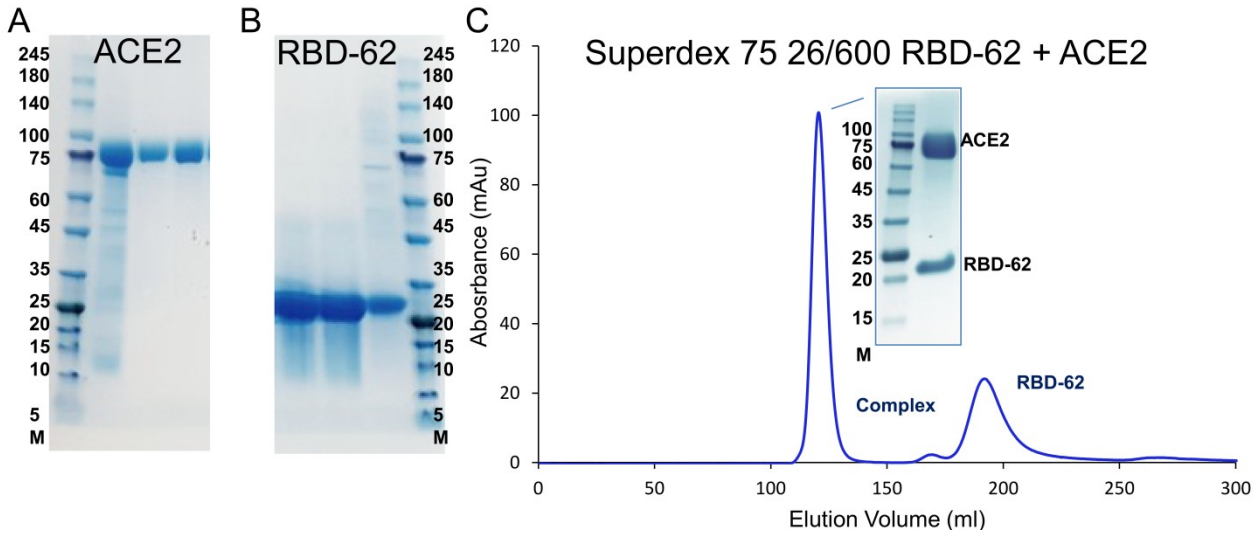
nature portfolio

https://doi.org/10.1038/s41564-021-00954-4

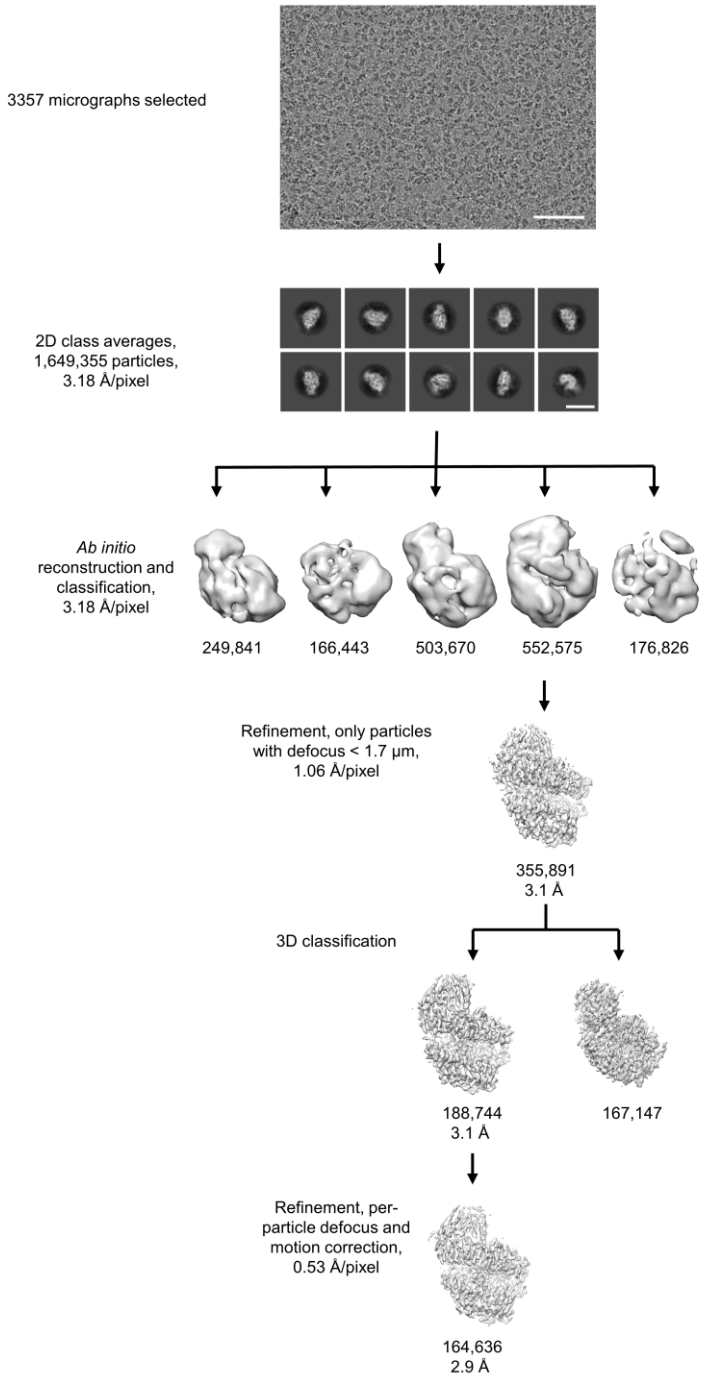
Supplementary information

SARS-CoV-2 variant prediction and antiviral drug design are enabled by RBD in vitro evolution

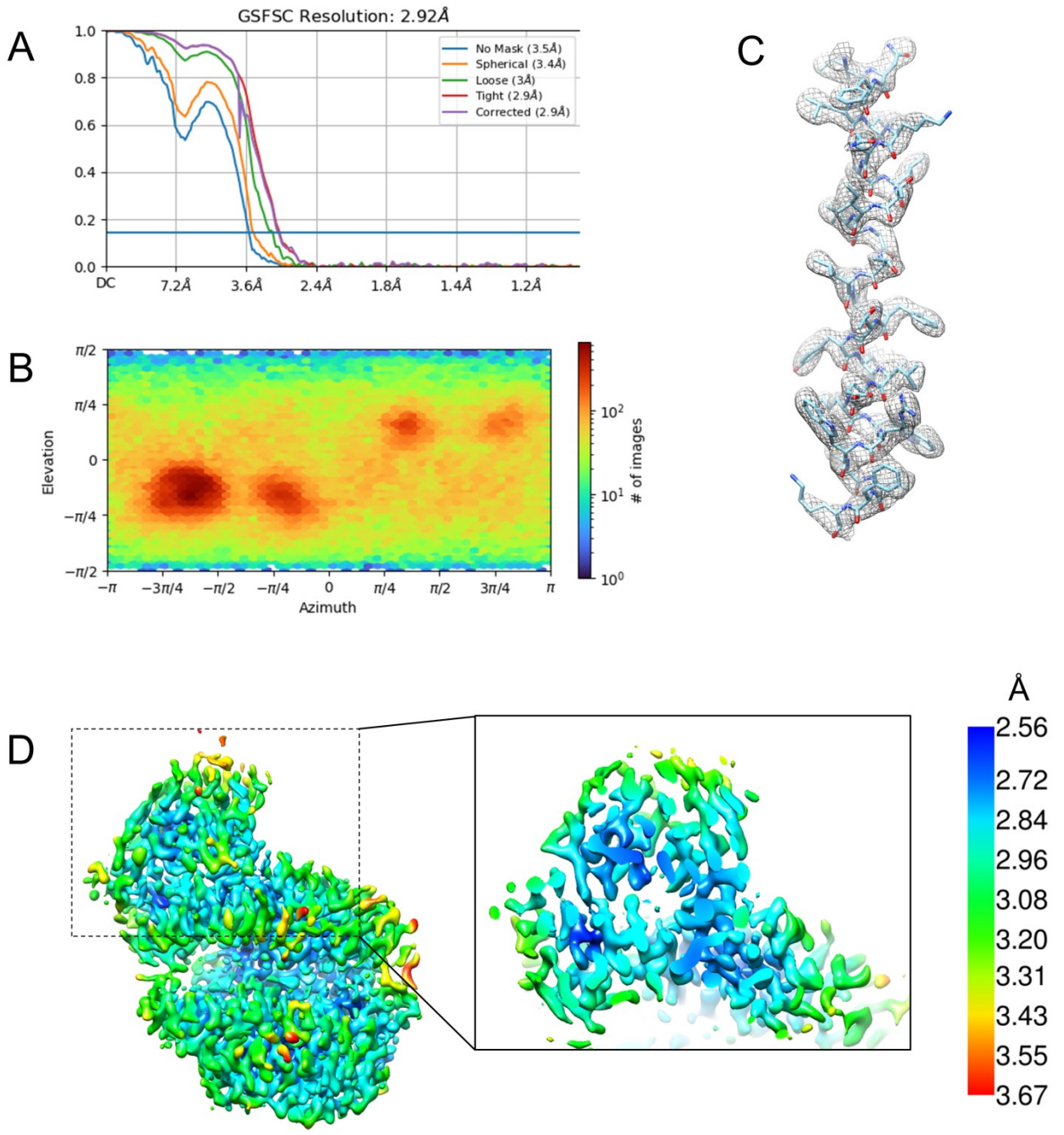
In the format provided by the authors and unedited



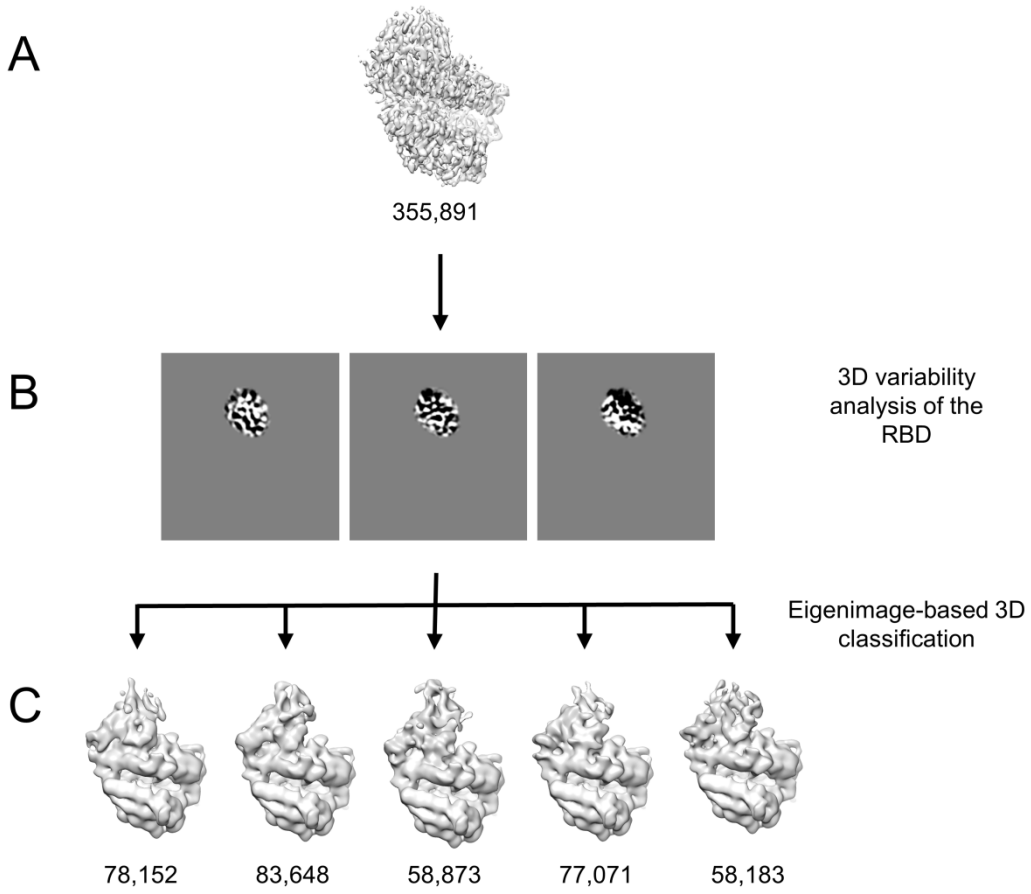
Supplementary Fig. 1. Protein purification of ACE2, RBD-62 and the complex between the two. Both proteins were expressed in Expi293F cells and secreted to cell culture media. (A) SDS-PAGE analysis after NiNTA agarose purification of the ACE2 receptor extracellular domain (AA Q18 – S740). M – Molecular weigth marker [PM2700] ExcelBand™ 3-color Broad Range (5-245 kDa, Mw corresponding to given band is shown (kDa)). (B) SDS-PAGE analysis from NiNTA agarose purification of RBD-62 (AA 333-528). (C) The ACE2 + RBD-62 complex was purified by gel filtration chromatography prior to CryoEM. The ACE2 protein was mixed with an excess of RBD-62 (1:1.5), incubated 1 h on ice, and applied on the chromatography column by using an ÄKTA pure FPLC system. The first peak corresponds to the complex (SDS-gel inset) and the second peak represents excess RBD-62.



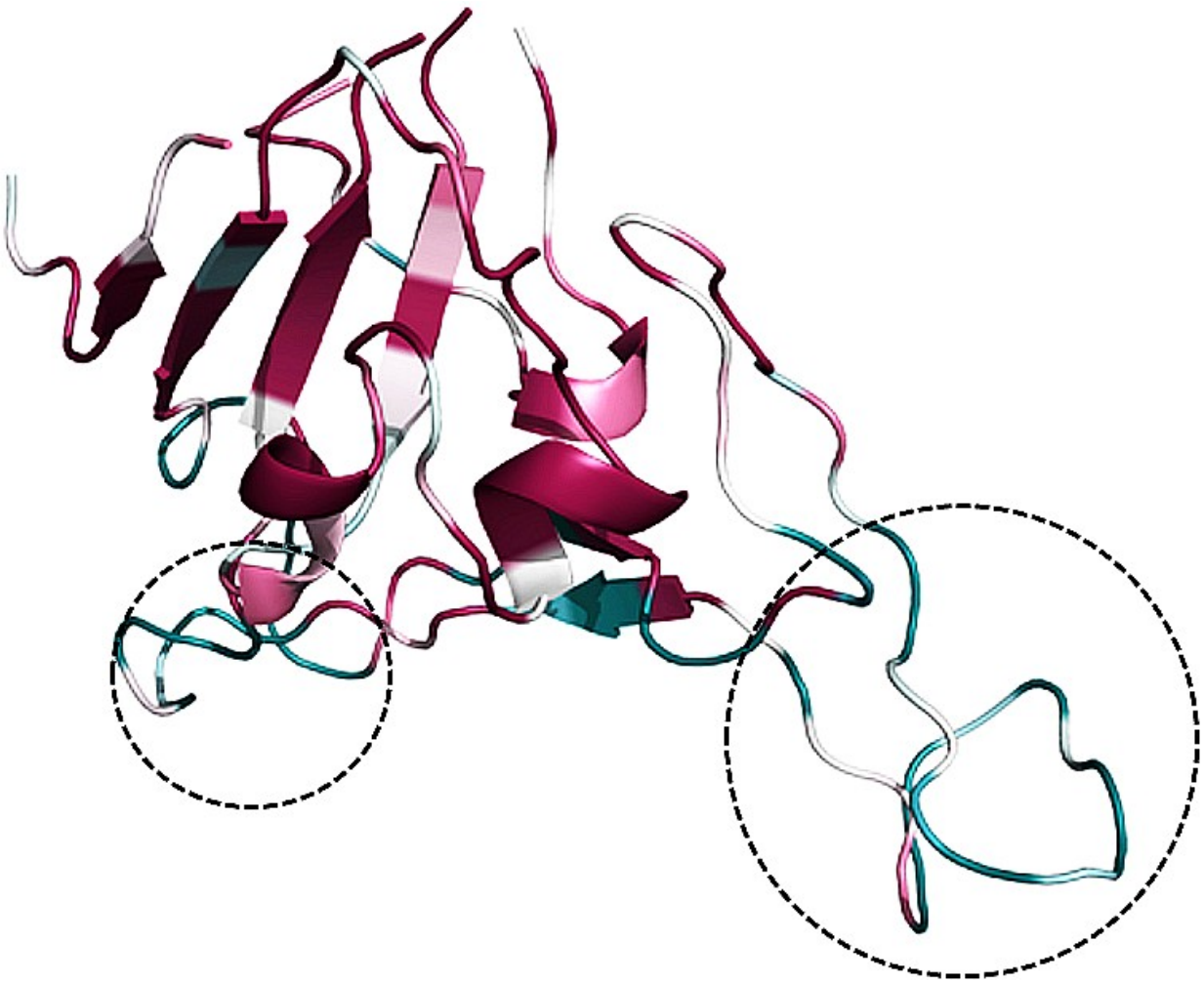
Supplementary Fig. 2. Single-particle Cryo-EM processing scheme. The details of the process are described in the Methods section under "Cryo-EM image processing". The number of particles in each map is indicated under the map's image, along with the map's resolution where relevant. The scale bar (white line) in the upper micrograph corresponds to 50 nm and the scale bar in the panel below to 10 nm.



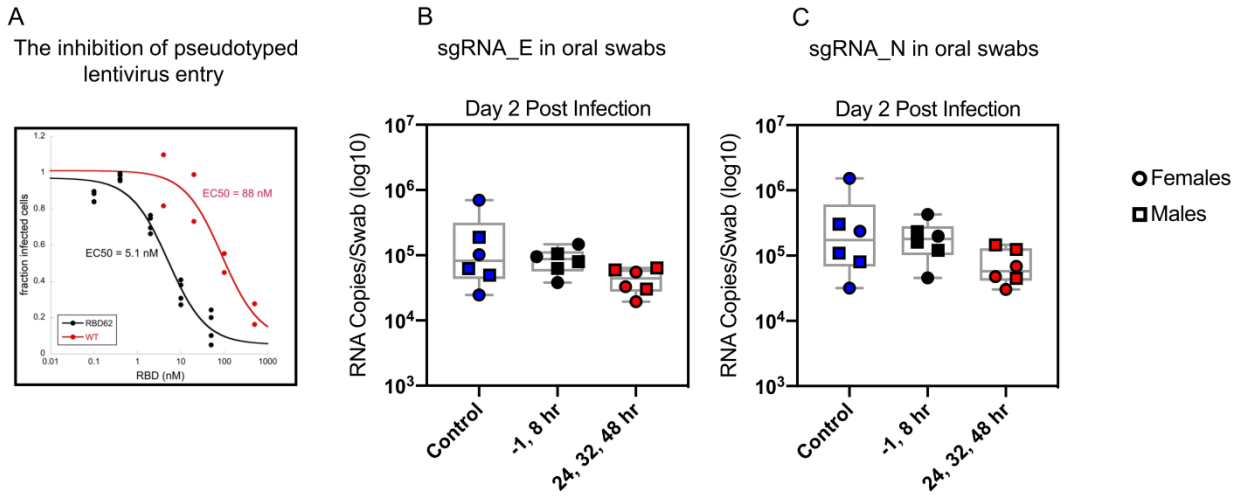
Supplementary Fig. 3. Resolution estimate and angular distribution for the ACE2-RBD-62 Cryo-EM map. (A) Fourier Shell Correlation (FSC) curves. (B) Angular distribution plot. (C) An alpha-helical segment showing the map density and fitted atomic coordinates. (D) Cryo-EM map colored according to local resolution estimate. The inset shows a slice through the RBD-ACE2 interface.



Supplementary Fig. 4. Variability analysis of the RBD. (A) Particle images from the well-resolved 3D class were subjected to 3D variability analysis. (B) Central slices through the three eigenimages calculated with a binary mask around the RBD region. (C) Five 3D classes, which were calculated based on the eigenimages. The maps show variable density for the RBD.



Supplementary Fig. 5. Analysis of conserved positions computed by ConSurf server depicted on the RBD-62 structure. The amino acids are colored by their conservation grades with turquoise-through-maroon indicating variable-through-conserved by the ConSurf server (Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. *Nucleic Acids Res.* **38**, W529–W533 (2010)). .



Supplementary Fig. 6. Inhibition of pseudo-virus entry by RBD-WT and RBD-62. **(**A) The pseudotyped lentivirus displayed the SΔC19 Spike variant in this experiment [28]. (B) The number of sgRNA_E copies identified in oral swabs isolated at the second day post-challenge. (C) The number of sgRNA_N copies identified in oral swabs isolated at the second day post-challenge. The box plots show the minimal, maximal and center bounds (25-75%) results for the six animals in each treatment taken at day 2 post infection.

Supplementary Table 1 – Comparison of different RBD domain spans and their influence on domain properties

Construct	Position[a]	Number of AA	Size [kDa]	Yeast expression [mean FL1*10^3][b]	Yeast display estimated K_D [nM][c] [replicates]	Melting temperature [°C]
RBDcon1	336-516	181	20.5	16.4	3.2 ± 0.1 [3]	
RBDcon2	336-528	193	21.7	32.9	1.7 ± 0.1 [8]	
RBDcon3	333-528	196	22.1	37.9	1.1 ± 0.3 [3]	53.5 ± 0.3
RBDcon4	319-541	223	25.1	67.6	1.2 ± 0.2 [3]	53.8 ± 0.2

[a] numbers are according to UniProtKB- P0DTC2

[b] measured in pJYDC1 (eUnaG2 fluorescence signal)

[c] Binding affinity against ACE2 was determined by FACS, with the relevant construct expressed on yeast surface.

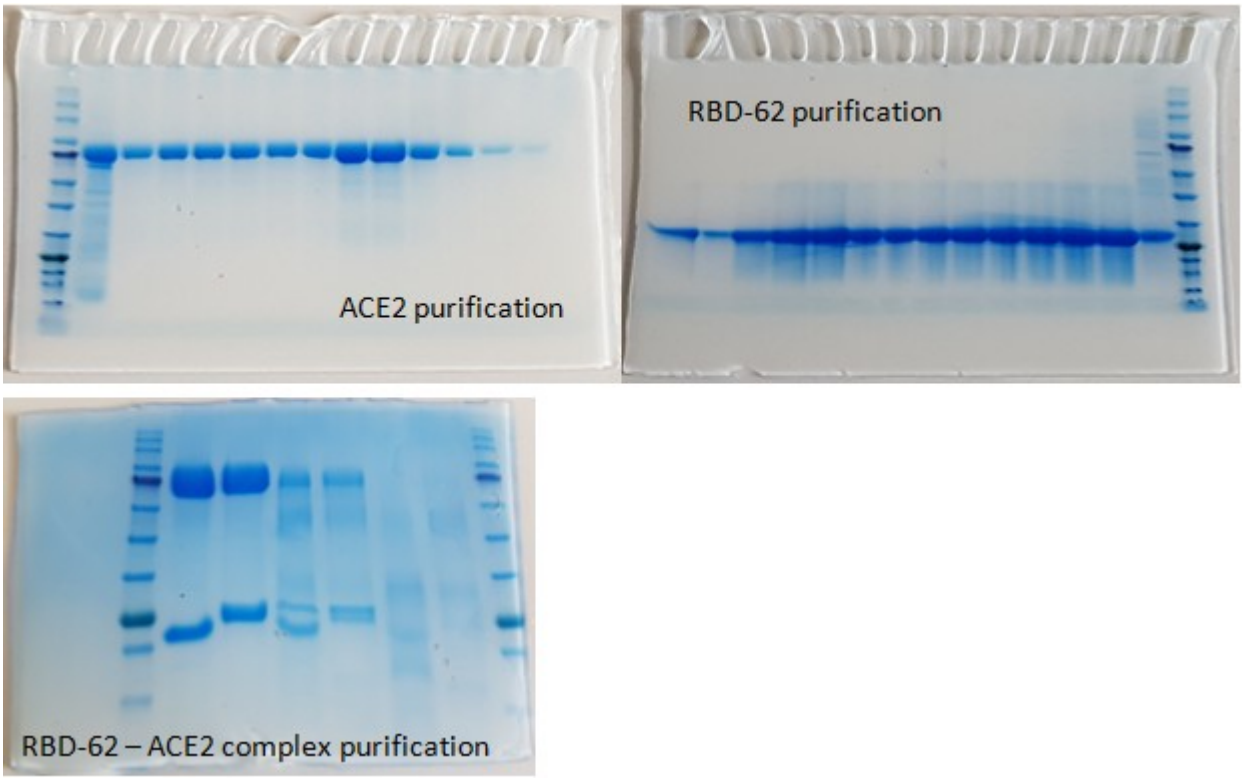
Supplementary Table 2 - Cryo-EM data collection, refinement and validation statistics of the ACE2-RBD62 complex

	RBD-62 ACE2 complex (EMD-12187) (PDB 7BH9)
Data collection and processing	
EM equipment	Titan Krios (Thermo Fisher Scientific)
Detector	K3 (Gatan)
Energy filter	BioQuantum (Gatan), 20 eV slit
Magnification	94,518
Voltage (kV)	300
Electron exposure (e–/Å^2)	70
Defocus range (μm)	0.4 – 2.5
Pixel size (Å)	0.53
Symmetry imposed	C1
Number of collected micrographs	4,470
Number of selected micrographs	2,535
Initial particle images (no.)	2,419,995
Final particle images (no.)	164,636
Map resolution (Å)	2.9
FSC threshold	0.143
Map resolution range (Å)	2.6 – 3.7
Map sharpening B factor (Å2)	-83
Refinement	
Initial model used (PDB code)	6M0J
Model resolution (Å)	2.9
FSC threshold	
Model resolution range (Å)	2.5/2.6/3.0
Map sharpening *B* factor (Å^2)	
Model composition	
Non-hydrogen atoms	5,870
Protein residues	729
Sugar	3
Zn	1
B factors (Å^2)	
Protein	33.27/92.04/52.19
Ligand	57.90/78.27/67.70
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.606
Validation	
MolProbity score	1.42
Clashscore	3.86
Poor rotamers (%)	0.0
Ramachandran plot	
Favored (%)	96.37
Allowed (%)	3.63
Disallowed (%)	0.0

Supplementary Table 3 – Primer list

Sequencing and random mutagenesis primers	
pCTCON_seq_FW	GCAGCCCCATAAACACACAGTAT
GaL1b_F	cctctatactttaacgtcaaggag
Aga2p_F	GCGGTTCTCACCCCTCAACAAC
pCT_seq_R	catgggaaaacatgttgtttacggag
RBD-bind-F	CTACAAACTGCCTGATGACTTCAC
RBD-int_R	GTCCAGGTTGTTGCTGTTCCAG
S-CoV mut_F	GGCTGTGTGATTGCCTGGAAC
S-CoV_mut_R	GCACCACCACCCTGTATGGTTG
RBD domain span testing	
S-CoV_Adap_F	TGCCCTTTTGGTGAAGTTTTTAACGCCACCAGGTTTGCCTCTGTC
S-CoV_Adap_R	TTCAAAAGAAAGTACTACTACTCTGTATGGTTGGTAGCCCACTCCATTGGTTGG
S-CoV_pJYDC_F	ggcggtagcggaggcggagggtcggctagccatTGCCCTTTTGGTGAAGTTTTTAACG
S-CoV_pJYDC_R2	ctattacaagtcctcttcagaaataagcttttgttcggatccCTTTGGTCCACACACTGTGGC
RBD4_pJYDC_F	ggcggtagcggaggcggagggtcggctagccatAGGGTCCAACCAACAGAGAG
RBD4_pJYDC_R	ctattacaagtcctcttcagaaataagcttttgttcggatccGAAGTTCACACACTTGTTCTTCACC
RBD3_pJYDC_F	ggcggtagcggaggcggagggtcggctagccatCCAAACATCACCAACCTGTGTC
RBD3_pJYDC_R	ctattacaagtcctcttcagaaataagcttttgttcggatccCTTTGGTCCACACACTGTGGC
Kr_RBDmut_F	CCATGTTCGTGTTTCTGGTGCTGCTGCCTCTGGTGTCCAGCACCCCAAACATCACCAACCTGTGCCCTTTTGGTGAAGTTTTTAACG
Kr_RBDmut_R	CGACTTAAGATCGATGCGGCCGCGAGCTCGAATTTTATCAATGGTGATGGTGATGGTGCTTTGGTCCACACACTGTGGC
RBD-WT mutagenesis	
RBD_S477N_R	GCCCTCCACTCCATTACATGGTGTGTTGCCAGCCTGGTAAATCTCTG
RBD_N501Y_R	GTATGGTTGGTAGCCCACTCCATAGGTTGGTTGGAAGCCATAGGATTG
RBD_N484K_R	GCAGCACACCATGTAATGGAGTGAAGGGCTTCAACTGTTACTTTCCACTCC
RBD_S494P_R	CCATTGGTTGGTTGGAAGCCATAGGGTTGGAGTGGAAAGTAACAGTTGAAGCC
RBD_V445K_F	GATTGCCTGGAACAGCAACAACCTGGACAGCAAGAAGGGAGGCAACTACAACTACCTCTACAG
RBD_N460K_F	GCAACTACAACTACCTCTACAGACTGTTCAGGAAGAGCAAACTGAAACCATTTGAGAGGGACATC
RBD_Q498R_F	CTGTTACTTTCCACTCCAATCCTATGGCTTCCGACCAACCAATGGAGTGGGCTAC
RBD_T470M_F	GCAACCTGAAACCATTTGAGAGGGACATCAGCATGGAGATTTACCAGGCTGGCAGC
JZ_RBD_E484R_F	CCAGGCTGGCAGCACACCATGTAATGGAGTGAGAGGCTTCAACTGTTACTTTCCACTCC
RBD_I468T_F	GGAAGAGCAACCTGAAACCATTTGAGAGGGACACCAGCACAGAGATTTACCAGGCTG
RBD_K417T_F	GAGACAGATTGCCCCTGGACAAACAGGCACTATTGCTGACTACAACTACAAACTGC
RBD_K417N_F	GAGACAGATTGCCCCTGGACAAACAGGCAATATTGCTGACTACAACTACAAACTGC
RBD_E484K_F	CCAGGCTGGCAGCACACCATGTAATGGAGTGAAAGGCTTCAACTGTTACTTTCCACTCC
RBD_N439K_F	GACTTCACAGGCTGTGTGATTGCCTGGAACAGCAAGAACCTGGACAGCAAGGTGGG
RBD_S477R_F	GGACATCAGCACAGAGATTTACCAGGCTGGCAGAACACCATGTAATGGAGTGGAGG
RBD_I358F_F	GGTTTGCCTCTGTCTATGCCTGGAACAGGAAGAGGATTAGCAACTGTGTGGCTGACTAC
RBD_T478K_F	GGACATCAGCACAGAGATTTACCAGGCTGGCAGCAAACCATGTAATGGAGTGGAGGGC
RBD_T478I_F	GGACATCAGCACAGAGATTTACCAGGCTGGCAGCATTCCATGTAATGGAGTGGAGGGC
RBD_G446V_F	GCCTGGAACAGCAACAACCTGGACAGCAAGGTGGTTGGCAACTACAACTACCTCTACAGAC
RBD_A475V_F	CCATTTGAGAGGGACATCAGCACAGAGATTTACCAGGTTGGCAGCACACCATGTAATGG
RBD_N501T_F	CCACTCCAATCCTATGGCTTCCAACCAACCACTGGAGTGGGCTACCAACCATAC

RBD-62 mutagenesis	
B62_S694P_F	GTGAAGGGCTTCAACTGTTACTTTCCACTCCAACCTTATGGCTTCCGACCAACCTATG
RBD_N354E_F	CGCCACCAGGTTTGCCTCTGTCTATGCCTGGGAAAGGAAGAGGTTTAGCAACTGTGTG
RBD_Y365W_F	CAGGAAGAGGTTTAGCAACTGTGTGGCTGACTGGTCTGTGCTCTACAACTCTGCC
RBD_V367W_F	GAAGAGGTTTAGCAACTGTGTGGCTGACTACTCTTGGCTCTACAACTCTGCCTCCTTCAG
RBD_T385R_F	CTTCAGCACCTTCAAGTGTTATGGAGTGAGCCCACGTAAACTGAATGACCTGTGTTTCACCAATG
RBD_R408D_F	CTGACTCCTTTGTGATTAGGGGAGATGAGGTGGATCAGATTGCCCCTGGACAAACAG
RBD_Q414A_F	GGGAGATGAGGTGAGACAGATTGCCCCTGGAGCCACAGGCAAGATTGCTGACTACAAC
RBD_K417V_F	GGTGAGACAGATTGCCCCTGGACAAACAGGCGTTATTGCTGACTACAACTACAAACTGCC
RBD_Y453F_F	GACAGCAAGGTGGGAGGCAACTACAACTACCTCTTTAGACTGTTCAGGAAGAGCAAGC
RBD_L455R_F	GGTGGGAGGCAACTACAACTACCTCTACAGACGTTTCAGGAAGAGCAAGCTGAAACC
RBD_Q493M_F	GAGTGAAGGGCTTCAACTGTTACTTCCCACTCATGTCCTATGGCTTCCGACCAACC
RBD_E,K484R_F	CCAGGCTGGCAGCACACCATGTAATGGAGTGCGCGGCTTCAACTGTTACTTCCCACTC
RBD_Q,R498H_F	CTGTTACTTCCCACTCCAATCCTATGGCTTCCATCCAACCTATGGAGTGGGCTACC
RBD_N,Y501F_F	CTTCCCACTCCAATCCTATGGCTTCCGACCAACCTTTGGAGTGGGCTACCAACCATAC
RBD_Y505W_F	CCTATGGCTTCCGACCAACCTATGGAGTGGGCTGGCAACCATACAGGGTGGTGGTGC
RBD_L517M_F	CCAACCATACAGGGTGGTGGTGCTGTCCTTTGAAATGCTCCATGCCCCTGCCACAGTG
B62_L455R_F	GCAAGAAGGGAGGCAACTACAACTACCTCTACAGACGTTTCAGGAAGAGCAAACTGAAACCATTTG
RBD cloning Inko pCAGGS	
RBD_con2-kr_F	TTTCTGGTGCTGCTGCCTCTGGTGTCCAGCACCAACCTGTGTCCATTTGGA
RBD_con2-kr_R	CGAATTTTATCAATGGTGATGGTGATGGTGCTTTGGTCCACACACTGTGG
RBD_seq-kr_F	TTTCCTACAGCTCCTGGGCAAC
RBD_seq-kr_R	AGCTATGACTGGGAGTAGTC



Source Data Supplementary Fig. 1. Uncropped scans of SDS-PAGE gels