**Virologica Sinica**

**Supplementary Data**

**Structures of SARS-CoV-2 spike protein alert noteworthy sites for the potential approaching variants**

Xiaorui Xing a,b,c,1, Lei Wang c,d,1, Zhen Cui c,d,1, Wangjun Fu c,d,1, Tao Zheng c,d, Lili Qin e, Pingju Ge e, Aidong Qian a,b,\* , Nan Wang c,\*, Shuai Yuan f,\*

a *College of Veterinary Medicine, Jilin Agricultural University**,* *Changchun**130118, China*

b *College of Animal Science and Technology, Jilin Provincial Engineering Research Center of Animal Probiotics, Key Lab of Animal Production, Product Quality and Security, Joint Laboratory of Modern Agricultural Technology International Cooperation, Ministry of Education, Jilin Agricultural University, Changchun 130118, China*

c *CAS Key Laboratory of Infection and Immunity, National Laboratory of Macromolecules, I**nstitute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China*

d *University of Chinese Academy of Sciences, Beijing 100101, China*

e *Acrobiosystems, Beijing 100101, China*

f *State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China*

\* Corresponding authors.

*E-mail addresses:* yuanshuai@wh.iov.cn (S.Yuan); wangnan161@ibp.ac.cn (N. Wang); qianaidong0115@163.com (A. Qian).

*ORCID: 0000-0003-4037-3429 (S.Yuan); 0000-0002-9936-9179 (N. Wang); 0000-0002-9640-4653 (A. Qian)*

1 Xiaorui Xing, Lei Wang, Zhen Cui and Wangjun Fu contributed equally to this work.

**Methods and Materials**

**Protein expression and purification**

The plasmids encoding the full-length spike (S) protein (residues 11028) of wild-type SARS-CoV-2 (GenBank: MN908947) was used as the template for the construction of the *S* gene of Lambda (G75V, T76I, RSYLTPG246252del, D253N, L452Q, F490S, D614G, T859N), Mu (ACROBiosystems, Cat No. SPN-C52Ha) (T95I, Y144S, Y145N, R346K, E484K, N501Y, D614G, P681H, D950N), C.1.2 (ACROBiosystems, Cat No. SPN-C82Eq) (C136F, Y144del, R190S, D215G, LA242-243del, Y449H, E484K, N501Y, D614G, H65) and B.1.620 (ACROBiosystems, Cat No. SPN-C82Ed) (P26S, H69del, V70del, V126A, Y144del, LLA241-243del, H245Y, S477N, E484K, D614G, P681H, T1027I, D1118H) by overlapping PCR. Six proline substitutions at residues 817, 892, 899, 942, 986, and 987 and two alanine substitutions at residues 683 and 685 and a C-terminal T4 fibritin foldon domain were also introduced into all the full-length S gene constructs to facilitate the protein expression and stabilization of the trimer conformation (ACROBiosystems, Cat No. SPN-C52Hz). All the constructs described above were attached with a C-terminal six-His for protein purification. For protein expression, the plasmids constructed above were transiently transfected into HEK293F cells grown in suspension at 37 °C in an incubator supplied with 8% CO2, rotating at 130 rpm. The supernatant was harvested and concentrated after 72-hour incubation, and exchanged into the binding buffer by tangential flow filtration cassette. The protein of interest was separated by affinity chromatography using resin attached with Ni-NTA and additionally purified by size exclusion chromatography using a Superose 6 10/300 column (GE Healthcare) with the buffer containing 20 mmol/L Tris pH 8.0 and 200 mmol/L NaCl.

**Cryo-EM sample preparation and data collection**

For cryo-EM samples preparation of variants, a 3 µL aliquot of purified variant spike protein was transferred onto a freshly glow-discharged C-flat 1.2/1.3 Au grid and blotted for 3 s in 100% relative humidity for plunge-freezing (Vitrobot; Thermo Fisher Scientific) in liquid ethane. Cryo-EM datasets were collected with 300 kV Titan Krios (Thermo Fisher Scientific) equipped with K3 detector (Gatan, Pleasanton, CA). Movies (32 frames, 0.2 s per frame, total dose 60 e-Å-2) were recorded with defocus range from 1.5 to 2.5 µm with SerialEM and yields a final pixel size of 1.07 Å.

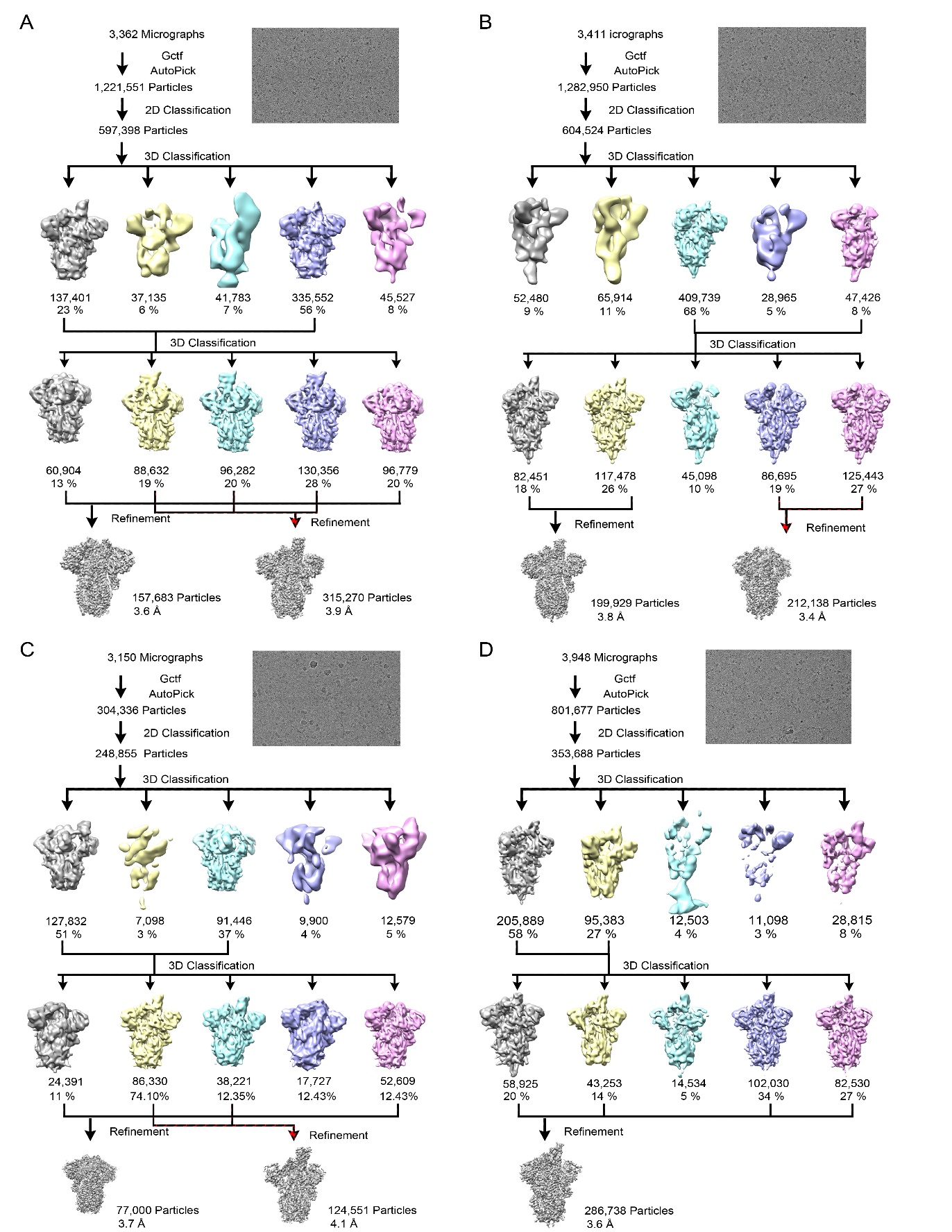
**Image processing, three-dimensional reconstruction, model building and refinement**

To detect the structures of SARS-CoV-2 variant spike proteins, 3,362 movies of B.1.620, 3,411 movies of C.1.2, 3,150 movies of Mu, and 3,982 movies of Lambda variants were recorded. The micrographs from each movie were aligned and averaged for the correction of beam-induced drift using MOTIONCOR2 (Zheng et al., 2017). The contrast transfer function parameters for each micrograph were estimated by Gctf (Zhang, 2016). Particles from micrographs were picked automatically, extracted, and screened with 2D classification. A total of 597,398 particles of B.1.620, 604,524 particles of C.1.2, 248,855 particles of Mu, and 353,688 particles of Lambda variant spike protein were selected and sorted with two rounds of 3D classifications. All the particles in major classes (157,683 particles for B.1.620 close state, 315,270 particles for B.1.620 one RBD up state, 212,138 particles for C.1.2 close state, 199,929 particles for C.1.2 one RBD up state, 77,000 particles for Mu close state, 124,551 particles for Mu one RBD up state, and 286,738 particles for Lambda one RBD up state) were picked out and used for further high-resolution refinement. Future more, postprocessing program (estimate the B-factor automatically) was used to generate the final Cryo-EM maps of B.1.620 close state at 3.6 Å, B.1.620 one up state at 3.9 Å, C.1.2 close state at 3.4 Å, C.1.2 one up state at 3.8 Å, Mu close state at 3.7 Å, Mu one up state at 4.1 Å, and Lambda one up state at 3.6 Å. All the resolutions were obtained and evaluated on the basis of the gold-standard Fourier shell correlation (threshold = 0.143).

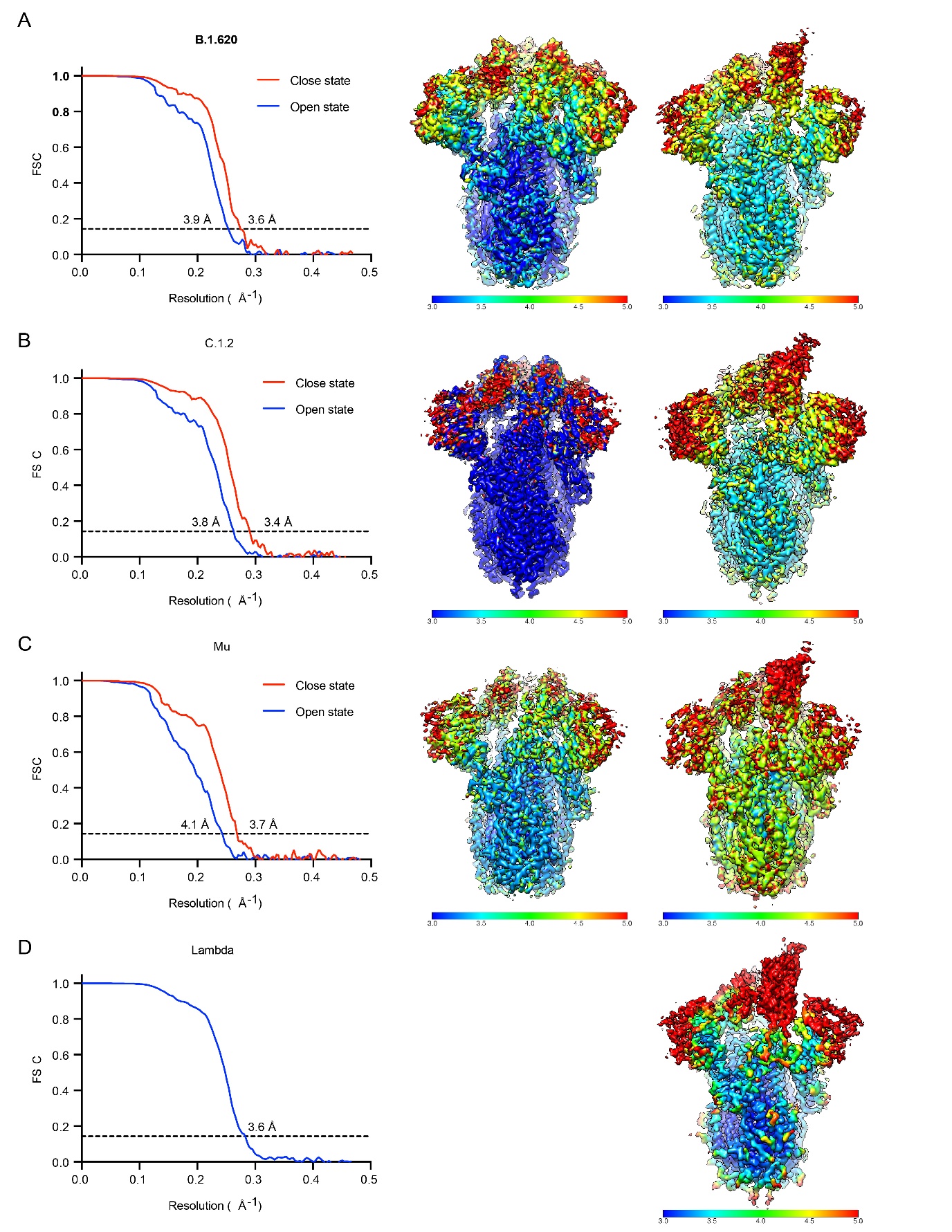
All procedures above were performed using Relion (Scheres, 2016) and Cryosparc. The local resolution was evaluated by ResMap (Kucukelbir et al., 2014). The structures of the SARS-CoV-2 S trimer (Protein Data Bank ID: 7WEA) were manually fitted into the refined EM maps of SARS-CoV-2 variants S trimers using the program Chimera (Yang et al., 2012) and further corrected manually by real-space refinement in COOT (Emsley and Cowtan, 2004). The atomic model was further refined by positional and B-factor refinement in real space using Phenix (Afonine et al., 2012) and rebuilding in COOT iteratively. The final models were evaluated by Molprobity (Chen et al., 2010). The datasets and refinement statistics are summarized in Supplementary Table S1.

**Surface Plasmon Resonance**

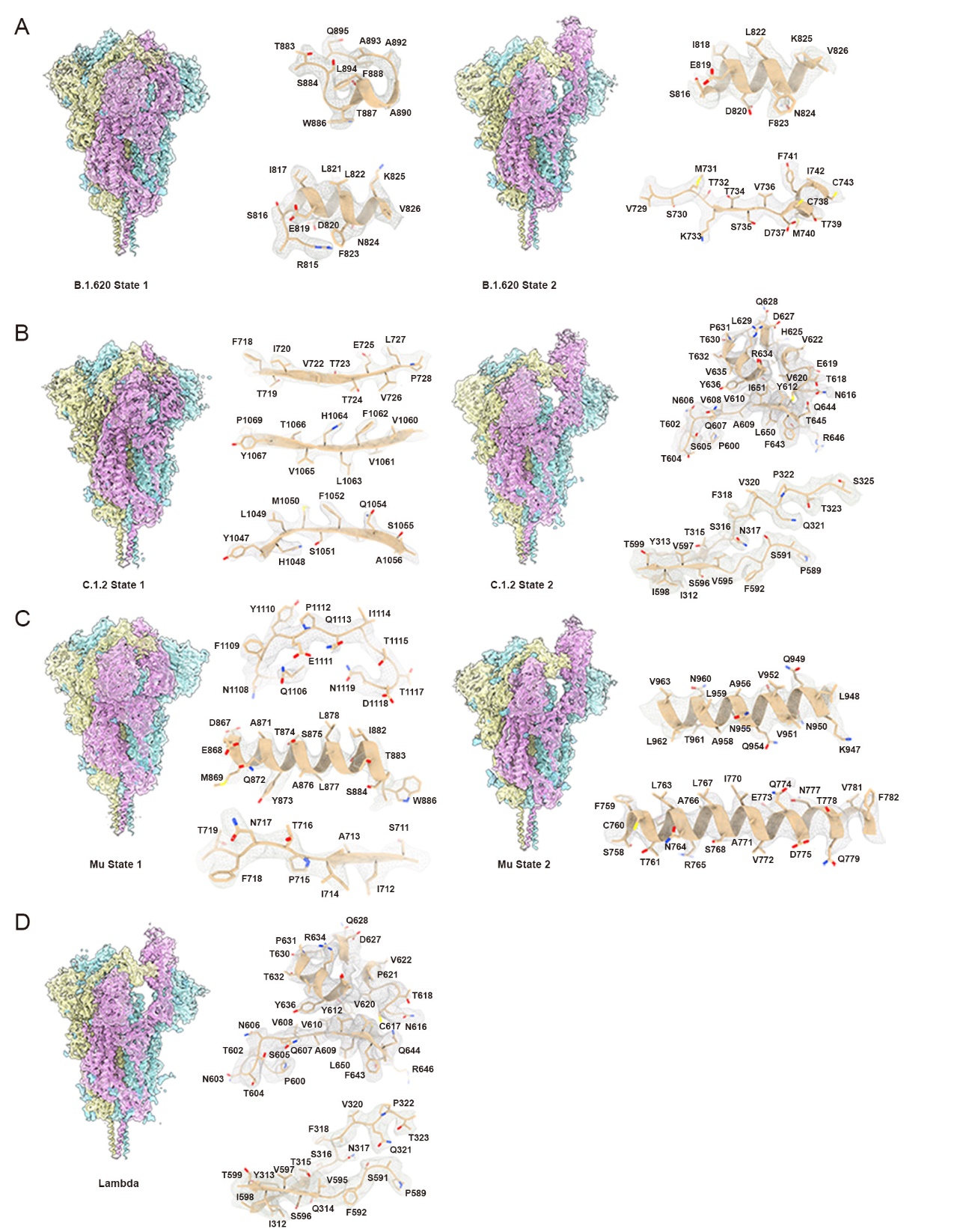
Human ACE2 was immobilized onto CM5 sensor chips using a Biacore 8K (GE Healthcare). Serial dilutions of purified RBD of wild-type SARS-CoV-2, C.1.2, Mu, and Lambda variants vawere injected, ranging in concentrations from 125 to 7.8 nM. The response units were recorded at room temperature using BIAcore 8K Evaluation Software (v3.0.12.15655; GE Healthcare), and the resulting data were fitted to a 1:1 binding model using BIAcore 8K Evaluation Software (v3.0.12.15655; GE Healthcare).

****

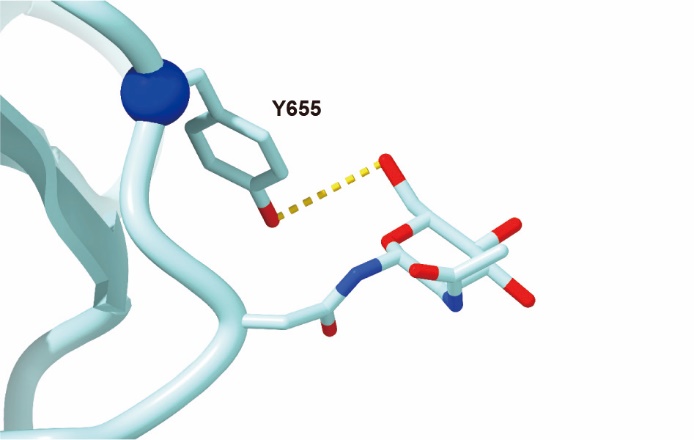
**Figure S1.** Cryo-EM images and data processing flowcharts of SARS-CoV-2 variants. **A** B.1.620; **B** C.1.2; **C** Mu; **D** Lambda.

****

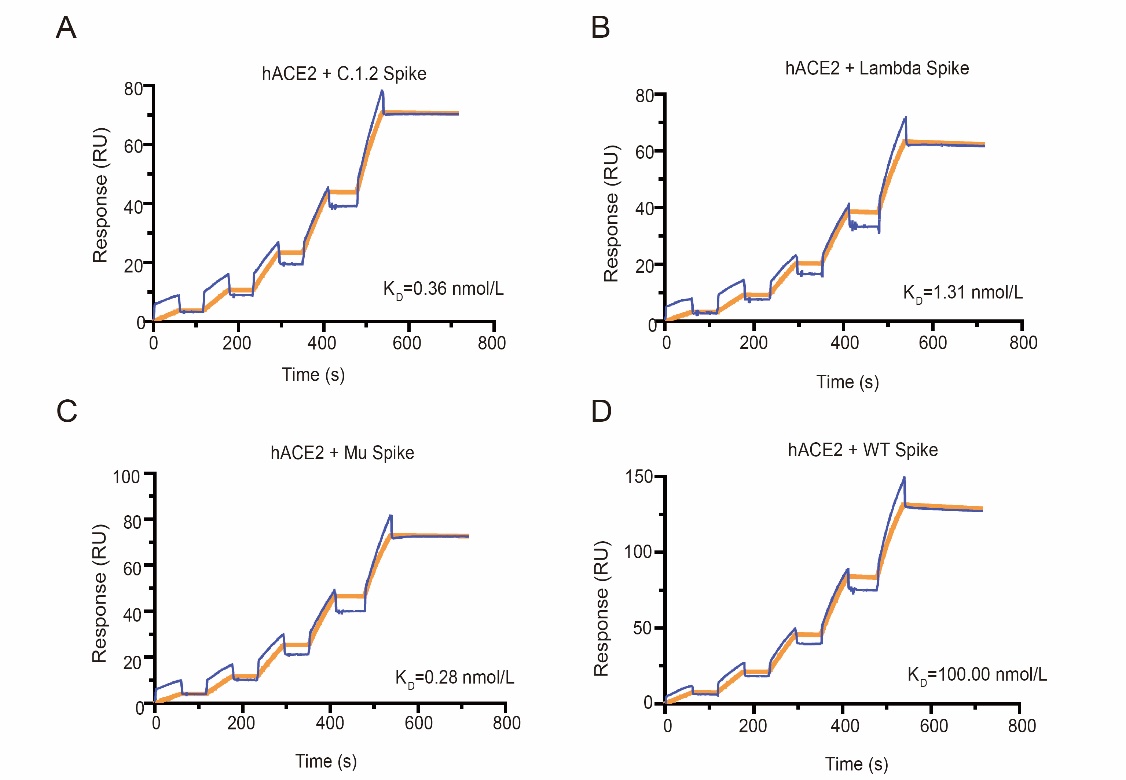
**Figure S2.** Resolution evaluation and local resolution assessments of the EM maps of variants. The gold-standard FSC curves (left) and ResMap assessed local resolution (right) of final maps are shown.

****

**Figure S3.** Density maps and atomic models. Cryo-EM maps of SARS-CoV-2 S trimer of B.1.620 (**A**), C.1.2 (**B**), Mu (**C**) and Lambda (**D**). The right panels present the density maps (mesh) and sidechains of related atomic models shown as sticks.

****

**Figure S4.** The zoom-in view of Y655 and the fixed NAG.



**Figure S5.** Binding affinities of variants spike to hACE2 measured by SPR. SPR analyses were conducted in biological duplicates. The results of SARS-CoV-2 S trimer binding affinities of C.1.2 (**A**), Lambda (**B**), Mu (**C**) and Wild-type (**D**).

**Table S1** Cryo-EM data collection and refinement statistics

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | B.1.620  (Close) | B.1.620  (One open) | C.1.2  (Close) | C.1.2  (One open) | Mu  (Close) | Mu  (One open) | Lambda |
| **Data collection and processing** |  |  |  |  |  |  |  |
| Magnification | 22,500 | | 22,500 | | 22,500 | | 22,500 |
| Voltage (kV) | 300 | | 300 | | 300 | | 300 |
| Electron exposure (e–/Å2) | 60 | | 60 | | 60 | | 60 |
| Defocus range (μm) | -1.5-2.5 | | -1.5-2.5 | | -1.5-2.5 | | -1.5-2.5 |
| Pixel size (Å) | 1.07 | | 1.07 | | 1.07 | | 1.07 |
| Symmetry imposed | C1 | C3 | C3 | C1 | C3 | C1 | C1 |
| Initial particle images (no.) | 1,221,551 | | 1,282,950 | | 304,336 | | 801,677 |
| Final particles images (no.) | 157,683 | 315,270 | 212,138 | 199,929 | 77,000 | 124,551 | 286,738 |
| Map resolution (Å) | 3.6 | 3.9 | 3.4 | 3.8 | 3.7 | 4.1 | 3.6 |
| FSC threshold | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 |
| Map resolution range (Å) | 3.660 | 3.960 | 3.460 | 3.860 | 3.760 | 4.160 | 3.660 |
| **Refinement** |  |  |  |  |  |  |  |
| Initial model used (PDB code) | 7WEA | 7WEA | 7WEA | 7WEA | 7WEA | 7WEA | 7WEA |
| Model resolution (Å) | 3.3 | 3.3 | 3.3 | 3.3 | 3.3 | 3.3 | 3.3 |
| FSC threshold | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 |
| Model resolution range (Å) | 3.360 | 3.360 | 3.360 | 3.360 | 3.360 | 3.360 | 3.360 |
| Map sharpening *B* factor (Å2) | 153.2 | 195.4 | 107.9 | 195.3 | 185.7 | 298.9 | 150.6 |
| Model composition |  |  |  |  |  |  |  |
| Non-hydrogen atoms | 27,288 | 26,685 | 27,456 | 26,559 | 27,702 | 26,798 | 26,966 |
| Protein residues | 3,285 | 3,248 | 3,309 | 3,268 | 3,345 | 3,260 | 3,297 |
| Ligands | 114 | 91 | 114 | 73 | 114 | 95 | 95 |
| *B* factors (Å2) |  |  |  |  |  |  |  |
| Protein | 65.67 | 62.66 | 23.11 | 40.38 | 65.19 | 107.40 | 114.35 |
| Ligand | 120.27 | 108.49 | 36.78 | 75.63 | 110.23 | 168.15 | 139.26 |
| R.m.s. deviations |  |  |  |  |  |  |  |
| Bond lengths (Å) | 0.005 | 0.004 | 0.003 | 0.002 | 0.003 | 0.003 | 0.010 |
| Bond angles (°) | 0.782 | 0.661 | 0.667 | 0.538 | 0.666 | 0.572 | 0.643 |
| Validation |  |  |  |  |  |  |  |
| MolProbity score | 1.92 | 1.78 | 1.78 | 1.59 | 1.79 | 1.85 | 1.94 |
| Clashscore | 10.38 | 10.22 | 8.1 | 7.10 | 8.85 | 11.75 | 7.19 |
| Poor rotamers (%) | 0.03 | 0.00 | 0.07 | 0.04 | 0.00 | 0.07 | 0.07 |
| Ramachandran plot |  |  |  |  |  |  |  |
| Favored (%) | 94.38 | 96.27 | 95.16 | 96.80 | 95.54 | 96.00 | 92.75 |
| Allowed (%) | 5.53 | 3.70 | 4.74 | 3.17 | 4.37 | 3.94 | 7.19 |
| Disallowed (%) | 0.09 | 0.03 | 0.09 | 0.03 | 0.09 | 0.06 | 0.06 |

**References:**

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 68, 352-367.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 66, 12-21.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 60, 2126-2132.

Kucukelbir, A., Sigworth, F.J., Tagare, H.D., 2014. Quantifying the local resolution of cryo-EM density maps. Nat Methods. 11, 63-65.

Scheres, S.H., 2016. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. Methods Enzymol. 579, 125-157.

Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C.C., Pettersen, E.F., Goddard, T.D., Meng, E.C., Sali, A., Ferrin, T.E., 2012. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J Struct Biol 179, 269-278.

Zhang, K., 2016. Gctf: Real-time CTF determination and correction. J Struct Biol. 193, 1-12.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., Agard, D.A., 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 14, 331-332.