










# Stabilizing the closed SARS-CoV-2 spike trimer

Jarek Juraszek <sup>1,3</sup>, Lucy Rutten <sup>1,3</sup>, Sven Blokland <sup>1</sup>, Pascale Bouchier <sup>1</sup>, Richard Voorzaat <sup>1</sup>, Tina Ritschel <sup>1</sup>, Mark J. G. Bakkers <sup>1</sup>, Ludovic L. R. Renault <sup>2</sup> & Johannes P. M. Langedijk <sup>1</sup>✉

The trimeric spike (S) protein of SARS-CoV-2 is the primary focus of most vaccine design and development efforts. Due to intrinsic instability typical of class I fusion proteins, S tends to prematurely refold to the post-fusion conformation, compromising immunogenic properties and prefusion trimer yields. To support ongoing vaccine development efforts, we report the structure-based design of soluble S trimers with increased yields and stabilities, based on introduction of single point mutations and disulfide-bridges. We identify regions critical for stability: the heptad repeat region 1, the SD1 domain and position 614 in SD2. We combine a minimal selection of mostly interprotomeric mutations to create a stable S-closed variant with a 6.4-fold higher expression than the parental construct while no longer containing a heterologous trimerization domain. The cryo-EM structure reveals a correctly folded, predominantly closed pre-fusion conformation. Highly stable and well producing S protein and the increased understanding of S protein structure will support vaccine development and serological diagnostics.

<sup>1</sup>Janssen Vaccines & Prevention B.V., Archimedesweg 4-6, Leiden, The Netherlands. <sup>2</sup>NeCEN, Leiden University, Einsteinweg 55, Leiden, The Netherlands. <sup>3</sup>These authors contributed equally: Jarek Juraszek, Lucy Rutten. ✉email: [hlangedi@its.jnj.com](mailto:hlangedi@its.jnj.com)

**D**evelopment of effective preventative interventions against the SARS-CoV-2 virus that causes the ongoing COVID-19 pandemic<sup>1,2</sup> is urgently needed. The viral surface spike (S) protein is a key target for prophylactic measures as it is critical in the viral life cycle and the primary target of neutralizing antibodies<sup>3–6</sup>. S is a large, trimeric glycoprotein that mediates both binding to host cell receptors and fusion of virus and host cell membranes, through its S1 and S2 subunits, respectively<sup>7–9</sup>. The S1 subunit comprises two distinct domains: an N-terminal domain (NTD) and a host cell receptor-binding domain (RBD). For infection, S requires proteolytic cleavage by a furin-like protease between the S1 and S2 subunits (S1/S2), and by TMPRSS2 at a conserved site directly preceding the fusion peptide (S2')<sup>10,11</sup>. In the prefusion state, the S-protein's RBD domains alternate between open ('up') and closed ('down') conformations. The receptor-binding site, which can bind to human angiotensin-converting enzyme 2 (ACE2), is transiently exposed in the 'up' conformation. Like other class I fusion proteins, the SARS-CoV-2 S protein is intrinsically metastable as a consequence of its ability to undergo extensive conformational changes that are required to drive fusion.

The prefusion conformation of S as present on the infectious particle contains the epitopes for neutralizing antibodies and thus holds most promise as a vaccine immunogen<sup>3–6,12,13</sup>. Prefusion stabilization typically increases the recombinant expression of viral glycoproteins, which facilitates the production of protein for (subunit) vaccines and improves the immune response elicited by recombinant protein and viral vector vaccines<sup>14</sup>. In recent years, efforts have been made to stabilize various class I fusion proteins through structure-based design (for a review see ref. 14). A particularly successful approach to enhance prefusion stability was shown to be the stabilization of the so-called hinge loop preceding the central helix (CH), which has been applied to a range of viral fusion glycoproteins<sup>15–20</sup>. Stabilization of the hinge loop of the S proteins of SARS-CoV and MERS-CoV has been achieved by mutation of two consecutive residues in the S2 subunit between the central helix (CH) and heptad repeat 1 (HR1)<sup>21,22</sup> to proline and this approach (2P) has successfully been applied to the SARS-CoV-2 S protein<sup>9</sup>. However, the SARS-CoV-2 S protein carrying these substitutions and additional furin site mutations (S-2P) remains unstable and strategies to further improve its stability have recently been described<sup>12,23–25</sup>. Comparison of the structure of SARS-CoV-2 S-2P<sup>9,26</sup> with that of native SARS-CoV-2 S<sup>27,28</sup> shows that the former adopts a more open conformation with one or more of the RBDs in the 'up' conformation. Although neutralizing antibodies have been mapped to the RBD in the up as well as down state, the antibodies that bind the conserved epitopes on the RBD in the down state were described to have the highest neutralizing potency<sup>4,6,13</sup>. Therefore, stabilizing S in its closed (3 RBDs down) state and arresting the first step in the conformational change may result in an improved vaccine immunogen.

In this work, using structure-based design, we find stabilizing mutations in both the S1 and S2 subdomains. Combining several of the mutations results in a highly stable S trimer, S-closed, with increased expression that remains stable in the absence of a heterologous trimerization domain that is typically required in soluble S designs<sup>9,21,26,29</sup>. Assessment of its antigenicity and high-resolution EM confirm that this trimer adopts a closed conformation.

## Results

To stabilize the SARS-CoV-2 S protein in the closed pre-fusion state, we took a rational approach based on the structure of S<sup>9</sup>. We searched for cavity-filling substitutions, buried charges, and

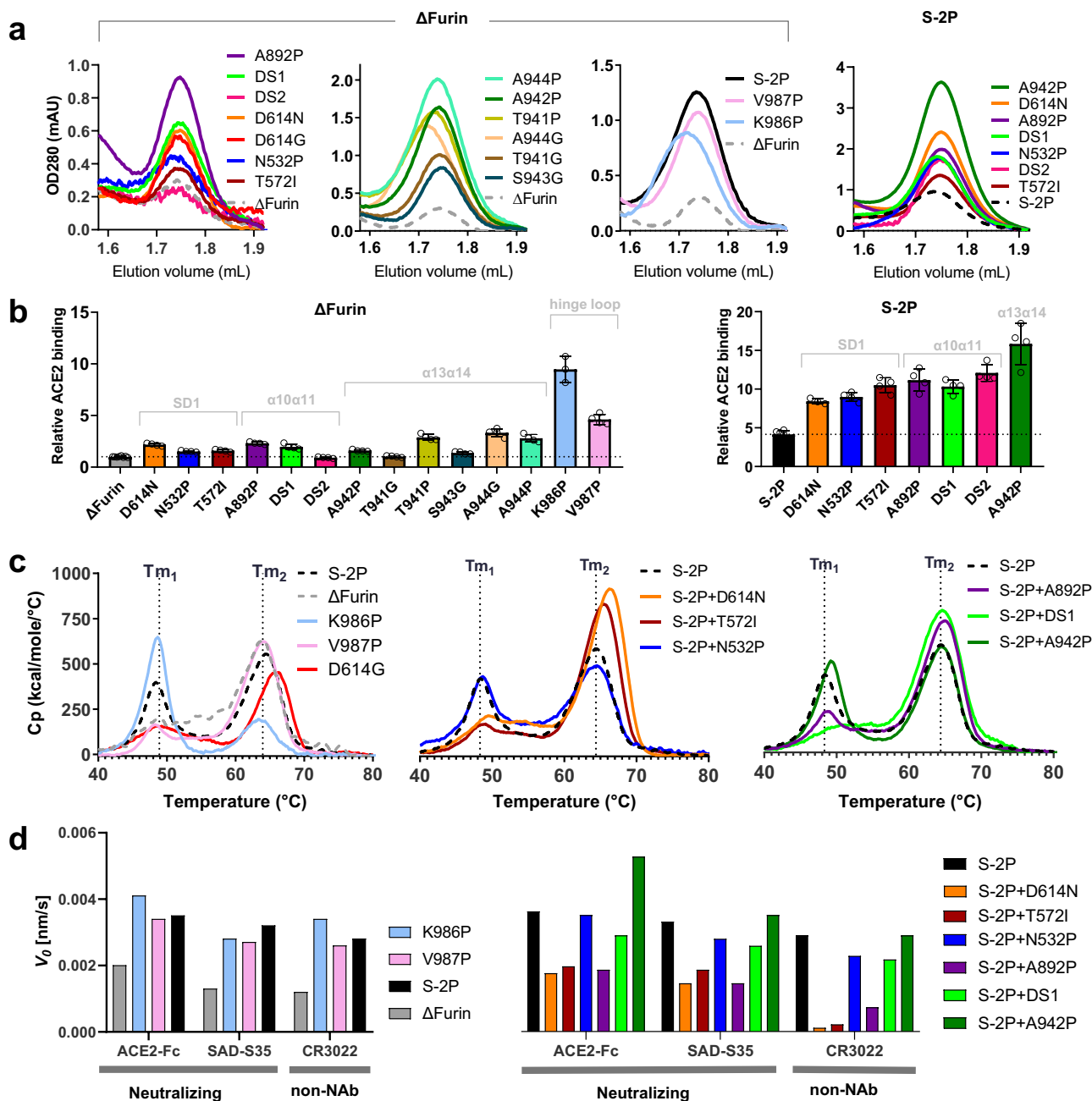
possibilities for forming disulfides. We expanded our search for proline and glycine mutations, beyond the previously described hinge loop in the S-2P variant. Mutations were identified computationally with Rosetta's mutant design<sup>30</sup> and Bioluminate cysteine bridge scanning<sup>31</sup> followed by visual inspection of molecular interactions. Selected mutations are presented in Supplementary Fig. 1 and fall in two structural categories—SD1 head mutations N532P, T572I, D614N, and S2 loop mutations A942P, T941G, T941P, S943G, A944P, A944G (loop  $\alpha$ 13 $\alpha$ 14), and A892P (loop  $\alpha$ 10 $\alpha$ 11). Disulfides F888C + G880C (DS1) and S884C + A893C (DS2) were identified in loop  $\alpha$ 10 $\alpha$ 11.

S ectodomain variants with mutations according to Supplementary Fig. 1 were expressed as single chain by mutation of the furin site and addition of a C-terminal foldon (Fd) trimerization domain without ( $\Delta$ Furin) or with the two previously described stabilizing prolines in the hinge loop (S-2P)<sup>9</sup>. Supernatants of Expi293F cells transfected with plasmids encoding the S variants were tested for trimer content (Fig. 1a) and for RBD exposure of the S protein by ACE2 binding (Fig. 1b). All mutations significantly increased trimer yields and ACE2 binding of the S protein for both the  $\Delta$ Furin and S-2P variants. A strong effect was observed with T941P, A942P, and A944P. A942P showed an ~11-fold increase in expression for  $\Delta$ Furin, and ~3-fold for S-2P (Fig. 1a). For T941P, A944G, and K986P shorter retention times were observed, likely indicating an opening of the trimers. T941P and A944G showed the highest ACE2 binding among the  $\alpha$ 13 $\alpha$ 14 loop mutations and K986P resulted in ~10-fold higher ACE2 binding, whereas the trimer yield was only ~3-fold higher than  $\Delta$ Furin. This indicates that the RBD domains are more exposed.

The stability of the single point mutants was further characterized by purified proteins using differential scanning calorimetry (DSC). First, we tested the contribution to the stability of the individual proline mutations of the S-2P variant (Fig. 1c left panel, Supplementary Table 1). All curves showed two melting events ( $T_{m50}$ 's), albeit with different ratios.  $\Delta$ Furin and additional V987P show a major  $T_{m50}$  at 64 °C and a minor  $T_{m50}$  around 49 °C. This is inverted for the K986P mutant in which the lower temperature transition is dominant. The combination of both prolines reduced the lower and increased the higher  $T_{m50}$ . Among the S1 mutants (Fig. 1c central panel and Supplementary Table 1) D614N and the natural variant D614G diminished the peak height of the first  $T_{m50}$  and increased the second by almost 2 °C in their respective backgrounds, which was subsequently confirmed with differential scanning fluorimetry (DSF) for D614N (Supplementary Fig. 2). A similar effect was observed for T572I and the loop-stabilizing mutations A892P and DS1, albeit to a lesser extent (Fig. 1c right panel and Supplementary Table 1). A942P, which increased the trimer yields, hardly affected the thermal stability in DSC (Fig. 1c right panel) or DSF (Supplementary Fig. 2).

RBD exposure was characterized by binding of ACE2, neutralizing antibody SAD-S35 and non-neutralizing antibody CR3022 that competes with ACE2. ACE2 and SAD-S35 can only bind RBD in the 'up' configuration and CR3022 can only bind when 2 RBDs are in the 'up' configuration<sup>5</sup>. The variant with K986P showed higher binding of SAD-S35, ACE2, and CR3022 than S-2P, measured with BioLayer Interferometry (Fig. 1d left panel), in accordance with the results obtained with SEC and AlphaLISA, both in cleared crude supernatants. D614N and T572I showed very low binding to SAD-S35 and ACE2 and almost no CR3022 binding (Fig. 1d right panel), indicating more closed trimers. A892P improved trimer closure to a lesser extent, while A942P seemed to increase its opening. These mutants likely exhibit a mixture of closed, 1-up, and 2-up structures.

The purified S variants with higher  $T_{m50}$  and lower Ab binding compared to S-2P also showed longer retention times in SEC

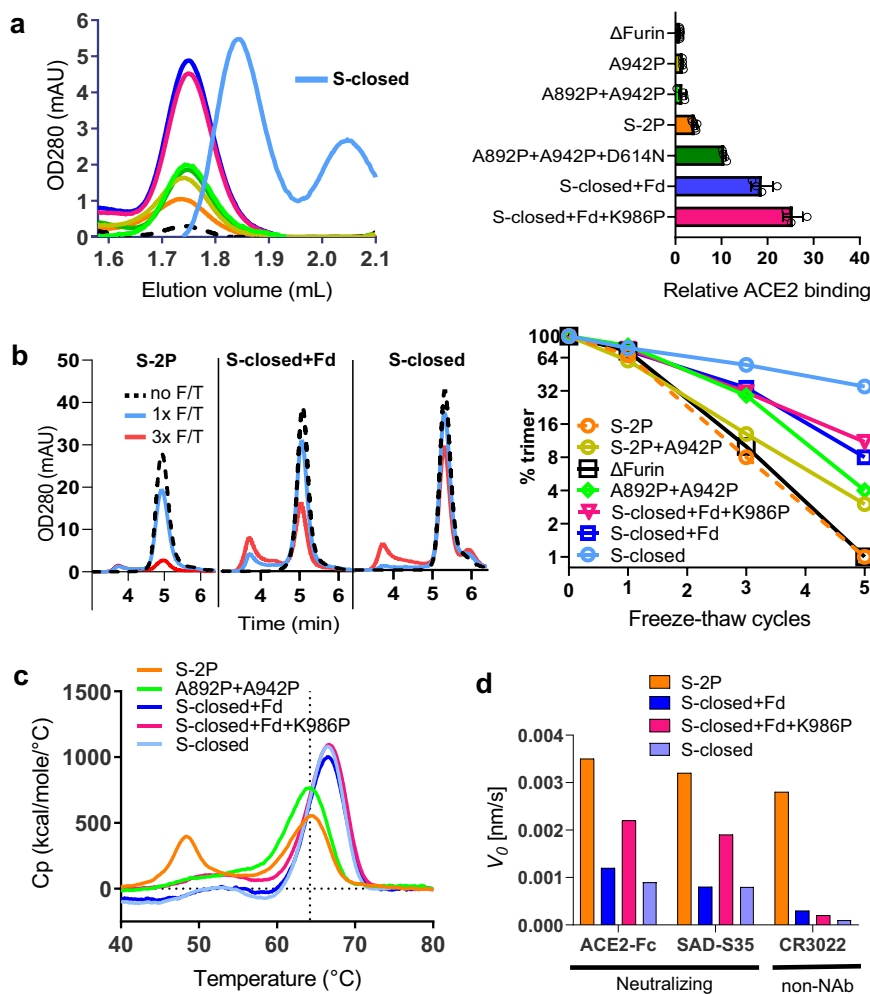


**Fig. 1** Characterization of SARS-CoV-2 S mutants containing single stabilizing mutations. **a** Analytical SEC of S variants on an SRT-10C SEC-500 15 cm column showing the trimer peak (solid line) relative to the control, which is the un-mutated backbone (dashed line). **b** ACE2-Fc binding to S protein mutants based on AlphaLISA of  $\Delta$ Furin variants (left panel) and S-2P (right panel). Data are represented as mean + SD of  $n = 4$  biologically independent samples in one experiment. Mutants were grouped according to the structural regions indicated in light gray. **c** Temperature stability of purified S trimers as measured by DSC. Two melting events are indicated by Tm1 and Tm2. **c** (left panel) Uncleaved SARS2-S variants with furin site mutations ( $\Delta$ Furin), with one stabilizing proline mutation in the hinge loop ( $\Delta$ Furin K986P or  $\Delta$ Furin V987P), and both proline mutations in the hinge loop (S-2P); (middle panel)  $\Delta$ Furin variants with indicated mutations in S1 and (right panel)  $\Delta$ Furin variants with indicated mutations in S2. **d** Binding of SAD-S35, ACE2, and CR3022 to purified S proteins measured with BiLayer Interferometry, showing the initial slope  $V_0$  at the start of binding. Source data are provided as a Source Data file.

(Supplementary Fig. 3), in agreement with a more compact structure. The largest shift was caused by D614N. Since 614 corresponds to a position in the S protein subject to the most extensive adaptation to the human host-D614G<sup>32</sup>, this variant was also tested for expression and stability. The results show that the D614G change has a very similar effect as D614N. Both increase trimer yields (Fig. 1a), increase Tm by ~2 °C (Fig. 1c), and reduce SAD-S35, ACE2, and CR3022 binding substantially

(Supplementary Fig. 5). Interestingly, both D614G and D614N full-length variants show increased fusogenicity compared to wild-type D614 (Supplementary Fig. 4).

Next, we made variants with combinations of selected mutations to evaluate if their effects on expression, trimer closure, and stability are additive. Because of their positive effect on the stability of the spike and location in different protein domains, both D614N and A892P were selected. In addition, A942P was selected



**Fig. 2** Characterization of SARS-CoV-2 S mutants containing combinations of stabilizing mutations. **a** Analysis of cell culture supernatant after transfection using analytical SEC using an SRT-10C SEC-500 15 cm column (left panel) and ACE2-Fc binding based on AlphaLISA (right panel) of  $\Delta$ Furin S with combinations of stabilizing mutations. Data are represented as mean + SD of  $n = 4$  biologically independent samples in one experiment. **b** Freeze-thaw stability of purified uncleaved S trimers with indicated stabilizing mutations as measured by analytical SEC. Chromatograms are shown for non-frozen (dashed black line), 1×frozen (light blue line), and 3×frozen (red line). **c** Temperature stability of purified combined S trimer variants by DSC. **d** Binding of SAD-S35, ACE2, and CR3022 to purified combined S trimer variants measured with BiLayer Interferometry, showing the initial slope  $V_0$  at the start of binding. Source data are provided as a Source Data file.

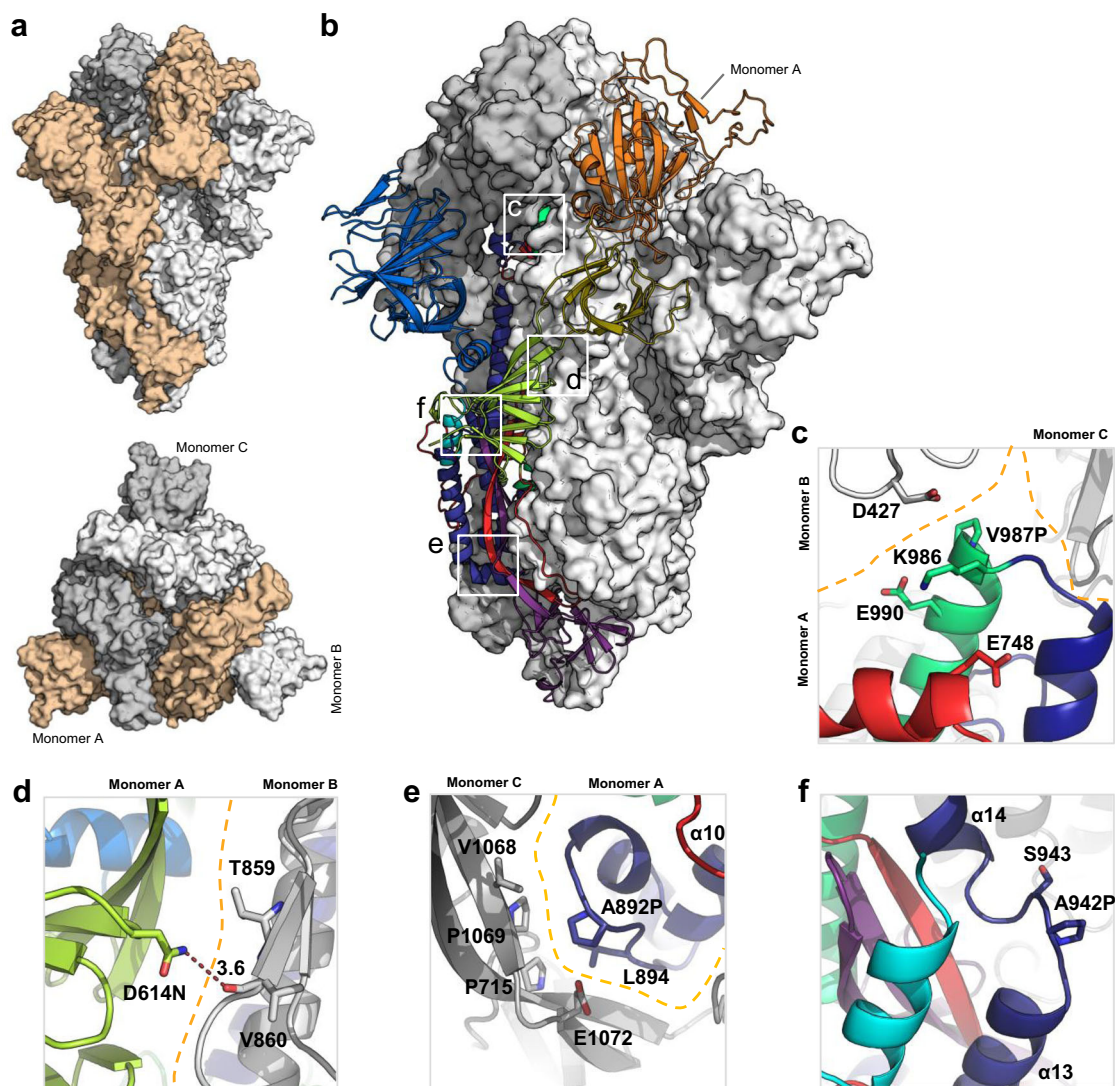
for its strong enhancement of protein yields (Fig. 1a). These mutations were combined with hinge loop prolines resulting in two combos, a quadruple mutant S-closed + Fd containing D614N + A892P + A942P + V987P and a quintuple mutant S-closed + Fd + K986P. Although both combos showed a similar, approximately 5-fold improvement in yields compared to S-2P, the addition of K986P increased ACE2 binding with similar expression levels (Fig. 2a). Interestingly, the quadruple mutant in which the Fd trimerization domain is deleted (S-closed), showed a 6.4-fold improvement in yield compared to S-2P. Its trimer peak was shifted towards a longer retention time due to the smaller size which was confirmed by MALS analysis (Fig. 2a and Supplementary Table 2).

S-closed with and without Fd were further characterized for resilience during repeated freeze-thaw cycles (Fig. 2b and Supplementary Table 3). Analytical SEC showed that only 8% of the original S-2P trimer was present after 3, and only 1% after 5 cycles (Fig. 2b, right panel). The trimer content is significantly improved for S-closed, which retained 55% of intact trimers after 3, and 35% after 5 freeze-thaw cycles. All combos displayed higher thermal stability with only a single  $T_{m50}$  at about 66 °C

(Fig. 2c) and decreased levels of ACE2 and antibody binding compared to S-2P control (Fig. 2d and Supplementary Fig. 5).

The S-closed+Fd quadruple mutant was then imaged by cryo-EM. A 2-steps 3D classification illustrates that out of 833,000 classified particles, ~80% was closed with all RBDs in the down state and 38% was categorized into a well-defined closed class while ~20% showed 1 RBD-up (Supplementary Fig. 6). Further processing of the 320,000 closed conformation particles allowed us to obtain a 2.8 Å electron potential map for the closed conformation and a 3.0 Å electron potential map for the 1 RBD-up (one up) conformation (Supplementary Fig. 7). An atomic model that was built into the 2.8 Å electron potential map confirmed that S retains the prefusion spike conformation (Fig. 3). The NTD and RBD density is less defined than for the rest of the map, suggesting flexibility in these regions. The closed structure (see Fig. 3a) is highly reminiscent of the one previously solved by Walls et al.<sup>26</sup>. The two structures differ by 2.2 Å all-atom RMSD and there are no significant differences in terms of the relative position of domains or domain conformation. The RBD is relatively more defined and the NTD less defined compared to the closed trimer described by Walls et al. The stabilizing mutations





**Fig. 3 Structural characterization of S-closed-Fd.** **a** Cryo-EM structure of the most abundant trimer class of S-closed-Fd—a closed S protein trimer (PDB accession code: 7A4N). Monomers are colored in light orange, white and gray and the spike is shown from the side (upper panel) and from the top (lower panel) view. **b** S-closed-Fd trimer with monomer A plotted as cartoon. Structural domains are colored according to the same color code as in Supplementary Fig. 1. Areas where stabilizing mutations are introduced are indicated with white squares. **c–f** Each of the four single-point mutations introduced in S-closed-Fd is shown in detail. Domains of the new structure are colored according to the same color code as used in Supplementary Fig. 1. The boundary between monomers has been additionally indicated with an orange dashed line where applicable. Secondary structure  $\alpha 10$ ,  $\alpha 13$ , and  $\alpha 14$  are indicated in panels (**e** and **f**).

do not significantly affect the backbone conformation of the closed trimer.

### Discussion

SARS-CoV-2 S protein is unstable<sup>9,12,23–27</sup> and although the introduction of a double proline (K986P and V987P) in the hinge loop at the C-terminus of HR1 was shown to improve the stability of the prefusion conformation<sup>9,26</sup>, the S protein still suffered from instability (Figs. 1c and 2b). We show that while each proline mutation increases trimer expression, the K986P variant also reduces the interaction with the S1 head, releasing the RBD to the ‘up’ configuration based on increased ACE2 and CR3022 binding (Figs. 1d and 2d) and a leftward shift in the SEC profile (Fig. 1a). The effect is partially compensated by the introduction of the second stabilizing mutation—V987P. We describe here two groups of substitutions that further stabilized the S-2P prefusion trimers and reduced RBD exposure. The first group of stabilizing mutations was identified in the S2 HR1 region that undergoes an

extensive conformational change during fusion. Introduction of stabilizing proline and glycine mutations in loops  $\alpha 10\alpha 11$  and  $\alpha 13\alpha 14$  and disulfides in loop  $\alpha 10\alpha 11$  and between the central helix and C-terminus of HR1 resulted in significantly increased trimer yields. It is likely that the mutations in S2 facilitate folding of the protein during expression by fixing loop conformations with otherwise intrinsic alpha-helical propensity that is necessary for driving the fusion conformational change. This strategy greatly improved prefusion S protein expression levels (Fig. 1a). Recently, Hsieh et al., in parallel to our study, demonstrated the stabilizing effect of the A942P and A892P mutations<sup>12</sup>. Mutations identified in the head region clustered near the SD1 subdomain. Interestingly, introduction of two of these mutations (D614N and T572I) into the S-2P backbone resulted, apart from increased expression, in a striking improvement of thermal stability (Fig. 1c) and very low RBD exposure (Fig. 1d).

Based on the biochemical characterization of single-point mutants and their subdomain placement we selected four

substitutions with low surface exposure (D614N, A892P, A942P, and V987P) to create S-closed. This construct does not contain a heterologous trimerization domain and exhibits a 6.4-fold increase in expression (Fig. 2a), high thermal and freeze–thaw stability (Fig. 2b, c), and antigenicity reminiscent of a closed trimer (Fig. 2d). Both D614G and D614N variants show increased fusogenicity and stability (Fig. 1c and Supplementary Fig. 4), which may be explained by a decrease in premature shedding of S1<sup>33</sup>. The interaction of D614 with the fusion peptide proximal region (FPPR) (residues 828–853)<sup>27</sup> may play a role in stabilizing the spike during expression, as recently elucidated by a low pH structure<sup>34</sup>, in which D614 is in direct proximity of D839 and forms a hydrogen bond with the carbonyl of V860, suggesting aspartate protonation. In our new structure, N614 forms two interprotomeric hydrogen bonds with the sidechain of T859 and backbone carbonyl of V860 (Fig. 3b), effectively mimicking the stabilizing effect of the protonated aspartate residue, without the pH dependence. It remains unknown why the D614G has a similar positive effect on stability, but perhaps the high cluster of negative charges in the interface between S1 and S2 destabilizes the protein, and the 614N and 614G mutations reduce this repulsion. None of the stabilizing proline mutations modify the local structure of the protein, while A892P adds new interprotomeric interactions with the neighboring hybrid sheet, composed of both S1 and S2 strands (Fig. 3b). The stabilizing mutations and presence of K986, which interacts with Asp427 in the RBD of the neighboring monomer and Glu748 in S2 (Fig. 3b), allow the S trimer to maintain predominantly the closed prefusion configuration.

The viral spike is mostly closed<sup>28</sup> and structurally very similar to other known closed S-2P spike conformations<sup>25,26</sup> and especially the closed wild-type structure of Xiong et al.<sup>25</sup>, with an all-atom RMSD of 1.7 Å. The FPPR is not visible as in the so-called locked S structure of Xiong et al.<sup>25</sup> or the structure of Cai et al.<sup>27</sup>, both showing tighter packing of the head domains against S2. Although many neutralizing antibodies are directed against the RBD, the antibodies with the highest neutralizing activity bind the RBD in the down state<sup>4,6,13</sup>. Similar to other class I fusion proteins, a closed conformation may be more reminiscent of a transmitted virus, and antibodies that recognize the closed state may be more important for protection as has been shown for HIV<sup>35</sup>. Furthermore, the closed conformation can potentially induce antibodies with higher cross-reactivity since the downstate surface is more conserved. A stable closed S trimer with minimal non-exposed mutations and without a Fd that shows a significant increase in expression levels may advance the development of novel (subunit) vaccine immunogens and further improve genetic vaccines, diagnostics, or isolation of antibodies.

## Methods

**Protein expression and purification.** A plasmid corresponding to the semi-stabilized SARS-CoV-2 S-2P protein<sup>9</sup> was synthesized, codon-optimized, cloned into pCDNA2004, and sequenced at GenScript (Piscataway, NJ 08854), where also all the variants with different amino acid substitutions were generated. A variant with a HIS-tag and a variant with a C-tag were purified. The expression platform used was the Expi293F cells. The cells were transiently transfected using ExpiFectamine (Life Technologies) according to the manufacturer's instructions and cultured for 6 days at 37 °C and 10% CO<sub>2</sub>. The culture supernatants were harvested and spun for 5 minutes at 300 × g to remove cells and cellular debris. The spun supernatants were subsequently sterile filtered using a 0.22 μm vacuum filter and stored at 4 °C until they were purified within 1–2 days of harvest. HIS-tagged SARS-CoV-2 S trimers were purified using a two-step purification protocol by 1- or 5-ml complete HIS-tag columns (Roche). Proteins were further purified by size-exclusion chromatography using a HiLoad Superdex 200 16/600 column (GE Healthcare).

**Antibodies and reagents.** SAD-S35 was purchased at Acro Biosystems. ACE2-Fc was made according to Liu et al. (2018)<sup>36</sup>, *Kidney international*. For CR3022, the heavy and light chains were cloned into a single IgG1 expression vector to express a

fully human IgG1 antibody. CR3022 was made by transfecting the IgG1 expression construct using the ExpiFectamine™ 293 Transfection Kit (ThermoFisher) in Expi293F (ThermoFisher) cells according to the manufacturer's specifications. CR3022 was purified from serum-free culture supernatants using mAb Select SuRe resin (GE Healthcare) followed by rapid desalting using a HiPrep 26/10 Desalting column (GE Healthcare). The final formulation buffer was 20 mM NaAc, 75 mM NaCl, and 5% Sucrose pH 5.5.

**Differential scanning calorimetry (DSC).** Melting temperatures for S protein variants were determined using a PEAQ-DSC system. In all, 325 μl of 0.3 mg/ml protein sample was used per measurement. The measurement was performed with a start temperature of 20 °C and a final temperature of 100 °C at a scan rate 100 °C/h and the feedback mode; Low (= signal amplification).

**Differential scanning fluorometry (DSF).** A total of 0.2 mg of purified protein in 50 μl PBS pH 7.4 (Gibco) was mixed with 15 μl of 20 times diluted SYPRO orange fluorescent dye (5000×stock, Invitrogen S6650) in a 96-well optical qPCR plate. A negative control sample containing the dye only was used for reference subtraction. The measurement was performed in a qPCR instrument (Applied Biosystems ViiA 7) using a temperature ramp from 25 to 95 °C with a rate of 0.015 °C per second. Data were collected continuously. The negative first derivative was plotted as a function of temperature. The melting temperature corresponds to the lowest point in the curve.

**BioLayer Interferometry (BLI).** A solution of SAD-S35 at a concentration of 0.5 μg/ml and ACE2-Fc and CR3022 at a concentration of 10 μg/ml was used to immobilize the ligand on anti-hIgG (AHC) sensors (FortéBio, cat. #18–5060) in 1×kinetics buffer (FortéBio, cat. #18–1092) in 96-well black flat-bottom polypropylene microplates (FortéBio, cat. #3694). The experiment was performed on an Octet RED384 instrument (Pall-FortéBio) at 30 °C with a shaking speed of 1000 rpm. Activation was 600 s, immobilization of antibodies 900 s, followed by washing for 600 s, and then binding the S proteins for 300 s. The data analysis was performed using the FortéBio Data Analysis 12.0 software (FortéBio).

## Cryo-EM

**Cryo-EM Grid Preparation and Data Collection.** SARS-CoV-2 S protein samples were prepared in 20 mM Tris, 150 mM NaCl, pH 7 buffer at a concentration of 0.15 mg/ml and applied to glow discharged Quantifoil R2/2 200 mesh grids before being double side blotted for 3 s in a Vitrobot Mark IV (Thermo Fisher Scientific) and plunge frozen into liquid ethane cooled. Grids were loaded into a Titan Krios electron microscope (Thermo Fisher Scientific) operated at 300 kv, equipped with a Gatan K3 BioQuantum direct electron detector. A total of 9760 movies were collected over two microscopy sessions at The Netherlands Center for Electron Nanoscopy (NeCEN). Detailed data acquisition parameters are summarized in Supplementary Table 4.

**Cryo-EM image processing.** Collected movies were imported into RELION-3.1-beta<sup>37</sup> and subjected to beam-induced drift correction using MotionCor2<sup>38</sup> and CTF estimation by CTFFIND-4.1.18<sup>39</sup>. Detailed steps of the image processing workflow are illustrated in Supplementary Figure 7. Final reconstructions were sharpened and locally filtered in RELION post-processing.

**Model building and refinement.** The SARS-CoV-2 S PDBID 6VXX and 6VSB structures<sup>9,26</sup> were used as starting models. PHENIX-1.18.261<sup>40</sup>, Coot<sup>41</sup>, and the Namdinator webserver<sup>42</sup> were iteratively used to build atomic models. Geometry and statistics are given in Supplementary Table 4 and Supplementary Table 5. Final maps were displayed using UCSF ChimeraX<sup>43</sup>.

**AlphaLISA.** Crude supernatants were diluted 300 times in AlphaLISA buffer (PBS + 0.05% Tween-20 + 0.5 mg/mL BSA). Then, 10 μl of each dilution was transferred to a 96-well plate and mixed with 40 μl acceptor beads, donor beads, and ACE2-Fc. The donor beads were conjugated to ProtA (Cat. #AS102M, Perkin Elmer), which binds to ACE2Fc. The acceptor beads were conjugated to an anti-Flag antibody (Cat. #AL112M, Perkin Elmer), which binds to the Flag-tag of the construct. The mixture of the supernatant containing the expressed S protein, the ACE-2-Fc, donor beads, and acceptor beads was incubated at room temperature for 2 h without shaking. Subsequently, the chemiluminescent signal was measured with an EnSight plate reader instrument (Perkin Elmer). The average background signal attributed to mock-transfected cells was subtracted from the AlphaLISA counts. Subsequently, the whole data set was divided by the signal measured for the SARS CoV-2 S protein having the S backbone sequence signal to normalize the signal for each of the S variants tested to the backbone.

**Analytical SEC.** An ultra-high-performance liquid chromatography system (Vanquish, Thermo Scientific) and μDAWN TREOS instrument (Wyatt) coupled to an Optilab μT-rEX Refractive Index Detector (Wyatt), in combination with an in-line Nanostar DLS reader (Wyatt), was used for performing the analytical SEC experiment. The cleared crude cell culture supernatants were applied to an



SRT-10C SEC-500 15 cm column (Sepax Cat. #235500–244615) with the corresponding guard column (Sepax) equilibrated in running buffer (150 mM sodium phosphate, 50 mM NaCl, pH 7.0) at 0.35 mL/min. When analyzing supernatant samples,  $\mu$ MALS detectors were offline and analytical SEC data were analyzed using Chromeleon 7.2.8.0 software package. The signal of supernatants of non-transfected cells was subtracted from the signal of supernatants of S transfected cells. When purified proteins were analyzed using SEC-MALS,  $\mu$ MALS detectors were inline and data were analyzed using Astra 7.3 software package. For the protein component, a dn/dc (mL/g) value of 0.1850 was used and for the glycan component a value of 0.1410. Molecular weights were calculated using the RI detector as [C] source and mass recoveries using UV as [C] source.

**Cell-cell fusion assay.** Quantitative cell–cell fusion assays were performed to ascertain the relative fusogenicity of the different D614 S protein variants by using the NanoBIT complementation system (Promega). Donor HEK293 cells were transfected with full-length S and the 11 S subunit in 96-well white flat-bottom TC-treated microtest assay plates. Acceptor HEK293 cells were transfected in six-well plates (Corning) with ACE2, TMPRSS2, and the PEP86 subunit, or just the PEP86 subunit (“Mock”) as a negative control. All proteins were expressed from pcDNA2004 plasmids using Trans-IT transfection reagent according to the manufacturer’s instructions. Eighteen hours after transfection, the acceptor cells were released by 0.1% trypsin/EDTA and added to the donor cells at a 1:1 ratio for 4 h. Luciferase complementation was measured by incubating with Nano-Glo<sup>®</sup> Live Cell Reagent for 3 m, followed by read-out on an Ensiht plate reader (PerkinElmer).

### Data availability

The authors declare that the main data supporting the findings of this study are available within the article and its Supplementary Information files. Some data that support the findings of this study are deposited in the Protein Data Bank database with accession codes 7A4N and 7AD1. Source data are provided with this paper.

Received: 22 July 2020; Accepted: 24 November 2020;

Published online: 11 January 2021

### References

- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
- Chen, W. H., Hotez, P. J. & Bottazzi, M. E. Potential for developing a SARS-CoV receptor-binding domain (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease (COVID)-19. *Hum. Vaccin. Immunother.* <https://doi.org/10.1080/21645515.2020.1740560> (2020).
- Liu, L. et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).
- Yuan, M. et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**, 630–633 (2020).
- Brouwer, P. J. M. et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* <https://doi.org/10.1126/science.abc5902> (2020).
- Bosch, B. J., van der Zee, R., de Haan, C. A. & Rottier, P. J. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* **77**, 8801–8811 (2003).
- Li, F. Structure, function, and evolution of coronavirus spike. *Proteins Annu. Rev. Virol.* **3**, 237–261 (2016).
- Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
- Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280 e278 (2020).
- Bestle, D. et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* <https://doi.org/10.26508/lsa.202000786> (2020).
- Hsieh, C. L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* <https://doi.org/10.1126/science.abd0826> (2020).
- Robbiani, D. F. et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* <https://doi.org/10.1038/s41586-020-2456-9> (2020).
- Graham, B. S., Gilman, M. S. A. & McLellan, J. S. Structure-based vaccine antigen design. *Annu. Rev. Med.* **70**, 91–104 (2019).
- Battles, M. B. et al. Structure and immunogenicity of pre-fusion-stabilized human metapneumovirus F glycoprotein. *Nat. Commun.* **8**, 1528 (2017).
- Hastie, K. M. et al. Structural basis for antibody-mediated neutralization of Lassa virus. *Science* **356**, 923–928 (2017).
- Krurup, A. et al. A highly stable prefusion RSV F vaccine derived from structural analysis of the fusion mechanism. *Nat. Commun.* **6**, 8143 (2015).
- Rutten, L. et al. Structure-based design of prefusion-stabilized Filovirus Glycoprotein Trimers. *Cell Rep.* **30**, 4540–4550 e4543 (2020).
- Rutten, L. et al. A universal approach to optimize the folding and stability of prefusion-closed HIV-1 envelope trimers. *Cell Rep.* **23**, 584–595 (2018).
- Sanders, R. W. et al. Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. *J. Virol.* **76**, 8875–8889 (2002).
- Pallesen, J. et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl Acad. Sci. USA* **114**, E7348–E7357 (2017).
- Kirchdoerfer, R. N. et al. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci. Rep.* **8**, 15701 (2018).
- Henderson, R. et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* <https://doi.org/10.1038/s41594-020-0479-4> (2020).
- McCallum, M., Walls, A. C., Bowen, J. E., Corti, D. & Veesler, D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nat. Struct. Mol. Biol.* <https://doi.org/10.1038/s41594-020-0483-8> (2020).
- Xiong, X. et al. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* <https://doi.org/10.1038/s41594-020-0478-5> (2020).
- Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292 e286 (2020).
- Cai, Y. et al. Distinct conformational states of SARS-CoV-2 spike protein. *Science* <https://doi.org/10.1126/science.abd4251> (2020).
- Ke, Z. et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* <https://doi.org/10.1038/s41586-020-2665-2> (2020).
- Walls, A. C. et al. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* **531**, 114–117 (2016).
- Kuhlman, B. et al. Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364–1368 (2003).
- Salam, N. K., Adzhigirey, M., Sherman, W. & Pearlman, D. A. Structure-based approach to the prediction of disulfide bonds in proteins. *Protein Eng. Des. Sel.* **27**, 365–374 (2014).
- Korber, B. et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 1–16 (2020).
- Zhang, L. et al. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *bioRxiv* <https://doi.org/10.1101/2020.06.12.148726> (2020).
- Zhou, T. et al. A pH-dependent switch mediates conformational masking of SARS-CoV-2 spike. *bioRxiv* <https://doi.org/10.1101/2020.07.04.187989> (2020).
- Sanders, R. W. & Moore, J. P. Native-like Env trimers as a platform for HIV-1 vaccine design. *Immunol. Rev.* **275**, 161–182 (2017).
- Liu, P. et al. Novel ACE2-Fc chimeric fusion provides long-lasting hypertension control and organ protection in mouse models of systemic renin angiotensin system activation. *Basic Research* **94**, P114–125 (2018).
- Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* <https://doi.org/10.7554/eLife.42166> (2018).
- Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
- Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
- Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D.* **75**, 861–877 (2019).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
- Kidmose, R. T. et al. Namdinator-automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ.* **6**, 526–531 (2019).
- Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

### Acknowledgements

This work benefited from access to the Netherlands Centre for Electron Nanoscopy (NeCEN) at Leiden University, an Instruct-ERIC center with assistance from Wen Yang and Frederic Bonnet. We thank Wouter Koudstaal for advice and assistance. We thank Ilona Bisschop, Martijn de Man, Ava Sadi, Anne-Marie de Goovert, Lam Le, and Annemart Koornneef for technical support. We thank Xiaodi Yu and Sujata Sharma for support with structure refinement.

### Author contributions

J.J., L.R., M.J.G.B. and J.P.M.L. designed the study. J.J., L.R. and J.P.M.L. performed structure-based design of mutations. T.R., S.B., P.B. and R.V. planned and/or performed biochemical assays and purifications. L.L.R.R. performed EM sample preparation, data collection, data processing, and analysis. J.J., L.R., M.J.G.B., L.L.R.R. and J.P.M.L. wrote the paper.

### Competing interests

J.J., L.R., M.J.G.B. and J.P.M.L. are co-inventors on related vaccine patent applications. J.J., L.R., S.B., P.B., R.V., T.R., M.J.G.B. and J.P.M.L. are employees of Janssen Vaccines & Prevention B.V.; L.R., J.J., and J.P.M.L. hold stock of Johnson & Johnson.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-20321-x>.

**Correspondence** and requests for materials should be addressed to J.P.M.L.

**Peer review information** *Nature Communications* thanks Priyamvada Acharya and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021