 (2020).
182. Gu, L. *et al.* Multiplex single-molecule interaction profiling of DNA-barcoded proteins. *Nature* **515**, 554–557 (2014).
183. Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. & Plückthun, A. Tailoring in vitro evolution for protein affinity or stability. *Proceedings of the National Academy of Sciences* **98**, 75–80 (2001).
184. Matsuura, T. & Plückthun, A. Selection based on the folding properties of proteins with ribosome display. *FEBS Lett.* **539**, 24–28 (2003).
185. Chaput, J. C. & Szostak, J. W. Evolutionary optimization of a nonbiological ATP binding protein for improved folding stability. *Chem. Biol.* **11**, 865–874 (2004).
186. Smith, M. D. *et al.* Structural insights into the evolution of a non-biological protein: Importance of surface residues in protein fold optimization. *PLoS One* **2**, e467 (2007).
187. Morelli, A., Haugner, J. & Seelig, B. Thermostable artificial enzyme isolated by in vitro selection. *PLoS One* **9**, 1–8 (2014).
188. Seelig, B. mRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries. *Nat. Protoc.* **6**, 540–552 (2011).
189. Sierrecki, E. *et al.* A cell-free approach to accelerate the study of protein–protein interactions in vitro. *Interface Focus* **3**, 20130018 (2013).
190. Wiens, M. D. & Campbell, R. E. Surveying the landscape of optogenetic methods for detection of protein–protein interactions. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **10**, 1–15 (2018).
191. Porter, J. R., Stains, C. I., Jester, B. W. & Ghosh, I. A general and rapid cell-free approach for the interrogation of protein-protein, protein-DNA, and protein-RNA interactions and their antagonists utilizing split-protein reporters. *J. Am. Chem. Soc.* **130**, 6488–6497 (2008).
192. Campbell, S. T., Carlson, K. J., Buchholz, C. J., Helmers, M. R. & Ghosh, I. Mapping the BH3 Binding Interface of Bcl-x L , Bcl-2, and Mcl-1 Using Split-Luciferase Reassembly. *Biochemistry* **54**, 2632–2643 (2015).

193. Henchey, L. K., Porter, J. R., Ghosh, I. & Arora, P. S. High specificity in protein recognition by hydrogen-bond-surrogate α -helices: Selective inhibition of the p53/MDM2 complex. *ChemBiochem* **11**, 2104–2107 (2010).
194. Porter, J. R. *et al.* Profiling small molecule inhibitors against helix-receptor interactions: The Bcl-2 family inhibitor BH3I-1 potently inhibits p53/hDM2. *Chem. Commun.* **46**, 8020–8022 (2010).
195. Doi, N. *et al.* Novel fluorescence labeling and high-throughput assay technologies for in vitro analysis of protein interactions. *Genome Res.* **12**, 487–492 (2002).
196. Mureev, S., Kovtun, O., Nguyen, U. T. T. & Alexandrov, K. Species-independent translational leaders facilitate cell-free expression. *Nat. Biotechnol.* **27**, 747–752 (2009).
197. Sierrecki, E. *et al.* Rapid Mapping of Interactions between Human SNX-BAR Proteins Measured In Vitro by AlphaScreen and Single-molecule Spectroscopy. *Mol. Cell. Proteomics* **13**, 2233–2245 (2014).
198. Gambin Y. N. Ariotti K. A. McMahon M. Bastiani E. Sierrecki O. Kovtun M. E. Polinkovsky A. Magenau Jung, W. *et al.* Single-molecule analysis reveals self assembly and nanoscale segregation of two distinct cavin subcomplexes on caveolae. *Elife* **3**, e01434 (2013).
199. Han, S. P. *et al.* Cortactin scaffolds Arp2/3 and WAVE2 at the epithelial zonula adherens. *J. Biol. Chem.* **289**, 7764–7775 (2014).
200. Aydin, S. A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA. *Peptides* **72**, 4–15 (2015).
201. Ojima-Kato, T. *et al.* Rapid Generation of Monoclonal Antibodies from Single B Cells by Ecobody Technology. *Antibodies* **7**, 38 (2018).
202. Ojima-Kato, T., Hashimura, D., Kojima, T., Minabe, S. & Nakano, H. In vitro generation of rabbit anti-*Listeria monocytogenes* monoclonal antibody using single cell based RT-PCR linked cell-free expression systems. *J. Immunol. Methods* **427**, 58–65 (2015).
203. Beaudet, L. *et al.* AlphaLISA immunoassays: the no-wash alternative to ELISAs for research and drug discovery. *Nat. Methods* **5**, an8–an9 (2008).
204. Matsuoka, K., Komori, H., Nose, M., Endo, Y. & Sawasaki, T. Simple screening method for autoantigen proteins using the N-terminal biotinylated protein library produced by wheat cell-free synthesis. *J. Proteome Res.* **9**, 4264–4273 (2010).

205. Zhou, Y. *et al.* Engineering bacterial transcription regulation to create a synthetic in vitro two-hybrid system for protein interaction assays. *J. Am. Chem. Soc.* **136**, 14031–14038 (2014).
206. Ceroni, A. *et al.* Systematic analysis of the IgG antibody immune response against varicella zoster virus (VZV) using a self-assembled protein microarray. *Mol. Biosyst.* **6**, 1604–1610 (2010).
207. Wang, J. *et al.* A versatile protein microarray platform enabling antibody profiling against denatured proteins. *Proteomics Clin. Appl.* **7**, 378–383 (2013).
208. Sibani, S. & LaBaer, J. Immunoprofiling using NAPPA protein microarrays. *Methods Mol. Biol.* **723**, 149–161 (2011).
209. Liang, W. *et al.* Measurement of small molecule binding kinetics on a protein microarray by plasmonic-based electrochemical impedance imaging. *Anal. Chem.* **86**, 9860–9865 (2014).
210. Díez, P. *et al.* NAPPA as a real new method for protein microarray generation. *Microarrays (Basel)* **4**, 214–227 (2015).
211. Stoevesandt, O., Taussig, M. J. & He, M. Protein microarrays: high-throughput tools for proteomics. *Expert Rev. Proteomics* **6**, 145–157 (2009).
212. He, M., Stoevesandt, O. & Taussig, M. J. In situ synthesis of protein arrays. *Curr. Opin. Biotechnol.* **19**, 4–9 (2008).
213. Nand, A. *et al.* In situ protein microarrays capable of real-time kinetics analysis based on surface plasmon resonance imaging. *Anal. Biochem.* **464**, 30–35 (2014).
214. Garcia-Cordero, J. L. & Maerkl, S. J. Mechanically Induced Trapping of Molecular Interactions and Its Applications. *J. Lab. Autom.* **21**, 356–367 (2016).
215. Cui, N. *et al.* A mix-and-read drop-based in vitro two-hybrid method for screening high-affinity peptide binders. *Sci. Rep.* **6**, 22575 (2016).
216. Layton, C. J., McMahon, P. L. & Greenleaf, W. J. Large-scale, quantitative protein assays on a high-throughput DNA sequencing chip. *Mol. Cell* **73**, 1075-1082.e4 (2019).
217. Svensen, N., Peersen, O. B. & Jaffrey, S. R. Peptide synthesis on a next-generation DNA sequencing platform. *ChemBiochem* **17**, 1628–1635 (2016).
218. Linciano, S., Pluda, S., Bacchin, A. & Angelini, A. Molecular evolution of peptides by yeast surface display technology. *Medchemcomm* **10**, 1569–1580 (2019).

219. Fujii, S. *et al.* Liposome display for in vitro selection and evolution of membrane proteins. *Nat. Protoc.* **9**, 1578–1591 (2014).
220. Luginbühl, B. *et al.* Directed Evolution of an Anti-prion Protein scFv Fragment to an Affinity of 1 pM and its Structural Interpretation. *J. Mol. Biol.* **363**, 75–97 (2006).
221. Hanes, J., Schaffitzel, C., Knappik, A. & Plückthun, A. Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat. Biotechnol.* **18**, 1287–1292 (2000).
222. Passioura, T. & Suga, H. A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem. Commun.* **53**, 1931–1940 (2017).
223. Plückthun, A. Ribosome display: a perspective. *Methods Mol. Biol.* **805**, 3–28 (2012).
224. Lipovsek, D. & Plückthun, A. In-vitro protein evolution by ribosome display and mRNA display. *J. Immunol. Methods* **290**, 51–67 (2004).
225. Zahnd, C., Amstutz, P. & Plückthun, A. Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods* **4**, 269–279 (2007).
226. Stafford, R. L. *et al.* In vitro Fab display: a cell-free system for IgG discovery. *Protein Eng. Des. Sel.* **27**, 97–109 (2014).
227. Binz, H. K. *et al.* High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* **22**, 575–582 (2004).
228. Zimmermann, I. *et al.* Synthetic single domain antibodies for the conformational trapping of membrane proteins. *Elife* **7**, (2018).
229. Custódio, T. F. *et al.* Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. *Nat. Commun.* **11**, 5588 (2020).
230. Zimmermann, I. *et al.* Generation of synthetic nanobodies against delicate proteins. *Nat. Protoc.* (2020) doi:10.1038/s41596-020-0304-x.
231. Chen, X., Gentili, M., Hacohen, N. & Regev, A. A cell-free nanobody engineering platform rapidly generates SARS-CoV-2 neutralizing nanobodies. *Nat. Commun.* **12**, 5506 (2021).

232. Veessler, D. *et al.* Crystal structure and function of a DARPIn neutralizing inhibitor of lactococcal phage TP901-1: comparison of DARPIn and camelid VHH binding mode. *J. Biol. Chem.* **284**, 30718–30726 (2009).
233. Zahnd, C. *et al.* A designed ankyrin repeat protein evolved to picomolar affinity to Her2. *J. Mol. Biol.* **369**, 1015–1028 (2007).
234. Galán, A. *et al.* Library-based display technologies: where do we stand? *Mol. Biosyst.* **12**, 2342–2358 (2016).
235. In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **414**, 405–408 (1997).
236. Wilson, D. S., Keefe, A. D. & Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3750–3755 (2001).
237. Xu, L. *et al.* Directed evolution of high-affinity antibody mimics using mRNA display. *Chem. Biol.* **9**, 933–942 (2002).
238. Suzuki, T. *et al.* Anti-survivin single-domain antibodies derived from an artificial library including three synthetic random regions by in vitro selection using cDNA display. *Biochem. Biophys. Res. Commun.* **503**, 2054–2060 (2018).
239. Sumida, T., Yanagawa, H. & Doi, N. In vitro selection of fab fragments by mRNA display and gene-linking emulsion PCR. *J. Nucleic Acids* **2012**, 371379 (2012).
240. Keefe, A. D. & Szostak, J. W. Functional proteins from a random-sequence library. *Nature* **410**, 715–718 (2001).
241. Huang, Y., Wiedmann, M. M. & Suga, H. RNA Display Methods for the Discovery of Bioactive Macrocycles. *Chem. Rev.* **119**, 10360–10391 (2019).
242. Yamaguchi, J. *et al.* cDNA display: a novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA-protein fusions. *Nucleic Acids Res.* **37**, e108 (2009).
243. Ishizawa, T., Kawakami, T., Reid, P. C. & Murakami, H. TRAP display: a high-speed selection method for the generation of functional polypeptides. *J. Am. Chem. Soc.* **135**, 5433–5440 (2013).
244. Kondo, T. *et al.* cDNA TRAP display for rapid and stable in vitro selection of antibody-like proteins. *Chem. Commun. (Camb.)* **57**, 2416–2419 (2021).

245. Odegrip, R. *et al.* CIS display: In vitro selection of peptides from libraries of protein-DNA complexes. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2806–2810 (2004).
246. Reiersen, H. *et al.* Covalent antibody display—an in vitro antibody-DNA library selection system. *Nucleic Acids Res.* **33**, e10 (2005).
247. Stein, V., Sielaff, I., Johnsson, K. & Hollfelder, F. A covalent chemical genotype-phenotype linkage for in vitro protein evolution. *Chembiochem* **8**, 2191–2194 (2007).
248. Houlihan, G., Gatti-Lafranconi, P., Lowe, D. & Hollfelder, F. Directed evolution of anti-HER2 DARPins by SNAP display reveals stability/function trade-offs in the selection process. *Protein Eng. Des. Sel.* **28**, 269–279 (2015).
249. Gan, R., Yamanaka, Y., Kojima, T. & Nakano, H. Microbeads display of proteins using emulsion PCR and cell-free protein synthesis. *Biotechnol. Prog.* **24**, 1107–1114 (2008).
250. Nord, O., Uhlén, M. & Nygren, P.-A. Microbead display of proteins by cell-free expression of anchored DNA. *J. Biotechnol.* **106**, 1–13 (2003).
251. Sepp, A., Tawfik, D. S. & Griffiths, A. D. Microbead display by in vitro compartmentalisation: selection for binding using flow cytometry. *FEBS Lett.* **532**, 455–458 (2002).
252. Nussinov, R. How do dynamic cellular signals travel long distances? *Mol. Biosyst.* **8**, 22–26 (2012).
253. Reinke, A. W., Baek, J., Ashenberg, O. & Keating, A. E. Networks of bZIP protein-protein interactions diversified over a billion years of evolution. *Science* **340**, 730–734 (2013).
254. Antebi, Y. E. *et al.* Combinatorial Signal Perception in the BMP Pathway. *Cell* **170**, 1184–1196.e24 (2017).
255. Harris, B. Z. & Lim, W. A. Mechanism and role of PDZ domains in signaling complex assembly. *J. Cell Sci.* **114**, 3219–3231 (2001).
256. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588 (2006).
257. Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196–1201 (2011).
258. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).

259. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
260. Tamsir, A., Tabor, J. J. & Voigt, C. A. Robust multicellular computing using genetically encoded NOR gates and chemical “wires.” *Nature* **469**, 212–215 (2011).
261. Siuti, P., Yazbek, J. & Lu, T. K. Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* **31**, 448–452 (2013).
262. Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P. & Endy, D. Amplifying genetic logic gates. *Science* **340**, 599–603 (2013).
263. Weinberg, B. H. *et al.* Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nat. Biotechnol.* **35**, 453–462 (2017).
264. Ausländer, S., Ausländer, D., Müller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocomputers. *Nature* **487**, 123–127 (2012).
265. Khalil, A. S. *et al.* A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150**, 647–658 (2012).
266. Roquet, N., Soleimany, A. P., Ferris, A. C., Aaronson, S. & Lu, T. K. Synthetic recombinase-based state machines in living cells. *Science* **353**, aad8559 (2016).
267. Andrews, L. B., Nielsen, A. A. K. & Voigt, C. A. Cellular checkpoint control using programmable sequential logic. *Science* **361**, eaap8987 (2018).
268. Angelici, B., Mailand, E., Haefliger, B. & Benenson, Y. Synthetic Biology Platform for Sensing and Integrating Endogenous Transcriptional Inputs in Mammalian Cells. *Cell Rep.* **16**, 2525–2537 (2016).
269. Lohmueller, J. J., Armel, T. Z. & Silver, P. A. A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Res.* **40**, 5180–5187 (2012).
270. Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925–939 (2014).
271. Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).
272. Rinaudo, K. *et al.* A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat. Biotechnol.* **25**, 795–801 (2007).

273. Wroblewska, L. *et al.* Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. *Nat. Biotechnol.* **33**, 839–841 (2015).
274. Park, S.-H., Zarrinpar, A. & Lim, W. A. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**, 1061–1064 (2003).
275. Howard, P. L., Chia, M. C., Del Rizzo, S., Liu, F.-F. & Pawson, T. Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11267–11272 (2003).
276. Groves, B., Khakhar, A., Nadel, C. M., Gardner, R. G. & Seelig, G. Rewiring MAP kinases in *Saccharomyces cerevisiae* to regulate novel targets through ubiquitination. *Elife* **5**, (2016).
277. Morsut, L. *et al.* Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **164**, 780–791 (2016).
278. Dueber, J. E., Yeh, B. J., Chak, K. & Lim, W. A. Reprogramming Control of an Allosteric Signaling Switch Through Modular Recombination. *Science* **301**, 1904–1908 (2003).
279. Dueber, J. E., Mirsky, E. A. & Lim, W. A. Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nat. Biotechnol.* **25**, 660–662 (2007).
280. Smith, A. J., Thomas, F., Shoemark, D., Woolfson, D. N. & Savery, N. J. Guiding Biomolecular Interactions in Cells Using de Novo Protein-Protein Interfaces. *ACS Synth. Biol.* **8**, 1284–1293 (2019).
281. Gao, X. J., Chong, L. S., Kim, M. S. & Elowitz, M. B. Programmable protein circuits in living cells. *Science* **361**, 1252–1258 (2018).
282. Fink, T. *et al.* Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat. Chem. Biol.* **15**, 115–122 (2019).
283. Chen, Z. *et al.* Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111 (2019).
284. Prehoda, K. E., Scott, J. A., Mullins, R. D. & Lim, W. A. Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex. *Science* **290**, 801–806 (2000).
285. Yu, B. *et al.* Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1. *Cell* **140**, 246–256 (2010).
286. Dyer, K. N. *et al.* High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. *Methods Mol. Biol.* **1091**, 245–258 (2014).

287. Ruotolo, B. T. & Robinson, C. V. Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.* **10**, 402–408 (2006).
288. VanAernum, Z. L. *et al.* Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. *Nat. Protoc.* (2020) doi:10.1038/s41596-019-0281-0.
289. Pause, A., Peterson, B., Schaffar, G., Stearman, R. & Klausner, R. D. Studying interactions of four proteins in the yeast two-hybrid system: structural resemblance of the pVHL/elongin BC/hCUL-2 complex with the ubiquitin ligase complex SKP1/cullin/F-box protein. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9533–9538 (1999).
290. Sandrock, B. & Egly, J. M. A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIH. *J. Biol. Chem.* **276**, 35328–35333 (2001).
291. Dixon, A. S. *et al.* NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
292. Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **5**, 1–8 (2015).
293. Maude, S. L. *et al.* Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. *N. Engl. J. Med.* **378**, 439–448 (2018).
294. Neelapu, S. S. *et al.* Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. *N. Engl. J. Med.* **377**, 2531–2544 (2017).
295. Fraietta, J. A. *et al.* Determinants of response and resistance to CD19 chimeric antigen receptor (CAR) T cell therapy of chronic lymphocytic leukemia. *Nat. Med.* **24**, 563–571 (2018).
296. June, C. H., O'Connor, R. S., Kawalekar, O. U., Ghassemi, S. & Milone, M. C. CAR T cell immunotherapy for human cancer. *Science* **359**, 1361–1365 (2018).
297. Long, A. H. *et al.* 4-1BB costimulation ameliorates T cell exhaustion induced by tonic signaling of chimeric antigen receptors. *Nat. Med.* **21**, 581–590 (2015).
298. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* **15**, 486–499 (2015).
299. John Wherry, E. *et al.* Molecular Signature of CD8+ T Cell Exhaustion during Chronic Viral Infection. *Immunity* **27**, (2007).

300. Pauken, K. E. *et al.* Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* **354**, 1160–1165 (2016).
301. Wilken, M. S. *et al.* Promoter keyholes enable specific and persistent multi-gene expression programs in primary T cells without genome modification. *bioRxiv* 2020.02.19.956730 (2020) doi:10.1101/2020.02.19.956730.
302. Boyken, S. E. *et al.* De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **352**, 680–687 (2016).
303. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).
304. Aranda-Díaz, A., Mace, K., Zuleta, I., Harrigan, P. & El-Samad, H. Robust Synthetic Circuits for Two-Dimensional Control of Gene Expression in Yeast. *ACS Synth. Biol.* **6**, 545–554 (2017).
305. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
306. Myers, J. K., Pace, C. N. & Scholtz, J. M. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148 (1995).
307. Bernadó, P., Blackledge, M. & Sancho, J. Sequence-specific solvent accessibilities of protein residues in unfolded protein ensembles. *Biophys. J.* **91**, 4536–4543 (2006).
308. Estrada, J., Bernadó, P., Blackledge, M. & Sancho, J. ProtSA: a web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble. *BMC Bioinformatics* **10**, 104 (2009).
309. Rambo, R. P. & Tainer, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559–571 (2011).
310. Schneidman-Duhovny, D., Hammel, M. & Sali, A. FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.* **38**, W540-4 (2010).
311. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974 (2013).
312. Schiestl, R. H. & Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* **16**, 339–346 (1989).

313. Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 9578–9582 (1991).
314. Bartel, P. L., Roecklein, J. A., SenGupta, D. & Fields, S. A protein linkage map of Escherichia coli bacteriophage T7. *Nat. Genet.* **12**, 72–77 (1996).
315. VanAernum, Z. L. *et al.* Surface-Induced Dissociation of Noncovalent Protein Complexes in an Extended Mass Range Orbitrap Mass Spectrometer. *Anal. Chem.* **91**, 3611–3618 (2019).
316. 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).
317. Jewett, M. C. & Swartz, J. R. Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **86**, 19–26 (2004).
318. Silverman, A. D., Kelley-Loughnane, N., Lucks, J. B. & Jewett, M. C. Deconstructing Cell-Free Extract Preparation for in Vitro Activation of Transcriptional Genetic Circuitry. *ACS Synth. Biol.* **8**, 403–414 (2019).
319. Koczula, K. M. & Gallotta, A. Lateral flow assays. *Essays Biochem.* **60**, 111–120 (2016).
320. Lu, R.-M. *et al.* Development of therapeutic antibodies for the treatment of diseases. *J. Biomed. Sci.* **27**, 1 (2020).
321. Mullard, A. 2020 FDA drug approvals. *Nat. Rev. Drug Discov.* **20**, 85–90 (2021).
322. Vazquez-Lombardi, R. *et al.* Transient expression of human antibodies in mammalian cells. *Nat. Protoc.* **13**, 99–117 (2018).
323. Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* (2020) doi:10.1038/s41586-020-2349-y.
324. Starr, T. N. *et al.* SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature* **597**, 97–102 (2021).
325. Copin, R. *et al.* The monoclonal antibody combination REGEN-COV protects against SARS-CoV-2 mutational escape in preclinical and human studies. *Cell* **184**, 3949-3961.e11 (2021).
326. Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014–1018 (2020).

- 327.Liu, Z. *et al.* Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* **29**, 477-488.e4 (2021).
- 328.Hershewe, J., Kightlinger, W. & Jewett, M. C. Cell-free systems for accelerating glycoprotein expression and biomanufacturing. *J. Ind. Microbiol. Biotechnol.* **47**, 977–991 (2020).
- 329.Ojima-Kato, T. *et al.* 'Zipbody' leucine zipper-fused Fab in E. coli in vitro and in vivo expression systems. *Protein Eng. Des. Sel.* **29**, 149–157 (2016).
- 330.Yasgar, A., Jadhav, A., Simeonov, A. & Coussens, N. P. Alphascreen-based assays: Ultra-high-throughput screening for small-molecule inhibitors of challenging enzymes and protein-protein interactions. in *Methods in Molecular Biology* vol. 1439 77–98 (2016).
- 331.Brouwer, P. J. M. *et al.* Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643–650 (2020).
- 332.Barnes, C. O. *et al.* Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* 2020.05.28.121533 (2020) doi:10.1016/j.cell.2020.06.025.
- 333.Du, S. *et al.* Structurally Resolved SARS-CoV-2 Antibody Shows High Efficacy in Severely Infected Hamsters and Provides a Potent Cocktail Pairing Strategy. *Cell* 1–11 (2020) doi:10.1016/j.cell.2020.09.035.
- 334.Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042.e21 (2020).
- 335.Kreye, J. *et al.* A Therapeutic Non-self-reactive SARS-CoV-2 Antibody Protects from Lung Pathology in a COVID-19 Hamster Model. *Cell* **183**, 1058-1069.e19 (2020).
- 336.Ju, B. *et al.* Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115–119 (2020).
- 337.Shi, R. *et al.* A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120–124 (2020).
- 338.Liu, L. *et al.* Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).

339. Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* (2020) doi:10.1038/s41586-020-2852-1.
340. Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor-binding domains of SARS-CoV-2 and SARS-CoV. *Science* **633**, eabb7269 (2020).
341. Wu, Y. *et al.* A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* **1278**, eabc2241 (2020).
342. Cao, Y. *et al.* Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. *Cell* (2020) doi:10.1016/j.cell.2020.05.025.
343. Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650–655 (2020).
344. Hansen, J. *et al.* Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **0827**, eabd0827 (2020).
345. Yuan, M. *et al.* Structural basis of a shared antibody response to SARS-CoV-2. *Science* **369**, 1119–1123 (2020).
346. Tortorici, M. A. *et al.* Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **4**, eabe3354 (2020).
347. Ding, R. *et al.* Rapid isolation of antigen-specific B-cells using droplet microfluidics. *RSC Adv.* **10**, 27006–27013 (2020).
348. Ritz, D., Lim, J., Reynolds, C. M., Poole, L. B. & Beckwith, J. Conversion of a peroxiredoxin into a disulfide reductase by a triplet repeat expansion. *Science* **294**, 158–160 (2001).
349. Dopp, J. L. & Reuel, N. F. Simple, functional, inexpensive cell extract for in vitro prototyping of proteins with disulfide bonds. *Biochem. Eng. J.* **164**, 107790 (2020).
350. Carter, P. *et al.* Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4285–4289 (1992).
351. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).

352. Greaney, A. J. *et al.* Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**, 463–476.e6 (2021).
353. Tan, C. W. *et al.* A SARS-CoV-2 surrogate virus neutralization test based on antibody-mediated blockage of ACE2–spike protein–protein interaction. *Nat. Biotechnol.* **38**, 1073–1078 (2020).
354. Wrobel, A. G. *et al.* Antibody-mediated disruption of the SARS-CoV-2 spike glycoprotein. *Nat. Commun.* **11**, 5337 (2020).
355. Kightlinger, W. *et al.* Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. *Nat. Chem. Biol.* **14**, 627–635 (2018).
356. Ojima-Kato, T., Nagai, S. & Nakano, H. N-terminal SKIK peptide tag markedly improves expression of difficult-to-express proteins in *Escherichia coli* and *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **123**, 540–546 (2017).
357. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
358. Rabe, B. A. & Cepko, C. A Simple Enhancement for Gibson Isothermal Assembly. *bioRxiv* 2020.06.14.150979 (2020) doi:10.1101/2020.06.14.150979.
359. Matsuda, T., Watanabe, S. & Kigawa, T. Cell-free synthesis system suitable for disulfide-containing proteins. *Biochem. Biophys. Res. Commun.* **431**, 296–301 (2013).
360. Bouvet, J.-P. Immunoglobulin Fab fragment-binding proteins. *Int. J. Immunopharmacol.* **16**, 419–424 (1994).
361. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281–292.e6 (2020).
362. Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).
363. Crook, Z. R., Nairn, N. W. & Olson, J. M. Miniproteins as a Powerful Modality in Drug Development. *Trends Biochem. Sci.* **45**, 332–346 (2020).
364. McCallum, M. *et al.* Molecular basis of immune evasion by the Delta and Kappa SARS-CoV-2 variants. *Science* **374**, 1621–1626 (2021).

365. McCallum, M. *et al.* SARS-CoV-2 immune evasion by the B.1.427/B.1.429 variant of concern. *Science* **373**, 648–654 (2021).
366. Dejnirattisai, W. *et al.* Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* **184**, 2939–2954.e9 (2021).
367. Chen, R. E. *et al.* Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**, 717–726 (2021).
368. Chen, R. E. *et al.* In vivo monoclonal antibody efficacy against SARS-CoV-2 variant strains. *Nature* **596**, 103–108 (2021).
369. Collier, D. A. *et al.* Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature* **593**, 136–141 (2021).
370. Cameroni, E. *et al.* Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **602**, 664–670 (2021).
371. VanBlargan, L. *et al.* An infectious SARS-CoV-2 B.1.1.529 Omicron virus escapes neutralization by therapeutic monoclonal antibodies. *Nature Medicine* **28**, 490–495 (2022).
372. McCallum, M. *et al.* Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. *bioRxiv* 2021.12.28.474380 (2021) doi:10.1101/2021.12.28.474380.
373. Pinto, D. *et al.* Broad betacoronavirus neutralization by a stem helix-specific human antibody. *Science* **373**, 1109–1116 (2021).
374. Jette, C. A. *et al.* Broad cross-reactivity across sarbecoviruses exhibited by a subset of COVID-19 donor-derived neutralizing antibodies. *Cell Rep.* **36**, 109760 (2021).
375. VanBlargan, L. A. *et al.* A potently neutralizing SARS-CoV-2 antibody inhibits variants of concern by utilizing unique binding residues in a highly conserved epitope. *Immunity* **54**, 2399–2416.e6 (2021).
376. Park, Y.-J. *et al.* Antibody-mediated broad sarbecovirus neutralization through ACE2 molecular mimicry. *Science* **375**, 449–454 (2022).
377. Cao, L. *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
378. Case, J. B. *et al.* Ultrapotent miniproteins targeting the SARS-CoV-2 receptor-binding domain protect against infection and disease. *Cell Host Microbe* **29**, 1151–1161.e5 (2021).

379. Javanmardi, K. *et al.* Rapid characterization of spike variants via mammalian cell surface display. *Molecular Cell* **81**, P5099-5111.E8 (2021).
380. Silverman, J. *et al.* Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains. *Nat. Biotechnol.* **23**, 1556–1561 (2005).
381. Detalle, L. *et al.* Generation and Characterization of ALX-0171, a Potent Novel Therapeutic Nanobody for the Treatment of Respiratory Syncytial Virus Infection. *Antimicrob. Agents Chemother.* **60**, 6–13 (2016).
382. Strauch, E. M. *et al.* Computational design of trimeric influenza-neutralizing proteins targeting the hemagglutinin receptor binding site. *Nat. Biotechnol.* **35**, 667–671 (2017).
383. Bracken, C. J. *et al.* Bi-paratopic and multivalent VH domains block ACE2 binding and neutralize SARS-CoV-2. *Nat. Chem. Biol.* **17**, 113–121 (2021).
384. Koenig, P.-A. *et al.* Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. *Science* **371**, eabe6230 (2021).
385. Xiang, Y. *et al.* Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2. *Science* **370**, 1479–1484 (2020).
386. Walser, M. *et al.* Highly potent anti-SARS-CoV-2 multi-DARPin therapeutic candidates. *bioRxiv* 2020.08.25.256339 (2020) doi:10.1101/2020.08.25.256339.
387. Schoof, M. *et al.* An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. *Science* **370**, 1473–1479 (2020).
388. Linsky, T. W. *et al.* De novo design of potent and resilient hACE2 decoys to neutralize SARS-CoV-2. *Science* **370**, 1208–1214 (2020).
389. Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).
390. Palomo, C. *et al.* Trivalency of a Nanobody Specific for the Human Respiratory Syncytial Virus Fusion Glycoprotein Drastically Enhances Virus Neutralization and Impacts Escape Mutant Selection. *Antimicrob. Agents Chemother.* **60**, 6498–6509 (2016).
391. Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).

392. Mammen, M., Choi, S.-K. & Whitesides, G. M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed Engl.* **37**, 2754–2794 (1998).
393. Pollard, T. D. & De La Cruz, E. M. Take advantage of time in your experiments: a guide to simple, informative kinetics assays. *Mol. Biol. Cell* **24**, 1103–1110 (2013).
394. Güthe, S. *et al.* Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. *J. Mol. Biol.* **337**, 905–915 (2004).
395. Gräwe, A. & Stein, V. Linker Engineering in the Context of Synthetic Protein Switches and Sensors. *Trends Biotechnol.* **39**, 731–744 (2020).
396. Ke, Z. *et al.* Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* **588**, 498–502 (2020).
397. Turoňová, B. *et al.* In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* **370**, 203–208 (2020).
398. Yao, H. *et al.* Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **183**, 730–738.e13 (2020).
399. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
400. Yeh, H. W. *et al.* Red-shifted luciferase-luciferin pairs for enhanced bioluminescence imaging. *Nat. Methods* **14**, 971–974 (2017).
401. Kim, B. B. *et al.* A red fluorescent protein with improved monomericity enables ratiometric voltage imaging with ASAP3. *Scientific Rep* **12**, 3678 (2020).
402. Beumer, J. *et al.* A CRISPR/Cas9 genetically engineered organoid biobank reveals essential host factors for coronaviruses. *Nat. Commun.* **12**, 5498 (2021).
403. Penninger, J. M., Mirazimi, A. & Montserrat, N. Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell* **181**, 905–913.e7 (2020).
404. Freedman, B. S. *et al.* Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids. *Nat. Commun.* **6**, 8715 (2015).
405. Alsoussi, W. B. *et al.* A Potently Neutralizing Antibody Protects Mice against SARS-CoV-2 Infection. *The Journal of Immunology* **205**, 915–922 (2020).

406. Errico, J. M. *et al.* Structural mechanism of SARS-CoV-2 neutralization by two murine antibodies targeting the RBD. *Cell Rep.* **37**, 109881 (2021).
407. Su, W. *et al.* Neutralizing Monoclonal Antibodies That Target the Spike Receptor Binding Domain Confer Fc Receptor-Independent Protection against SARS-CoV-2 Infection in Syrian Hamsters. *MBio* **12**, e0239521 (2021).
408. Wertz, G. W., Moudy, R. & Ball, L. A. Adding genes to the RNA genome of vesicular stomatitis virus: positional effects on stability of expression. *J. Virol.* **76**, 7642–7650 (2002).
409. Hassan, A. O. *et al.* A SARS-CoV-2 infection model in mice demonstrates protection by neutralizing antibodies. *Cell* **182**, 744–753.e4 (2020).
410. Walls, A. C. *et al.* Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **176**, 1026–1039.e15 (2019).
411. Lempp, F. A. *et al.* Lectins enhance SARS-CoV-2 infection and influence neutralizing antibodies. *Nature* **598**, 342–347 (2021).
412. Guo, L. *et al.* Engineered trimeric ACE2 binds viral spike protein and locks it in “Three-up” conformation to potently inhibit SARS-CoV-2 infection. *Cell Res.* **31**, 98–100 (2021).
413. Glasgow, A. *et al.* Engineered ACE2 receptor traps potently neutralize SARS-CoV-2. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 28046–28055 (2020).
414. Higuchi, Y. *et al.* Engineered ACE2 receptor therapy overcomes mutational escape of SARS-CoV-2. *Nat. Commun.* **12**, 3802 (2021).
415. Weinreich, D. M. *et al.* REGN-COV2, a Neutralizing Antibody Cocktail, in Outpatients with Covid-19. *N. Engl. J. Med.* **384**, 238–251 (2021).
416. Chen, P. *et al.* SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19. *N. Engl. J. Med.* **384**, 229–237 (2020).
417. O'Brien, M. P. *et al.* Subcutaneous REGEN-COV Antibody Combination to Prevent Covid-19. *N. Engl. J. Med.* **385**, 1184–1195 (2021).
418. Chevalier, A. *et al.* Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017).

419. Quijano-Rubio, A. *et al.* De novo design of modular and tunable protein biosensors. *Nature* **591**, 482–487 (2021).
420. Elledge, S. K. *et al.* Engineering luminescent biosensors for point-of-care SARS-CoV-2 antibody detection. *Nat. Biotechnol.* **39**, 928–935 (2021).
421. Crawford, K. H. D. *et al.* Protocol and Reagents for Pseudotyping Lentiviral Particles with SARS-CoV-2 Spike Protein for Neutralization Assays. *Viruses* **12**, 513 (2020).
422. Hassan, A. O. *et al.* An intranasal vaccine durably protects against SARS-CoV-2 variants in mice. *Cell Rep.* **36**, 109452 (2021).
423. Case, J. B. *et al.* Neutralizing Antibody and Soluble ACE2 Inhibition of a Replication-Competent VSV-SARS-CoV-2 and a Clinical Isolate of SARS-CoV-2. *Cell Host Microbe* **28**, 475–485.e5 (2020).
424. Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6**, e20161 (2011).
425. Hsia, Y. *et al.* Design of multi-scale protein complexes by hierarchical building block fusion. *Nat. Commun.* **12**, 2294 (2021).
426. Huang, P.-S. *et al.* RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS One* **6**, e24109 (2011).
427. Lin, Y.-R. *et al.* Control over overall shape and size in de novo designed proteins. *Proc. Nat. Acad. Sci. U. S. A.* **112**, E5478–E5485 (2015).
428. Rocklin, G. J. *et al.* Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* **357**, 168–175 (2017).
429. Dang, B. *et al.* De novo design of covalently constrained mesosize protein scaffolds with unique tertiary structures. *Proc. Nat. Acad. Sci. U. S. A.* **114**, 10852–10857 (2017).
430. Khatib, F. *et al.* Algorithm discovery by protein folding game players. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18949–18953 (2011).
431. Greaney, A. J. *et al.* Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44–57.e9 (2021).

432. Starr, T. N., Greaney, A. J., Dingens, A. S. & Bloom, J. D. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Rep Med* **2**, 100255 (2021).
433. Daddona, P. E. *et al.* Method of removing endotoxin contaminants. *Patent* WO1989003885A1 (1989).
434. Pallesen, J. *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E7348–E7357 (2017).
435. Kirchdoerfer, R. N. *et al.* Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci. Rep.* **8**, 15701 (2018).
436. Russo, C. J. & Passmore, L. A. Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. *Science* **346**, 1377–1380 (2014).
437. Suloway, C. *et al.* Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
438. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).
439. 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).
440. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
441. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
442. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).
443. Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
444. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
445. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

446. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
447. Wang, R. Y.-R. *et al.* Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *Elife* **5**, e17219 (2016).
448. Frenz, B., Walls, A. C., Egelman, E. H., Veerler, D. & DiMaio, F. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nat. Methods* **14**, 797–800 (2017).
449. Corman, V. M. *et al.* Comparison of seven commercial SARS-CoV-2 rapid point-of-care antigen tests: a single-centre laboratory evaluation study. *Lancet Microbe* **2**, e311–e319 (2021).
450. Bar-On, Y. M., Flamholz, A., Phillips, R. & Milo, R. SARS-CoV-2 (COVID-19) by the numbers. *Elife* **9**, 1–15 (2020).
451. Katsamba, P. S. *et al.* Kinetic analysis of a high-affinity antibody/antigen interaction performed by multiple Biacore users. *Anal. Biochem.* **352**, 208–221 (2006).
452. Firth, A. E. & Patrick, W. M. GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries. *Nucleic Acids Res.* **36**, 281–285 (2008).
453. Case, J. B., Bailey, A. L., Kim, A. S., Chen, R. E. & Diamond, M. S. Growth, detection, quantification, and inactivation of SARS-CoV-2. *Virology* **548**, 39–48 (2020).
454. Benda, A. *et al.* COVID-19 testing and diagnostics: A review of commercialized technologies for cost, convenience and quality of tests. *Sensors (Basel)* **21**, 6581 (2021).
455. Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A. & Kozlakidis, Z. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* **19**, 171–183 (2021).
456. Smithgall, M. C., Dowlatshahi, M., Spitalnik, S. L., Hod, E. A. & Rai, A. J. Types of assays for SARS-CoV-2 testing: A review. *Lab. Med.* **51**, e59–e65 (2020).
457. Hammerling, M. J., Warfel, K. F. & Jewett, M. C. Lyophilization of premixed COVID-19 diagnostic RT-qPCR reactions enables stable long-term storage at elevated temperature. *Biotechnol. J.* **16**, e2000572 (2021).
458. Pavelka, M. *et al.* The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia. *Science* **641**, eabf9648 (2021).

459. Mina, M. J., Peto, T. E., García-Fiñana, M., Semple, M. G. & Buchan, I. E. Clarifying the evidence on SARS-CoV-2 antigen rapid tests in public health responses to COVID-19. *Lancet* **397**, 1425–1427 (2021).
460. Lanser, L. *et al.* Evaluating the clinical utility and sensitivity of SARS-CoV-2 antigen testing in relation to RT-PCR Ct values. *Infection* **49**, 555–557 (2021).
461. Wölfel, R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).
462. Perchetti, G. A., Huang, M.-L., Mills, M. G., Jerome, K. R. & Greninger, A. L. Analytical sensitivity of the Abbott BinaxNOW COVID-19 Ag card. *J. Clin. Microbiol.* **59**, (2021).
463. Adamson, B., Sikka, R., Wyllie, A. L. & Premsrirut, P. Discordant SARS-CoV-2 PCR and rapid antigen test results when infectious: A December 2021 occupational case series. *bioRxiv* (2022) doi:10.1101/2022.01.04.22268770.
464. Frediani, J. K. *et al.* Multidisciplinary assessment of the Abbott BinaxNOW SARS-CoV-2 point-of-care antigen test in the context of emerging viral variants and self-administration. *Sci. Rep.* **11**, 14604 (2021).
465. Agarwal, D. K. *et al.* Highly sensitive and ultra-rapid antigen-based detection of SARS-CoV-2 using nanomechanical sensor platform. *Biosens. Bioelectron.* **195**, 113647 (2022).
466. Agarwal, D. K. *et al.* Detection of heart-type fatty acid-binding protein (h-FABP) using piezoresistive polymer microcantilevers functionalized by a dry method. *Appl. Nanosci.* **8**, 1031–1042 (2018).
467. Lavrik, N. V., Sepaniak, M. J. & Datskos, P. G. Cantilever transducers as a platform for chemical and biological sensors. *Rev. Sci. Instrum.* **75**, 2229–2253 (2004).
468. Agarwal, D. K., Kushagra, A., Ashwin, M., Shukla, A. S. & Palaparthi, V. Sensitive detection of cardiac troponin-I protein using fabricated piezoresistive microcantilevers by a novel method of asymmetric biofunctionalization. *Nanotechnology* **31**, 115503 (2020).
469. Fritz, J. *et al.* Translating biomolecular recognition into nanomechanics. *Science* **288**, 316–318 (2000).
470. Tamayo, J., Kosaka, P. M., Ruz, J. J., San Paulo, Á. & Calleja, M. Biosensors based on nanomechanical systems. *Chem. Soc. Rev.* **42**, 1287–1311 (2013).

471. Shekhawat, G. S. & Dravid, V. P. Nanomechanical sensors: Bent on detecting cancer. *Nature nanotechnology* vol. 8 77–78 (2013).
472. Mostafa, S. *et al.* Integrated MOSFET-embedded-cantilever-based biosensor characteristic for detection of anthrax simulant. *IEEE Electron Device Lett.* **32**, 408–410 (2011).
473. Miller, E. A., Baniya, S., Osorio, D., Al Maalouf, Y. J. & Sikes, H. D. Paper-based diagnostics in the antigen-depletion regime: High-density immobilization of rcSso7d-cellulose-binding domain fusion proteins for efficient target capture. *Biosens. Bioelectron.* **102**, 456–463 (2018).
474. Holstein, C. A. *et al.* Immobilizing affinity proteins to nitrocellulose: a toolbox for paper-based assay developers. *Anal. Bioanal. Chem.* **408**, 1335–1346 (2016).
475. Anderson, C. E. *et al.* Rapid diagnostic assay for intact influenza virus using a high affinity hemagglutinin binding protein. *Anal. Chem.* **89**, 6608–6615 (2017).
476. Jewett, M. C., Calhoun, K. A., Voloshin, A., Wu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* **4**, (2008).
477. Fung, B. *et al.* Direct comparison of SARS-CoV-2 analytical limits of detection across seven molecular assays. *J. Clin. Microbiol.* **58**, (2020).
478. Joung, J. *et al.* Detection of SARS-CoV-2 with SHERLOCK One-Pot Testing. *N. Engl. J. Med.* **382**, NEJMc2026172 (2020).
479. Broughton, J. P. *et al.* CRISPR-Cas12-based detection of SARS-CoV-2. *Nat. Biotechnol.* **38**, 870–874 (2020).
480. Fozouni, P. *et al.* Amplification-free detection of SARS-CoV-2 with CRISPR-Cas13a and mobile phone microscopy. *Cell* **184**, 323-333.e9 (2020).
481. Barak, N. *et al.* Lessons from applied large-scale pooling of 133,816 SARS-CoV-2 RT-PCR tests. *Sci. Transl. Med.* **13**, eabf2823 (2021).
482. Shekhawat, G., Tark, S.-H. & Dravid, V. P. MOSFET-Embedded microcantilevers for measuring deflection in biomolecular sensors. *Science* **311**, 1592–1595 (2006).
483. Jayanthi, V. S. P. K. S. A., Das, A. B. & Saxena, U. Recent advances in biosensor development for the detection of cancer biomarkers. *Biosens. Bioelectron.* **91**, 15–23 (2017).

484. Kosaka, P. M. *et al.* Detection of cancer biomarkers in serum using a hybrid mechanical and optoplasmonic nanosensor. *Nat. Nanotechnol.* **9**, 1047–1053 (2014).
485. Burton, D. R., Poignard, P., Stanfield, R. L. & Wilson, I. A. Broadly Neutralizing Antibodies Present New Prospects to Counter Highly Antigenically Diverse Viruses. *Science* **337**, 183–186 (2012).
486. Westendorf, K. *et al.* LY-CoV1404 (bebtelovimab) potently neutralizes SARS-CoV-2 variants. *bioRxiv* (2022) doi:10.1101/2021.04.30.442182.
487. Pegu, A., Hessel, A. J., Mascola, J. R. & Haigwood, N. L. Use of broadly neutralizing antibodies for HIV-1 prevention. *Immunol. Rev.* **275**, 296–312 (2017).
488. Corti, D. *et al.* Tackling influenza with broadly neutralizing antibodies. *Curr. Opin. Virol.* **24**, 60–69 (2017).
489. Piran, U. & Riordan, W. J. Dissociation rate constant of the biotin-streptavidin complex. *J. Immunol. Methods* **133**, 141–143 (1990).
490. Muñoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J. Mol. Biol.* **245**, 297–308 (1995).
491. Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* 2020.05.13.092619 (2020) doi:10.1038/s41586-020-2456-9.
492. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
493. Ueda, G. *et al.* Tailored design of protein nanoparticle scaffolds for multivalent presentation of viral glycoprotein antigens. *Elife* **9**, e57659 (2020).
494. Boyken, S. E. *et al.* De novo design of tunable, pH-driven conformational changes. *Science* **364**, 658–664 (2019).
495. Harbury, P. B., Kim, P. S. & Alber, T. Crystal structure of an isoleucine-zipper trimer. *Nature* **371**, 80–83 (1994).
496. Thomson, A. R. *et al.* Computational design of water-soluble α -helical barrels. *Science* **346**, 485–488 (2014).