

CORONAVIRUS

Structure-based design of prefusion-stabilized SARS-CoV-2 spikes

Ching-Lin Hsieh¹, Jory A. Goldsmith¹, Jeffrey M. Schaub¹, Andrea M. DiVenere², Hung-Che Kuo¹, Kamyab Javanmardi¹, Kevin C. Le², Daniel Wrapp¹, Alison G. Lee¹, Yutong Liu², Chia-Wei Chou¹, Patrick O. Byrne¹, Christy K. Hjorth¹, Nicole V. Johnson¹, John Ludes-Meyers¹, Annalee W. Nguyen², Juyeon Park¹, Nianshuang Wang¹, Dzifa Amengor¹, Jason J. Lavinder^{1,2}, Gregory C. Ippolito^{1,3}, Jennifer A. Maynard^{2*}, Ilya J. Finkelstein^{1,4,*}, Jason S. McLellan^{1*}

The coronavirus disease 2019 (COVID-19) pandemic has led to accelerated efforts to develop therapeutics and vaccines. A key target of these efforts is the spike (S) protein, which is metastable and difficult to produce recombinantly. We characterized 100 structure-guided spike designs and identified 26 individual substitutions that increased protein yields and stability. Testing combinations of beneficial substitutions resulted in the identification of HexaPro, a variant with six beneficial proline substitutions exhibiting higher expression than its parental construct (by a factor of 10) as well as the ability to withstand heat stress, storage at room temperature, and three freeze-thaw cycles. A cryo-electron microscopy structure of HexaPro at a resolution of 3.2 angstroms confirmed that it retains the prefusion spike conformation. High-yield production of a stabilized prefusion spike protein will accelerate the development of vaccines and serological diagnostics for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel beta-coronavirus that emerged in Wuhan, China in December 2019 and is the causative agent of the coronavirus disease 2019 (COVID-19) pandemic (1, 2). Effective vaccines, therapeutic antibodies, and small-molecule inhibitors are urgently needed, and the development of these interventions is proceeding rapidly. Coronavirus virions are decorated with a spike (S) glycoprotein that binds to host cell receptors and mediates cell entry via fusion of the host and viral membranes (3). Binding of the SARS-CoV-2 spike to the angiotensin-converting enzyme 2 (ACE2) receptor (4–6) triggers a large conformational rearrangement of the spike from a metastable prefusion conformation to a highly stable postfusion conformation, facilitating membrane fusion (7, 8). Attachment and entry are essential for the viral life cycle, making the S protein a primary target of neutralizing antibodies and a critical vaccine antigen (9, 10).

Prefusion stabilization tends to increase the recombinant expression of viral fusion glycoproteins, possibly by preventing triggering or misfolding that results from a tendency to adopt the more stable postfusion structure. Prefusion-stabilized viral glycoproteins are also superior immunogens to their wild-type counterparts (11–13). Structure-based design of prefusion-stabilized MERS-CoV and SARS-CoV spike ectodomains resulted in homoge-

neous preparations of prefusion spikes and greatly increased yields (11). These variants (S-2P) contained two consecutive proline substitutions in the S2 subunit in a turn between the central helix and heptad repeat 1 (HR1) that must transition to a single, elongated α helix in the postfusion conformation. These S-2P spikes have been used to determine high-resolution

structures by cryo-electron microscopy (cryo-EM) (14–17), including for SARS-CoV-2 (18, 19), and have accelerated the development of vaccine candidates. However, even with these substitutions, the SARS-CoV-2 S-2P ectodomain is unstable and difficult to produce reliably in mammalian cells, hampering biochemical research and development of subunit vaccines.

To generate a prefusion-stabilized SARS-CoV-2 spike protein that expresses at higher levels and is more stable than our original S-2P construct (18), we analyzed the SARS-CoV-2 S-2P cryo-EM structure (PDB ID 6VSB) and designed substitutions based on knowledge of class I viral fusion protein function and general protein stability principles. These strategies included the introduction of disulfide bonds to prevent conformational changes during the prefusion-postfusion transition, salt bridges to neutralize charge imbalances, hydrophobic residues to fill internal cavities, and prolines to cap helices or stabilize loops in the prefusion state. We cloned 100 single S-2P variants and characterized their relative expression levels (table S1), and for those that expressed well, we characterized their monodispersity, thermostability, and quaternary structure. Given that the S2 subunit undergoes large-scale refolding during the prefusion-postfusion transition, we exclusively focused our efforts on stabilizing S2. Substitutions of each category were identified that increased expression while

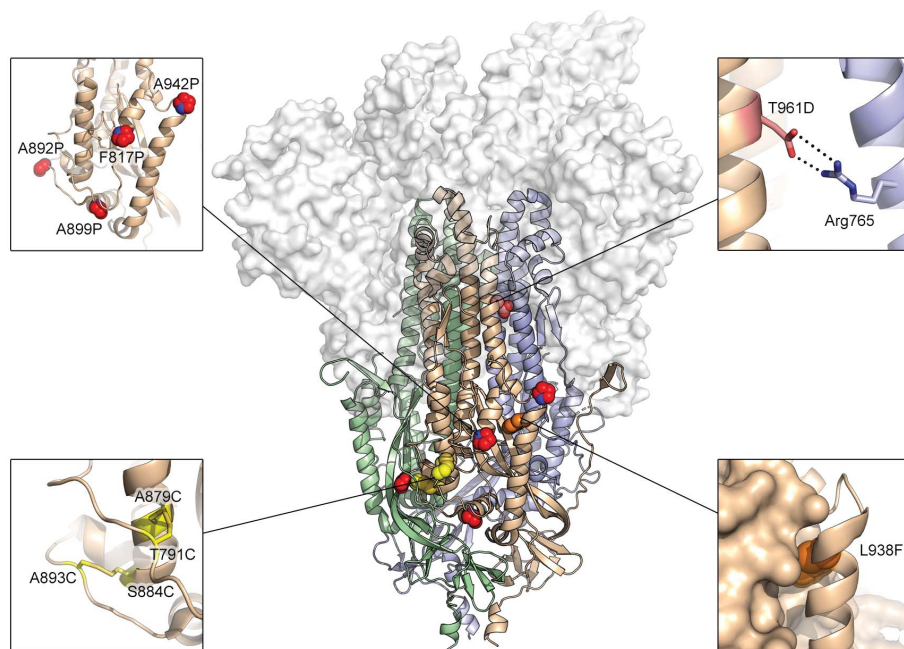


Fig. 1. Exemplary substitutions for SARS-CoV-2 spike stabilization. The main image is a side view of the trimeric SARS-CoV-2 spike ectodomain in a prefusion conformation (PDB ID 6VSB). The S1 domains are shown as a transparent molecular surface. The S2 domain for each protomer is shown as a ribbon diagram. Each inset corresponds to one of four types of spike modification: proline, salt bridge, disulfide, and cavity filling. Side chains in each inset are shown as red spheres (proline), yellow sticks (disulfide), red and blue sticks (salt bridge), and orange spheres (cavity filling).

¹Department of Molecular Biosciences, University of Texas, Austin, TX 78712, USA. ²Department of Chemical Engineering, University of Texas, Austin, TX 78712, USA. ³Department of Oncology, Dell Medical School, University of Texas, Austin, TX 78712, USA. ⁴Center for Systems and Synthetic Biology, University of Texas, Austin, TX 78712, USA.

*Corresponding author. Email: maynard@che.utexas.edu (J.A.M.); ilya@finkelsteinlab.org (I.J.F.); jmclellan@austin.utexas.edu (J.S.M.)

maintaining the prefusion conformation (Fig. 1 and Fig. 2A). Overall, 26 of the 100 single-substitution variants had higher expression than S-2P (table S1).

One common strategy to stabilize class I fusion proteins is to covalently link a region that undergoes a conformational change to a region that does not via a disulfide bond. For instance, the Q965C/S1003C (Gln⁹⁶⁵ → Cys, Ser¹⁰⁰³ → Cys) substitution aims to link HR1 to the central helix, whereas G799C/A924C (Gly⁷⁹⁹ → Cys, Ala⁹²⁴ → Cys) aims to link HR1 to the upstream helix. Relative to S-2P, these two variants boosted protein expression by factors of 3.8 and 1.3, respectively (Fig. 2B). However, the

size-exclusion chromatography (SEC) traces of both variants showed a leftward shift relative to S-2P, indicating that the proteins were running larger than expected, which agreed well with negative-stain electron microscopy (nsEM) results that showed partially misfolded spike particles (fig. S1). Although introduction of disulfide bonds has been successful in the case of HIV-1 Env (SOSIP) and RSV F (DS-Cav1) (22, 20), it generally had detrimental effects for SARS-CoV-2 S, but there were a few exceptions. The S884C/A893C (Ser⁸⁸⁴ → Cys, Ala⁸⁹³ → Cys) and T791C/A879C (Thr⁷⁹¹ → Cys, Ala⁸⁷⁹ → Cys) variants eluted on SEC at a volume similar to S-2P and were well-folded trimeric particles by

nsEM (Fig. 2E). These variants link the same α helix to two different flexible loops that pack against a neighboring protomer (Fig. 1). Notably, the expression of S884C/A893C was double that of S-2P, with slightly increased thermostability (Fig. 2, F and G).

Introducing a salt bridge at the HIV-1 gp120-gp41 interface has been shown to boost expression and enhance the binding of trimer-specific antibodies (21). On the basis of a similar principle, the T961D (Thr⁹⁶¹ → Asp) and G769E (Gly⁷⁶⁹ → Glu) substitutions were introduced to form interprotomeric electrostatic interactions with Arg⁷⁶⁵ and Arg¹⁰¹⁴, respectively (Fig. 1). Both variants increased expression

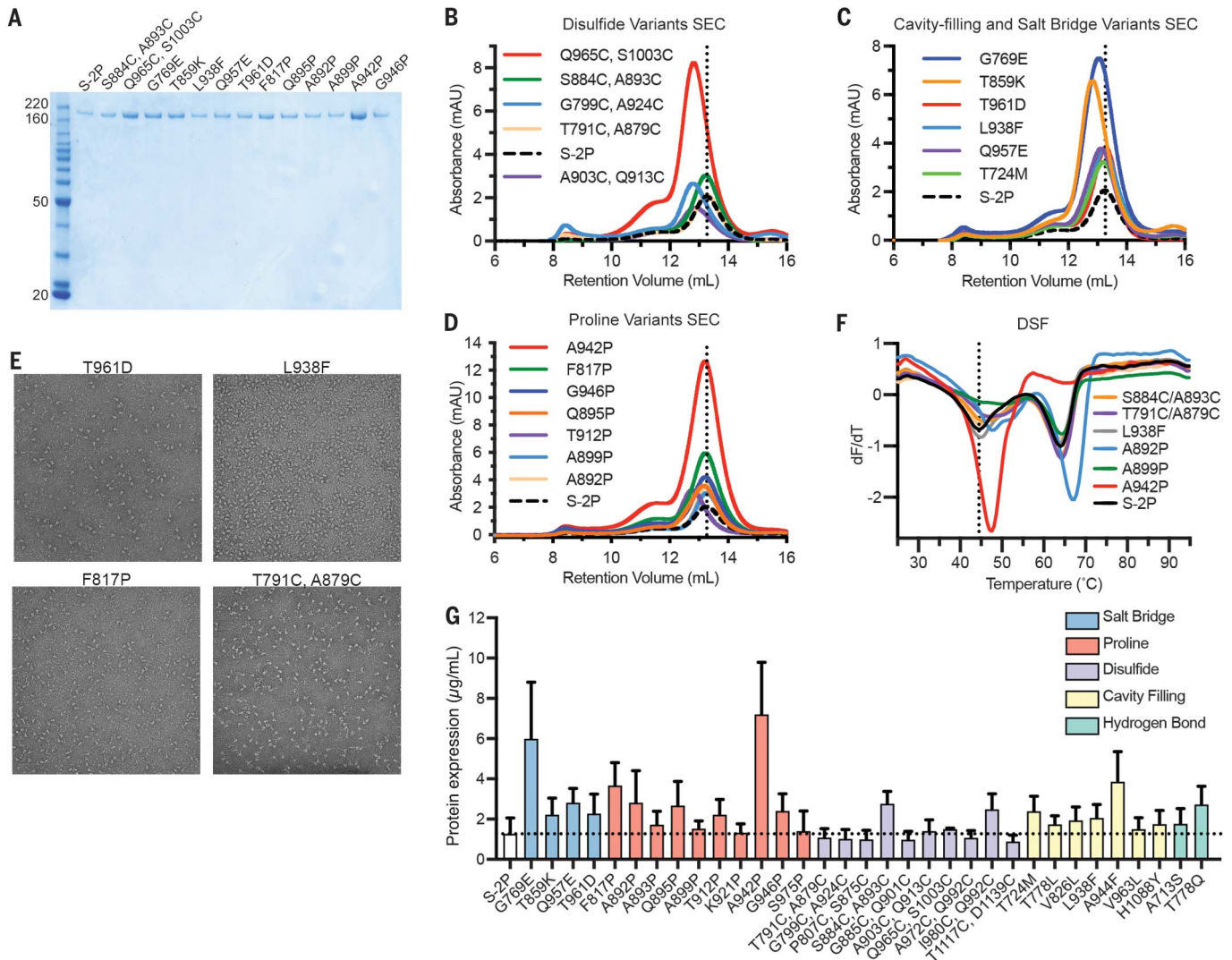


Fig. 2. Characterization of single-substitution spike variants. (A) SDS-polyacrylamide gel electrophoresis (PAGE) of SARS-CoV-2 S-2P and single-substitution spike variants. Molecular weight standards are indicated at the left in kDa. (B to D) Size-exclusion chromatography traces of purified spike variants, grouped by type: disulfide variants (B), cavity filling and salt bridge variants (C), and proline variants (D). A vertical dotted line indicates the characteristic peak retention volume for S-2P. (E) Representative negative-stain electron

micrographs for four variants. (F) Differential scanning fluorimetry analysis of spike variant thermostability. The vertical dotted line indicates the first apparent melting temperature for S-2P. (G) Expression levels of individual variants determined by quantitative biolayer interferometry. Variants are colored by type. The horizontal dotted line indicates the calculated concentration of S-2P, which was used as a control for comparison. Data are means \pm SD of three biological replicates.

and resembled well-folded trimeric spikes (Fig. 2, C and E, fig. S2, and table S1). In addition to salt bridges, filling loosely packed hydrophobic cores that allow the protein to refold can help to stabilize the prefusion state, as shown by previous cavity-filling substitutions in RSV F and HIV-1 Env (12, 20, 22). Here, the L938F (Leu⁹³⁸ → Phe) substitution was de-

signed to fill a cavity formed in part by HRI, the fusion peptide, and a β hairpin (Fig. 1). This substitution resulted in a factor of 2 increase in expression (Fig. 2C) that was additive in combination with disulfide or proline substitutions (table S2).

Previous successes using proline substitutions inspired us to investigate 14 individual

variants wherein a proline was substituted into flexible loops or the N termini of helices in the fusion peptide, HRI, and the region connecting them (CR) (Fig. 2, D and G, and table S1). As expected, multiple proline variants boosted the protein expression and increased the thermostability (Fig. 2, D, F, and G). Two of the most successful substitutions, F817P

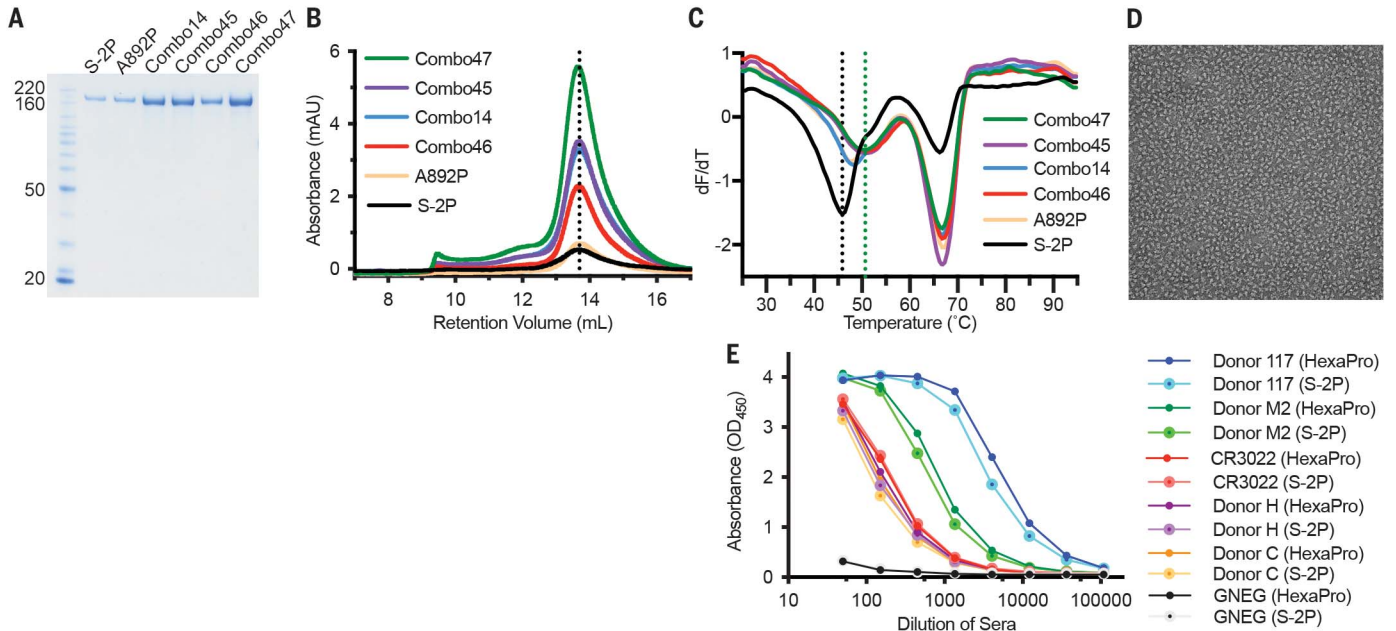


Fig. 3. Characterization of multi-substitution spike variants. (A) SDS-PAGE of SARS-CoV-2 Combo variants. Molecular weight standards are indicated at the left in kDa. (B) SEC traces for S-2P, A892P, and four Combo variants. The vertical dotted line indicates the peak retention volume for S-2P. (C) Differential scanning fluorimetry analysis of Combo variant thermostability. The black vertical dotted line indicates

the first apparent melting temperature for S-2P; the green vertical dotted line indicates the first apparent melting temperature for Combo47 (HexaPro). (D) Negative-stain electron micrograph of purified Combo47 (HexaPro). (E) Binding of S-2P or HexaPro to convalescent human sera, mAb CR3022, and negative control serum (GNEG), measured by enzyme-linked immunosorbent assay.

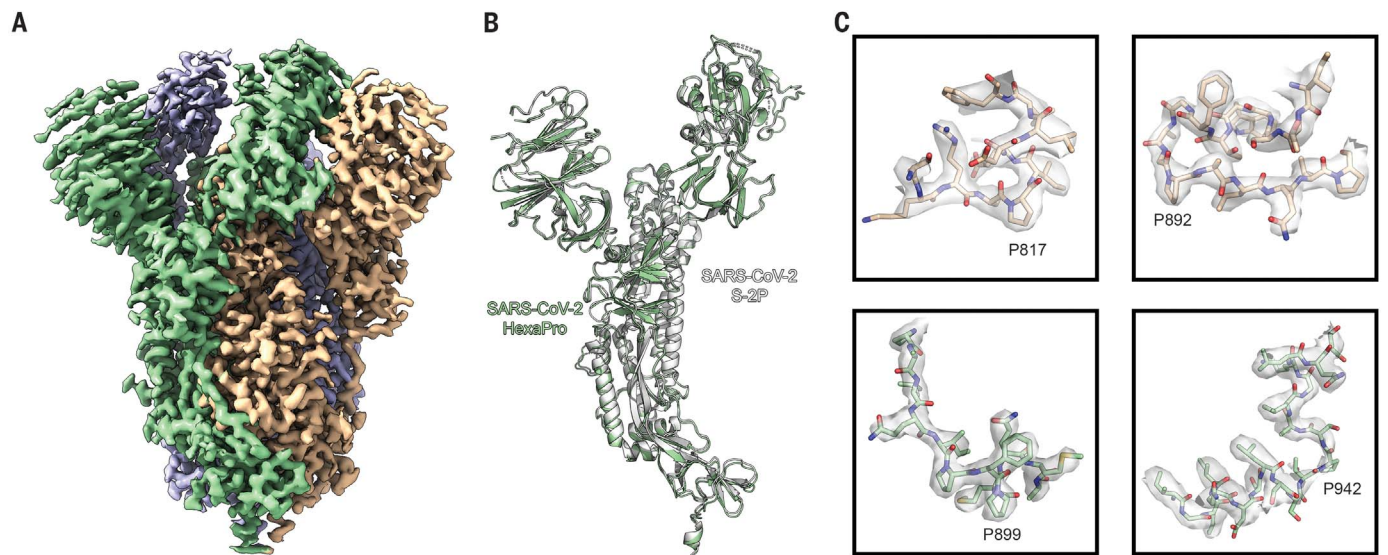


Fig. 4. High-resolution cryo-EM structure of HexaPro. (A) EM density map of trimeric HexaPro. Each protomer is shown in a different color; the protomer depicted in wheat adopts the RBD-up conformation. (B) Alignment of an RBD-down protomer from HexaPro (green ribbon) with an RBD-down protomer

from S-2P (white ribbon; PDB ID 6VSB). (C) Zoomed-in view of the four proline substitutions unique to HexaPro. The EM density map is shown as a transparent surface and individual atoms are shown as sticks. Blue, nitrogen atoms; red, oxygen atoms.

(Phe⁸¹⁷ → Pro) and A942P (Ala⁹⁴² → Pro), exhibited increases in protein yield relative to S-2P by factors of 2.8 and 6.0, respectively. The A942P substitution further increased the melting temperature (T_m) by ~3°C, and both variants appeared as well-folded trimers by nsEM (Fig. 2E and fig. S2). This result is reminiscent of previous successful applications of proline substitutions to class I fusion proteins including HIV-1 Env, influenza HA, RSV F, hMPV F, MERS-CoV S, Lassa GPC, and Ebola GP (11, 12, 22–26).

We next generated combination (“Combo”) variants that combined the best-performing substitutions from our initial screen. The expression of the Combo variants containing two disulfide bonds was generally half that of the single-disulfide variants, which suggests that they interfered with each other (table S2). Adding one disulfide (S884C/A893C) to a single proline variant (F817P) also reduced the expression level, although the quaternary structure of the spikes was well maintained (table S2, Combo40). The beneficial effect of a disulfide bond was most prominent when combined with L938F, a cavity-filling variant. Combo23 (S884C/A893C, L938F) had higher protein yields than either of its parental variants, but the T_m of Combo23 did not increase relative to S884C/A893C (fig. S3). In addition, mixing one cavity-filling substitution with one proline substitution (Combo20) increased the expression relative to L938F alone (table S2).

Combining multiple proline substitutions resulted in the most substantial increases in expression and stability (Fig. 3A). Combo14, containing A892P (Ala⁸⁹² → Pro) and A942P, had a factor of 6.2 increase in protein yield relative to A892P alone (Fig. 3B). Adding a third proline, A899P (Ala⁸⁹⁹ → Pro) (Combo45), increased thermostability ($T_m = +1.2^\circ\text{C}$) but did not further increase expression (Fig. 3C). Combo46 (A892P, A899P, F817P) had a factor of 3.4 increase in protein yield and a 3.3°C rise in T_m as compared to A892P. The most promising variant, Combo47, renamed HexaPro, contains all four beneficial proline substitutions (F817P, A892P, A899P, A942P) as well as the two proline substitutions in S-2P. HexaPro displayed higher expression than S-2P by a factor of 9.8, had a ~5°C increase in T_m , and retained the trimeric prefusion conformation (Fig. 3D). We focused on this construct for additional characterization.

To assess the viability of HexaPro as a potential vaccine antigen or diagnostic reagent, we comprehensively examined large-scale production in FreeStyle 293-F cells, feasibility of protein expression in ExpiCHO cells, epitope integrity, and protein stability. We were able to generate ~21 mg of HexaPro from 2 liters of FreeStyle 293-F cells, or 10.5 mg/liter, which represents an improvement over S-2P by more than an order of magnitude (18). Large-scale

HexaPro preparations retained a monodisperse SEC peak corresponding to the molecular weight of a glycosylated trimer (fig. S4A) and were indistinguishable from S-2P by nsEM (fig. S4B). Industrial production of recombinant proteins typically relies on CHO cells rather than HEK293 cells. We thus investigated HexaPro expression in ExpiCHO cells via transient transfection. ExpiCHO cells produced 1.3 mg of well-folded protein per 40 ml of culture, or 32.5 mg/liter (fig. S4, C and D). In addition, the binding kinetics of HexaPro to the human ACE2 receptor were comparable to those of S-2P (fig. S4, E and F), with affinities of 13.3 nM and 11.3 nM, respectively. HexaPro remained folded in the prefusion conformation after three cycles of freeze-thaw, 2 days of incubation at room temperature, or 30 min at 55°C (fig. S4, G and H). In contrast, S-2P showed signs of aggregation after three cycles of freeze-thaw and began unfolding after 30 min at 50°C. HexaPro reacted to human convalescent sera and to receptor-binding domain (RBD)-specific monoclonal antibody (mAb) CR3022 (27) similarly to S-2P, which suggests that the antigenicity of HexaPro is well preserved (Fig. 3E). Collectively, these data indicate that HexaPro is a promising candidate for SARS-CoV-2 vaccine and diagnostic development.

To confirm that the stabilizing substitutions did not lead to any unintended conformational changes, we determined the cryo-EM structure of SARS-CoV-2 S HexaPro. From a single dataset, we were able to obtain high-resolution 3D reconstructions for two distinct conformations of S: one with a single RBD in the up conformation and the other with two RBDs in the up conformation. This two-RBD-up conformation was not observed during previous structural characterization of SARS-CoV-2 S-2P (18, 19). Although it is tempting to speculate that the enhanced stability of S2 in HexaPro allowed us to observe this less stable intermediate, validating this hypothesis will require further investigation. Roughly one-third (30.6%) of the particles were in the two-RBD-up conformation, leading to a 3.20 Å reconstruction. The remaining particles were captured in the one-RBD-up conformation, although some flexibility in the position of the receptor-accessible RBD prompted us to remove a subset of one-RBD-up particles that lacked clear density for this domain, resulting in a final set of 85,675 particles that led to a 3.21 Å reconstruction (Fig. 4A and figs. S5 and S6). Comparison of our one-RBD-up HexaPro structure with the previously determined 3.46 Å S-2P structure revealed a RMSD of 1.2 Å over 436 C α atoms in S2 (Fig. 4B). The relatively high resolution of this reconstruction allowed us to confirm that the stabilizing proline substitutions did not distort the S2 subunit conformation (Fig. 4C).

The high yield and enhanced stability of HexaPro should enable industrial production

of subunit vaccines and could also improve DNA- or mRNA-based vaccines by producing more antigen per nucleic acid molecule, thus improving efficacy at the same dose or maintaining efficacy at lower doses. It is our hope that this work will accelerate the production of prefusion spikes to mitigate the public health emergency and has broad implications for next-generation coronavirus vaccine design.

REFERENCES AND NOTES

1. J. F. W. Chan et al., *Lancet* **395**, 514–523 (2020).
2. C. Huang et al., *Lancet* **395**, 497–506 (2020).
3. F. Li, *Annu. Rev. Virol.* **3**, 237–261 (2016).
4. M. Hoffmann et al., *Cell* **181**, 271–280.e8 (2020).
5. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, *J. Virol.* **94**, e00127 (2020).
6. P. Zhou et al., *Nature* **579**, 270–273 (2020).
7. B. J. Bosch, R. van der Zee, C. A. de Haan, J. M. Rottier, *J. Virol.* **77**, 8801–8811 (2003).
8. A. C. Walls et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11157–11162 (2017).
9. U. J. Buchholz et al., *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9804–9809 (2004).
10. H. Hofmann et al., *J. Virol.* **78**, 6134–6142 (2004).
11. J. Pallesen et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7348–E7357 (2017).
12. R. W. Sanders et al., *PLOS Pathog.* **9**, e1003618 (2013).
13. M. C. Crank et al., *Science* **365**, 505–509 (2019).
14. Y. J. Park et al., *Nat. Struct. Mol. Biol.* **26**, 1151–1157 (2019).
15. Z. Li et al., *eLife* **8**, e51230 (2019).
16. N. Wang et al., *Cell Rep.* **28**, 3395–3405.e6 (2019).
17. A. C. Walls et al., *Cell* **176**, 1026–1039.e15 (2019).
18. D. Wrapp et al., *Science* **367**, 1260–1263 (2020).
19. A. C. Walls et al., *Cell* **181**, 281–292.e6 (2020).
20. J. S. McLellan et al., *Science* **342**, 592–598 (2013).
21. L. Rutten et al., *Cell Rep.* **23**, 584–595 (2018).
22. A. Krarup et al., *Nat. Commun.* **6**, 8143 (2015).
23. L. Rutten et al., *Cell Rep.* **30**, 4540–4550.e3 (2020).
24. M. B. Battles et al., *Nat. Commun.* **8**, 1528 (2017).
25. H. Qiao et al., *J. Cell Biol.* **141**, 1335–1347 (1998).
26. K. M. Hastie et al., *Science* **356**, 923–928 (2017).
27. M. Yuan et al., *Science* **368**, 630–633 (2020).

ACKNOWLEDGMENTS

We thank members of the Maynard, Finkelstein, and McLellan Laboratories for providing helpful comments on the manuscript; T. Edwards and U. Baxa for cryo-EM data collection; and E. Fitch for providing helpful data analysis. **Funding:** Supported by NIH grants or contracts R01-A1127521 (J.S.M.), GM120554 and GM124141 (I.J.F.), A1122753 (J.A.M.), and 75N93019C00050 (J.J.L. and G.C.I.); the Bill & Melinda Gates Foundation (J.A.M., I.J.F., and J.S.M.); Welch Foundation grants F-1808 (I.J.F.) and F-1767 (J.A.M.); NSF grant 1453358 (I.J.F.); and the National Cancer Institute’s National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E. I.J.F. is a Cancer Prevention and Research Institute of Texas (CPRIT) Scholar in Cancer Research. The Sauer Structural Biology Laboratory is supported by the University of Texas College of Natural Sciences and by award RR160023 from CPRIT. **Author contributions:** Conceptualization, C.-L.H. and J.S.M.; investigation and visualization, C.-L.H., J.A.G., C.-W.C., A.M.D., K.J., H.-C.K., K.C.L., A.G.L., Y.L., J.M.S., D.W., P.O.B., C.K.H., N.V.J., J.L.-M., A.W.N., J.P., G.I., J.J.L., and D.A.; writing—original draft, C.-L.H., J.A.G., D.W., P.O.B., C.K.H., N.V.J., and J.S.M.; writing—reviewing & editing, C.-L.H., J.A.G., D.W., P.O.B., N.W., C.K.H., N.V.J., J.A.M., I.J.F., and J.S.M.; supervision, J.A.M., I.J.F., and J.S.M. **Competing interests:** N.W. and J.S.M. are inventors on U.S. patent application no. 62/412,703 (“Prefusion Coronavirus Spike Proteins and Their Use”). D.W., N.W., and J.S.M. are inventors on U.S. patent application no. 62/972,886 (“2019-nCoV Vaccine”). C.-L.H., J.A.G., J.M.S., C.-W.C., A.M.D., K.J., H.-C.K., D.W., P.O.B., C.K.H., N.V.J., N.W., J.A.M., I.J.F., and J.S.M. are inventors on U.S. patent application no. 63/032,502 [“Engineered Coronavirus Spike (S) Protein and Methods of Use Thereof”]. **Data and materials availability:** Atomic coordinates and cryo-EM maps of the reported structures have been deposited

in the Protein Data Bank under accession code 6XKL and in the Electron Microscopy Data Bank under accession codes EMD-22221 and EMD-22222. HexaPro plasmid is available from Addgene (ID: 154754) or from J.S.M. under a material transfer agreement with the University of Texas at Austin. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6510/1501/suppl/DC1
Materials and Methods

Figs. S1 to S6
Tables S1 to S3
References (28–33)

[View/request a protocol for this paper from Bio-protocol.](#)

30 May 2020; accepted 13 July 2020
Published online 23 July 2020
10.1126/science.abd0826

Structure-based design of prefusion-stabilized SARS-CoV-2 spikes

Ching-Lin HsiehJory A. GoldsmithJeffrey M. SchaubAndrea M. DiVenereHung-Che KuoKamyab JavanmardiKevin C. LeDaniel WrappAlison G. LeeYutong LiuChia-Wei ChouPatrick O. ByrneChristy K. HjorthNicole V. JohnsonJohn Ludes-MeyersAnnalee W. NguyenJuyeon ParkNianshuang WangDzifa AmengorJason J. LavinderGregory C. IppolitoJennifer A. MaynardIlya J. FinkelsteinJason S. McLellan

Science, 369 (6510),

Stabilizing the prefusion SARS-CoV-2 spike

The development of therapeutic antibodies and vaccines against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is focused on the spike (S) protein that decorates the viral surface. A version of the spike ectodomain that includes two proline substitutions (S-2P) and stabilizes the prefusion conformation has been used to determine high-resolution structures. However, even S-2P is unstable and difficult to produce in mammalian cells. Hsieh *et al.* characterized many individual and combined structure-guided substitutions and identified a variant, named HexaPro, that retains the prefusion conformation but shows higher expression than S-2P and can also withstand heating and freezing. This version of the protein is likely to be useful in the development of vaccines and diagnostics.

Science, this issue p. 1501

View the article online

<https://www.science.org/doi/10.1126/science.abd0826>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of think article is subject to the [Terms of service](#)

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works