iScience, Volume *24*

Supplemental information

Structural insight reveals SARS-CoV-2 ORF7a as an immunomodulating factor for human $CD14^+$ monocytes

Ziliang Zhou, Chunliu Huang, Zhechong Zhou, Zhaoxia Huang, Lili Su, Sisi Kang, Xiaoxue Chen, Qiuyue Chen, Suhua He, Xia Rong, Fei Xiao, Jun Chen, and Shoudeng Chen

Supplemental Information

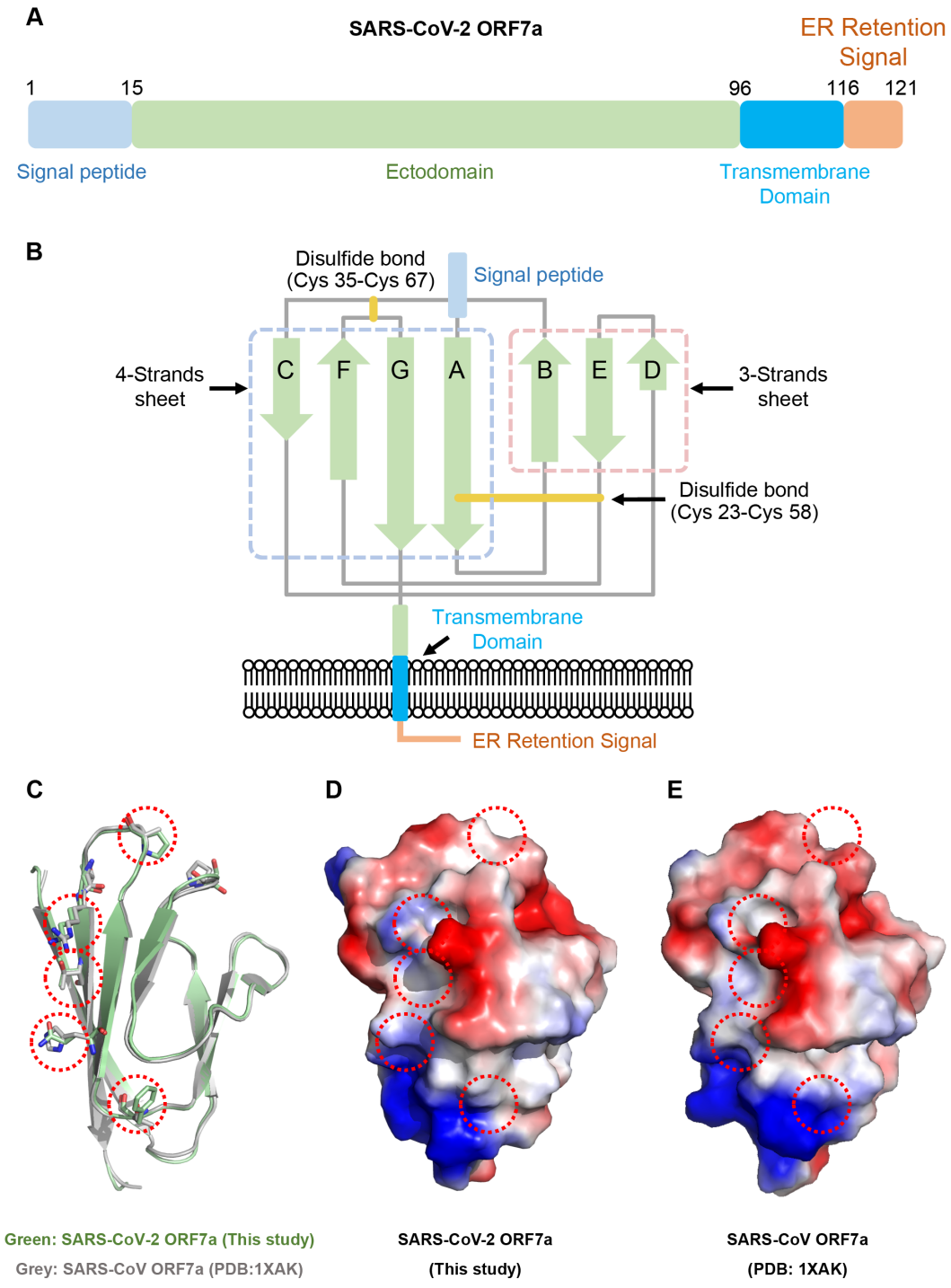

Figure S1. Features of secondary structure and electrostatic potential in SARS-CoV-2 ORF7a. Related to Figures 1 and 3.

(**A**) Domain architecture of SARS-CoV-2 ORF7a. The numbers denote the residue sites. (**B**) Topological illustration of the SARS-CoV-2 ORF7a structure. (**C**) Overall structural comparison of the residue variations in SARS-CoV-2 ORF7a and SARS-CoV ORF7a. (**D**–**E**) Different electrostatic potentials on the SARS-CoV-2 ORF7a surface (**D**) and the SARS-CoV ORF7a surface (**E**). Red dash circles indicate the varied residues.

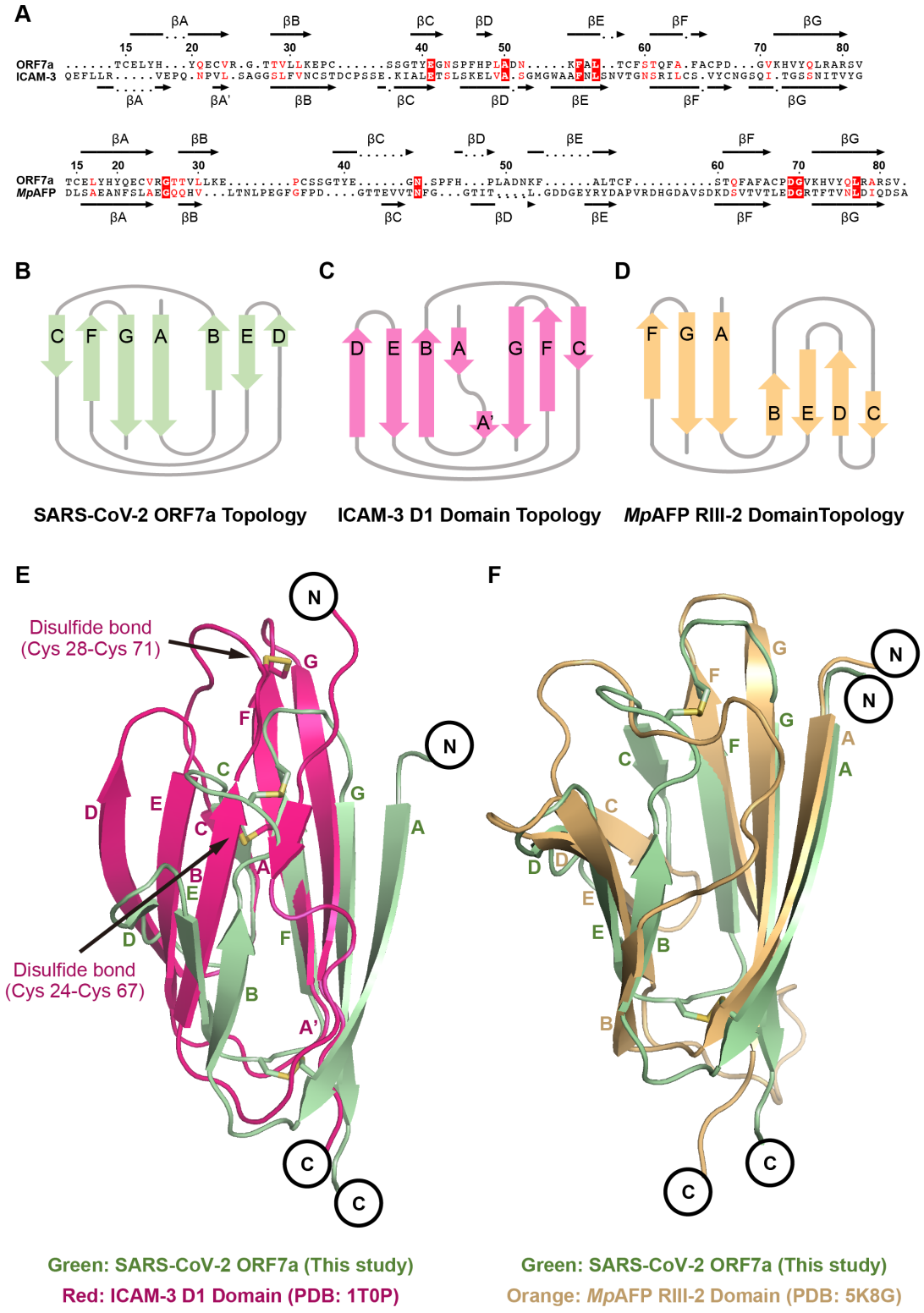

Figure S2. Comparison of SARS-CoV-2 ORF7a with structurally similar proteins. Related to Figure 4.

(**A**) Sequence alignments of SARS-CoV-2 ORF7a (UniProt: A0A6B9VSS2) with the ICAM3 D1 domain (UniProt: P32942) and *Mp*AFP RIII-2 domain (UniProt: A1YIY3). The arrows with labels denote the β-strand regions of the structure. (**B**–**C**) Topological diagrams of the structures of SARS-CoV-2 ORF7a (**B**), the ICAM3 D1 domain (**C**), and the *Mp*AFP RIII-2 domain (**D**). (**E**) Structural comparison of the SARS-CoV-2 ORF7a and ICAM3 D1 domain. (**F**) Structural comparison of the SARS-CoV-2 ORF7a and *Mp*AFP RIII-2 domain. Disulfide bonds are represented with sticks.

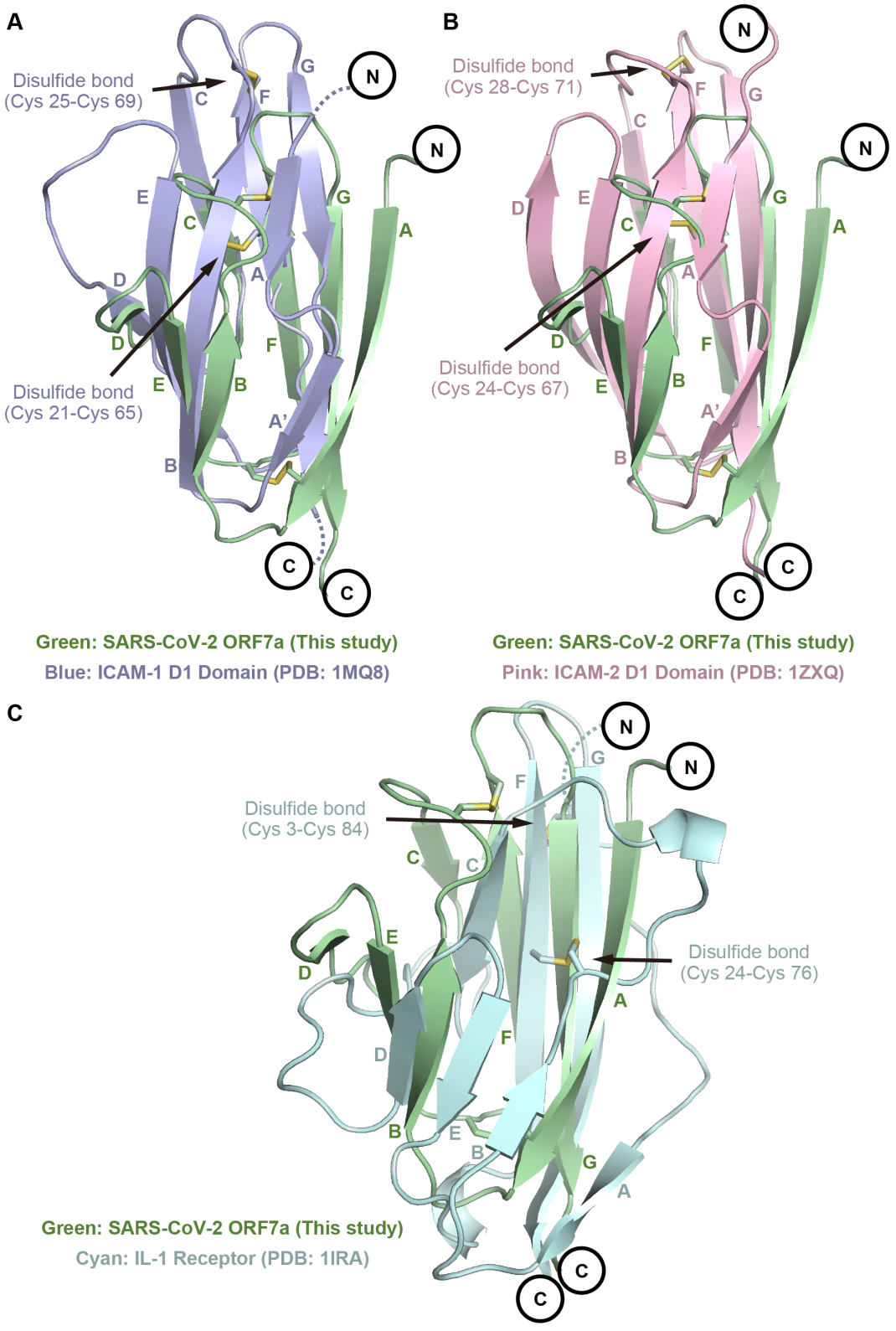

Figure S3. Structural superimpositions of SARS-CoV-2 ORF7a with the correlated Ig-like proteins. Related to Figure 4.

Structural comparisons of SARS-CoV-2 ORF7a and the ICAM1 D1 domain (**A**), ICAM2 D1 domains (**B**), and IL-1 Receptor (**C**), respectively. Disulfide bonds are represented with sticks.

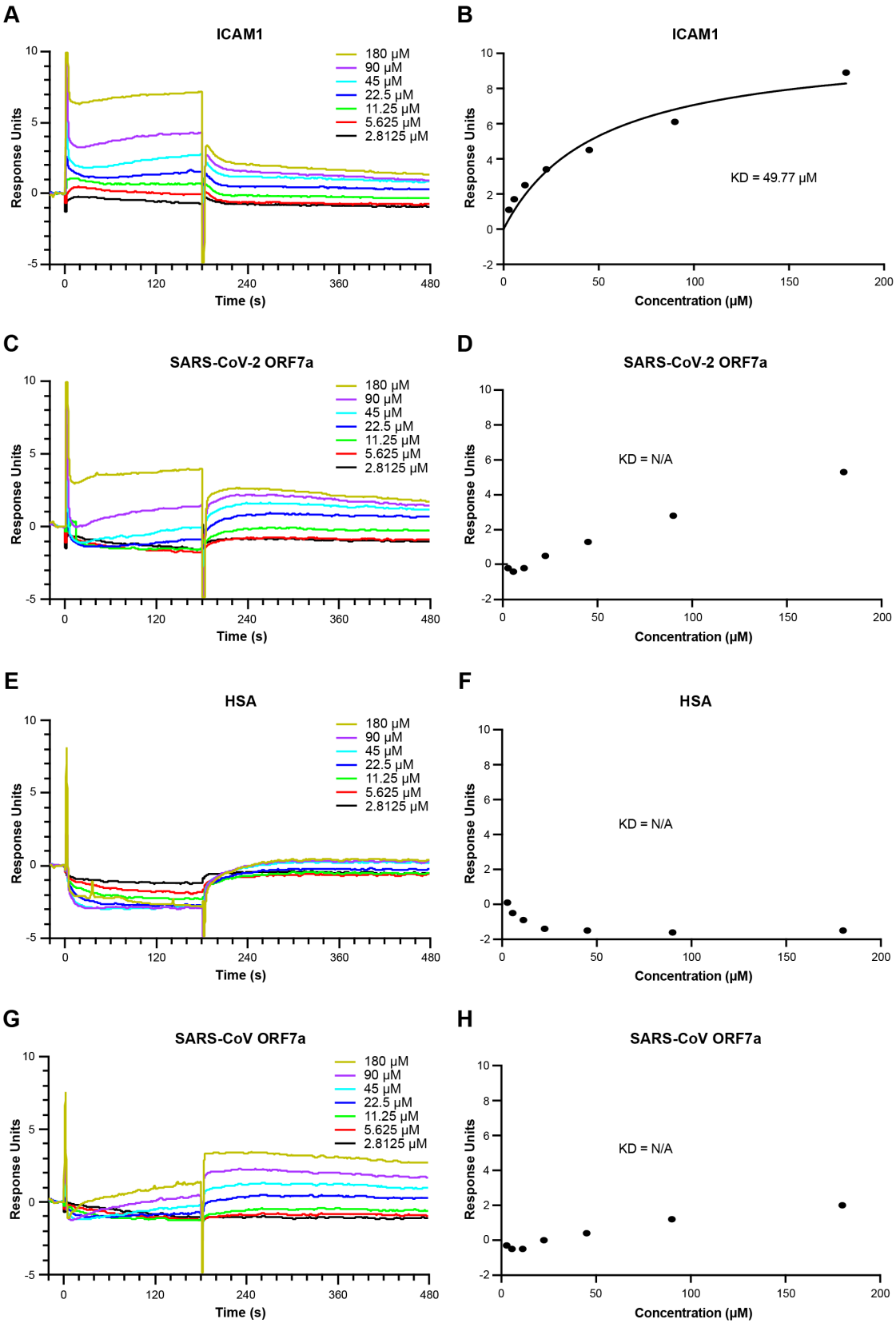
A
ICAM1
180 μM
90 μM
45 μM
22.5 μM
11.25 μM
5.625 μM
2.8125 μM
Response Units
Time (s)
B
ICAM1
KD = 49.77 μM
Response Units
Concentration (μM)
C
SARS-CoV-2 ORF7a
180 μM
90 μM
45 μM
22.5 μM
11.25 μM
5.625 μM
2.8125 μM
Response Units
Time (s)
D
SARS-CoV-2 ORF7a
KD = N/A
Response Units
Concentration (μM)
E
HSA
180 μM
90 μM
45 μM
22.5 μM
11.25 μM
5.625 μM
2.8125 μM
Response Units
Time (s)
F
HSA
KD = N/A
Response Units
Concentration (μM)
G
SARS-CoV ORF7a
180 μM
90 μM
45 μM
22.5 μM
11.25 μM
5.625 μM
2.8125 μM
Response Units
Time (s)
H
SARS-CoV ORF7a
KD = N/A
Response Units
Concentration (μM)

Figure S4. The Surface Plasmon Resonance (SPR) analysis for ICAM1, SARS-CoV-2 ORF7a and SARS-CoV ORF7a with the active (I-domain K287/294C) form of the LFA-1 α_L I-domain. Related to Figure 4.

(**A**) The sensorgrams of different concentrations of the active form of the LFA-1 α_L I-domain binding to ICAM1, showing the baselines, association, and dissociation processes. (**B**) The curve represents the binding affinity of ICAM1 and the active form of the LFA-1 α_L I-domain by matching the sensorgrams to a Langmuir binding rate equation. (**C–H**) The sensorgrams of the binding to SARS-CoV-2 ORF7a, HSA, and SARS-CoV ORF7a, respectively. The binding affinities are not applicable to be determined (N/A).

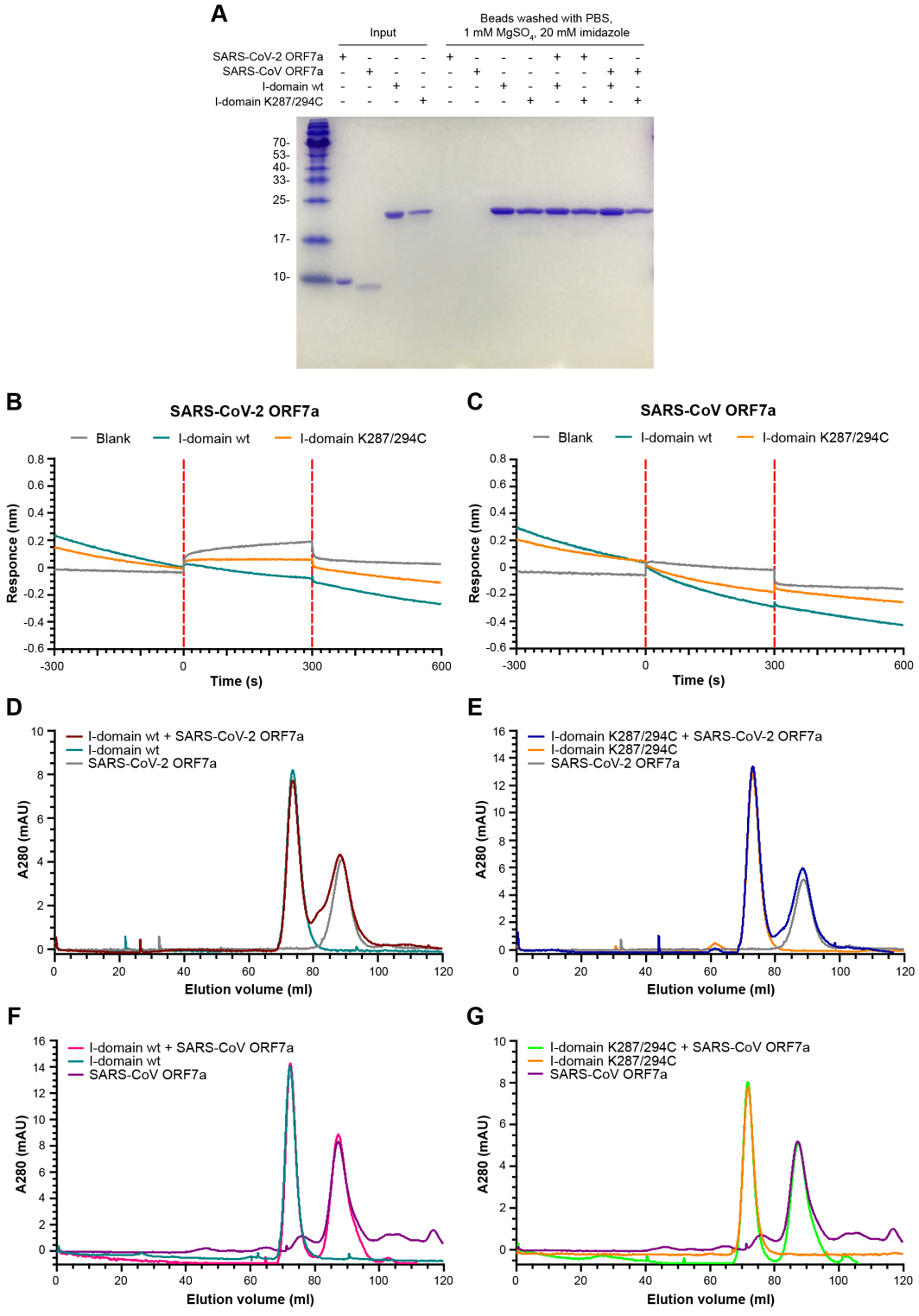
A
Input
Beads washed with PBS, 1 mM MgSO4, 20 mM imidazole
SARS-CoV-2 ORF7a
SARS-CoV ORF7a
I-domain wt
I-domain K287/294C
70-
53-
40-
33-
25-
17-
10-
B
SARS-CoV-2 ORF7a
Blank
I-domain wt
I-domain K287/294C
Responce (nm)
Time (s)
C
SARS-CoV ORF7a
D
I-domain wt + SARS-CoV-2 ORF7a
A280 (mAU)
Elution volume (ml)
E
I-domain K287/294C + SARS-CoV-2 ORF7a
F
I-domain wt + SARS-CoV ORF7a
G
I-domain K287/294C + SARS-CoV ORF7a

Figure S5. *In vitro* protein-protein interaction assays for SARS-CoV-2 ORF7a (9.4 kDa) and SARS-CoV ORF7a (9.4 kDa) with the inactive (I-domain wt) (21.3 kDa) or active (I-domain K287/294C) (21.3 kDa) form of the LFA-1 α_L I-domain. Related to Figure 4.

(**A**) SDS-PAGE analysis of proteins pulled down using Ni-NTA affinity beads. (**B**–**C**) BLI experimental plots showing the baselines, association, and dissociation processes. (**D**–**G**) SEC comparisons to analyze the complex formation between SARS-CoV-2 ORF7a or SARS-CoV ORF7a and the inactive or active form of the LFA-1 α_L I-domain.

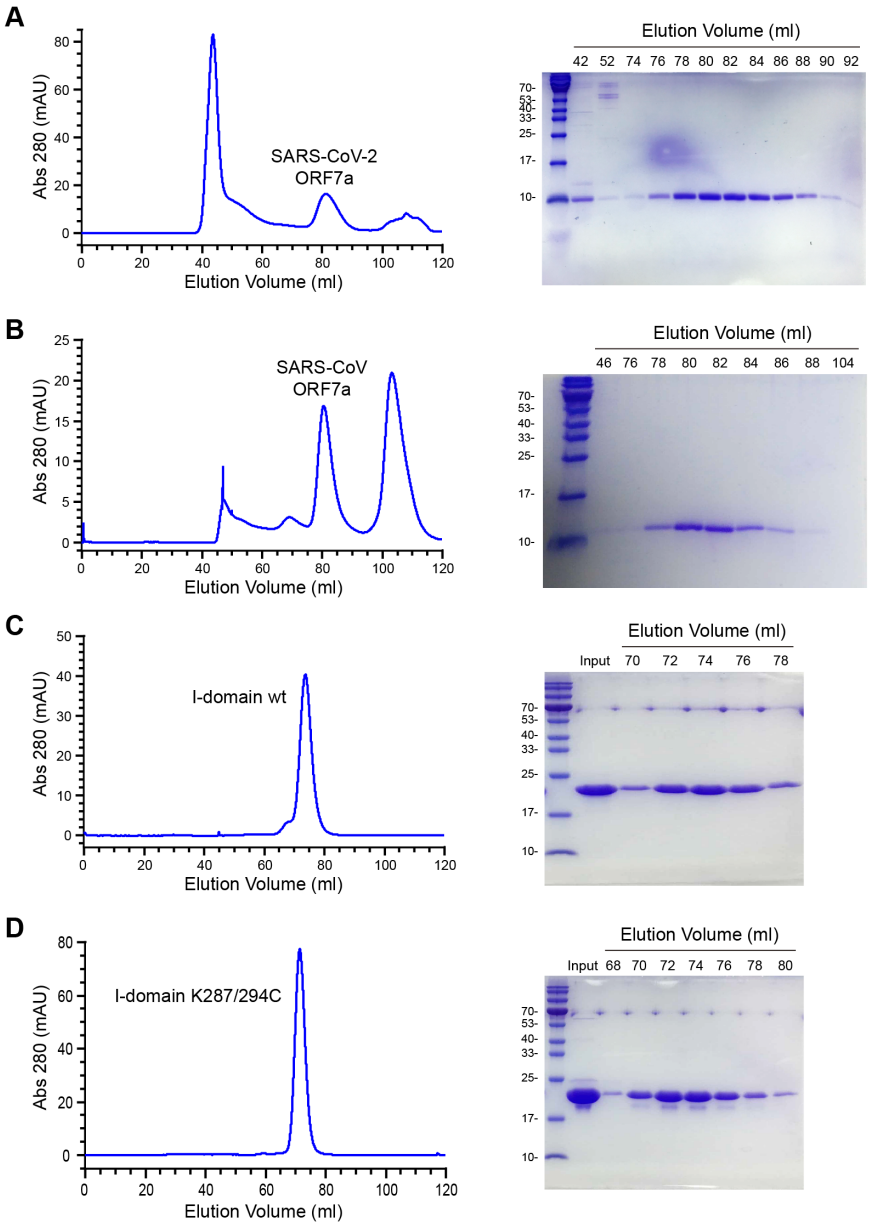

Figure S6. Protein purifications using SEC method. Related to Figure 4.

The SEC chromatography plots of SARS-CoV-2 ORF7a (**A**), SARS-CoV ORF7a (**B**), the inactive (**C**) and active (**D**) form of the LFA-1 α_L I-domain with the corresponding SDS-PAGE analysis. The soluble proteins were subjected to SEC on a HiLoad 16/600 Superdex 75 pg column.

Tables

Table S1. Data collection and structure refinement statistics. Related to Figure 2.

	SARS-CoV-2 Orf7a (PDB: 7CI3)
Data collection	
Space group	*P* 3_1
Cell dimensions	
a, *b*, *c* (Å)	37.55, 37.55, 56.16
α, *β*, *γ* (°)	90.00, 90.00, 120.00
Resolution (Å)	50.00-2.20 (2.24-2.20)*
R_{merge}**	0.103 (0.629)
$I / \sigma I$	15.48 (1.29)
Completeness (%)	99.70 (96.20)
Redundancy	4.5 (2.8)
Refinement	
Resolution (Å)	28.14-2.20 (2.28-2.20)
No. reflections	4462 (427)
R_{work}# / R_{free}##	0.201 / 0.246
No. atoms	549
Protein	542
Ligand/ion	0
Water	7
B-factors	61.34
Protein	61.35
Ligand/ion	0
Water	60.38
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.14
Ramachandran favored (%)	94.03
Ramachandran allowed (%)	5.97
Ramachandran outliers (%)	0

*Statistics for the highest-resolution shell are shown in parentheses.

**Rmerge = $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the *i* th reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl.

#R-work = $\sum_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \sum_{hkl} |F_{obs}(hkl)|$.

##R-free is calculated in an identical manner using 10% of randomly selected reflections that were not included in the refinement.

Table S2. DALI searching results. Related to Figure 4.

NO.	Chain	Z-score	RMSD	Identity (%)	Description
1	1YO4-A	10.9	1.3	87	Molecule: Hypothetical Protein X4
2	5K8G-A	6.2	2.5	11	Molecule: Antifreeze Protein
3	1T0P-B	6.1	2.5	8	Molecule: Integrin Alpha-L
4	5H7V-A	6.0	2.2	10	Molecule: Kunitz-type Protease
5	2YMO-A	6.0	2.9	11	Molecule: PF12
6	6FX6-A	5.9	3.0	11	Molecule: Satie-ted
7	6GYB-B	5.9	2.4	13	Molecule: VIRB7
8	5FM5-P	5.9	2.4	8	Molecule: Myomesin-1
9	LZ7Z-I	5.9	2.4	14	Molecule: Human Coxsackievirus A21
10	4BFI-A	5.7	2.5	10	Molecule: Cell Surface Glycoprotein CD200 Receptor 1
11	5HP5-A	5.4	2.3	13	Molecule: Protein-arginine Deiminase Type-1
12	2Z4T-A	5.3	2.6	8	Molecule: Beta-galactoside Alpha-2,6-sialyltransperase
13	1BQS-A	5.3	2.2	13	Molecule: Protein (Mucosal Addressin Cell Adhesion)
14	3KVQ-A	5.3	2.1	5	Molecule: Vascular Endothelial Growth Factor Receptor 2
15	6R0X-E	5.3	3.0	11	Molecule: Antibody Fab Fragment Heavy Chain
16	6NZS-A	5.3	2.2	8	Molecule: Dextranase
17	3ZMR-A	5.3	2.5	5	Molecule: Cellulase (Glycosyl Hydrolase Family 5)
18	6EY4-A	5.1	2.4	8	Molecule: GLDM
19	4WIQ-A	5.1	3.0	9	Molecule: Dystroglycan
20	3O3U-N	5.1	2.2	16	Molecule: Maltose-binding Periplasmic Protein, Advanced GLY

Table S3. Primers used in qRT-PCR assays. Related to Figure 5.

Primer	Primer sequence (5'-3')
hGAPDH-F	TTGCCCTCAACGACCACTTT
hGAPDH-R	CCACCACCCTGTTGCTGTAG
hIFNa-F	TTCAAAGACTCTCACCCCTGC
hIFNa-R	ACAGTGTAAAGGTGCACATGA
hIFNb-F	ACGCCGCATTGACCATCTAT
hIFNb-R	GTCTCATTCCAGCCAGTGCTA
hIL8-F	ACTCCAAACCTTTCCACCCC
hIL8-R	TTCTCAGCCCTCTTCAAAAACT
hIL6-F	CCACCGGGAACGAAAGAGAA
hIL6-R	CGAAGGCGCTTGTGGAGAA
hTNFa-F	CCCATGTTGTAGCAAACCCTC
hTNFa-R	TATCTCTCAGCTCCACGCCA
hIL1b-F	AGAAGTACCTGAGCTCGCCA
hIL1b-R	CTGGAAGGAGCACTTCATCTGT

Transparent Methods

Cell experiments

For cell binding assays, the purified recombinant SARS-CoV-2 ORF7a (residues 14–96), SARS-CoV ORF7a (residues 14–96) and human serum albumin (HSA) proteins were labeled with the green fluorescent dye Alexa Fluor 488 NHS Ester (AF488, Invitrogen). Human peripheral blood mononuclear cells (PBMCs) from the peripheral blood of healthy donors were purified by Ficoll. After incubating with or without protein-AF488 in 6-well plates for 2 hours, the PBMCs were incubated with human FcR blockers (422302, BioLegend) for 20 min on ice before staining with phycoerythrin (PE) anti-human CD14 antibody (63D3, BioLegend), PE anti-human CD19 antibody (4G7, BioLegend), allophycocyanin (APC) anti-human CD3 antibody (HIT3a, BioLegend), PE anti-human CD4 antibody (A161A1, BioLegend) or PE anti-human CD8 antibody (HIT8a, BioLegend) on ice for 30 min. The cells were washed and detected by CytoFLEX (Beckman).

For the immune cell function assays, the PBMCs were incubated with or without purified SARS-CoV-2 ORF7a protein in a 12-well plate for 24 hours, and then, the PBMCs were incubated with FcR blockers for 20 min on ice before staining with PE anti-human CD14 antibody (63D3, BioLegend), PE anti-human CD3 antibody (HIT3a, BioLegend), fluorescein isothiocyanate (FITC) anti-human CD69 antibody (FN50, BioLegend), FITC anti-human HLA-DR/DP/DQ antibody (39, BioLegend), FITC anti-human HLA-A/B/C antibody (W6/32, BioLegend) or FITC anti-human PD-1 antibody (EH12.2H7, BioLegend) on ice for 30 min. The cells were washed and detected with a CytoFLEX counter (Beckman).

For qRT-PCR, monocytes were isolated by an adhesion assay from PBMCs and then incubated with or without purified SARS-CoV-2 ORF7a protein for 24 hours. RNA was extracted, and target genes were detected with specific primers (**Table S3**) by qRT-PCR assay.

Cloning, expression and purification

The SARS-CoV-2 ORF7a gene-encoding plasmids were kindly provided by Prof. Peihui Wang from Shandong University and Prof. Xi Huang from Sun Yat-sen University. The ectodomain of SARS-CoV-2 ORF7a (residues ranging from 14 to 96) was cloned into a

pET3a expression vector with no tags. The protein was expressed in *E. coli* BL21 (DE3) pLysS cells and induced with 0.1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG, Sigma-Aldrich, St. Louis, MO, USA) in Luria-Bertani (LB) medium at 37°C for 3 hours. After the cells were harvested and sonicated, the protein was found in inclusion bodies. The inclusion bodies were dissolved in 6 M guanidinium-HCl; 100 mM Tris-HCl, pH 8.5; and 5 mM DTT at room temperature for 1 hour. The supernatant was rapidly diluted into 1 M arginine; 100 mM Tris-HCl, pH 8.5; 2 mM EDTA; 200 μM PMSF; 5 mM reduced glutathione; and 500 μM oxidized glutathione for 24 hours. The refolded protein was precipitated by adding an equal volume of saturated ammonium sulfate solution and solubilized in 20 mM HEPES pH 7.4 and 25 mM NaCl. The soluble protein was subjected to size exclusive chromatography (SEC) on a HiLoad 16/600 Superdex 75 pg column (GE Healthcare, Little Chalfont, Buckinghamshire, UK) equilibrated in 20 mM Tris-HCl pH 8.5 and 20 mM NaCl (**Figure S6A**). The fragment consisting of residues 14 to 96 of SARS-CoV ORF7a was subjected to codon optimization in *E. coli* and synthesized into a pET9a expression vector, and purified as previously described (**Figure S6B**) (Nelson et al., 2005).

The wild-type LFA-1 α_L I-domain and K287/294C mutant LFA-1 α_L I-domain genes were synthesized. Both genes were cloned into a pET9a vector with a 6×His tag in the C terminus (Sangon Biotech, Shanghai, China). Recombinant LFA-1 α_L I-domain proteins were expressed in *E. coli* BL21 (DE3) pLysS cells and induced by 0.1 mM ITPG at 37°C for 3 hours. The cells were collected and lysed by a homogenizer at 1,200 bar in 50 mM Tris-HCl, pH 8.0; 2 mM DTT; 5 mM benzamidine-HCl; and 2 mM EDTA. After centrifugation at 16,000 g for 30 min, the supernatant was discarded. The pellets were washed with lysis buffer three times and then subsequently washed with pure water. The prepared inclusion body pellets were subsequently dissolved in 6 M guanidine-HCl; 50 mM Tris-HCl, pH 8.0; and 20 mM DTT. The supernatant was diluted to 50 μg/ml with ice-cold refolding buffer (50 mM Tris-HCl, pH 8.5; 1 mM $MgSO_4$; 1 mM DTT; and 5% (v/v) glycerol) and incubated overnight at 4°C, and 1 mM o-phenanthroline and 0.2 mM $CuSO_4$ were added additionally to the refolding buffer when treating the K287/294C mutant I-domain. After filtration and concentration, the solution was dialyzed against refolding buffer overnight. The wild-type I-domain was purified using a HisTrap HP column (GE Healthcare), while the K287/294C mutant I-domain was purified using a HiTrap Q HP column (GE Healthcare). Finally, both the wild-type and K287/294C mutant I-domains were loaded onto a HiLoad 16/600 Superdex 75 pg column (GE Healthcare) equilibrated in PBS with 1 mM $MgSO_4$ added (**Figure S6C** and **S6D**).

Crystallization and data collection

Using the hanging drop vapor diffusion method, crystals were grown at 16°C for 5 days. The protein solution (10 mg/ml) was mixed with an equal volume of reservoir solution (100 mM sodium acetate, pH 5.0 and 10% PEG 3000) and equilibrated against 500 µl of the reservoir solution. The rod-like crystals were harvested in the reservoir solution supplemented with 20% ethylene glycol for 2 s and flash-frozen in liquid nitrogen. The X-ray diffraction data were collected from the beamline BL18U1 of Shanghai Synchrotron Radiation Facility (SSRF).

Structural determination and model refinement

The dataset was processed with HKL3000 (Minor et al., 2006). The 2.2 Å structure of SARS-CoV-2 ORF7a was solved by the molecular replacement method with SARS-CoV ORF7a (PDB: 1XAK) used as the search model with Phaser (McCoy et al., 2007) in the Phenix program suite (Adams et al., 2010). Model building was manually performed using Coot software (Emsley et al., 2010). The statistics of the X-ray diffraction data collection and structure refinement are summarized in **Table S1**.

Surface plasmon resonance (SPR) analysis

Early evidence suggested that replacing Lys287 and Lys294 with Cys of LFA-1 α_L I-domain generated a recombinant protein with high-affinity for ICAM1 (Shimaoka et al., 2003). Using a Biacore 8K with the CM5 sensor chip (GE Healthcare) at 25°C, the SPR experiments were performed. The surface of the CM5 chip was activated by NHS and EDC. ICAM1 (10346-HCCH, Sino Biological, Beijing, China), SARS-CoV-2 ORF7a, HSA, and SARS-CoV ORF7a were respectively immobilized on a channel of the CM5 chip up to about 100 response units. The remaining activated groups were blocked by ethanolamine. The running buffer contained 20 mM HEPES, pH 7.4; 100 mM NaCl; and 1 mM $MgCl_2$. The active (I-domain K287/294C) form of the LFA-1 α_L I-domain was used as the analyte and flowed through the channels at a flow rate of 30 µl/min. The analyte was injected for affinity analysis with 180 s contact time and 300 s dissociation time. We tested seven different concentrations of the analyte from 2.8125 µM to 180 µM to explore the dose-dependent affinity. A blank channel was used as a reference to correct the bulk refractive index by being subtracted from the response signals of the protein-immobilized channels. The equilibrium constant (KD) for the analyte binding to the immobilized proteins was determined from the association and dissociation curves of the sensorgrams.

Protein pull-down assay

Fresh Ni Sepharose 6 Fast Flow beads (GE Healthcare) were used to perform pull-down assays of the inactive (I-domain wild-type) or active (I-domain K287/294C) forms of the LFA-1 α_L I-domain and SARS-CoV-2 or SARS-CoV ORF7a. The beads were equilibrated in PBS containing 1 mM $MgSO_4$. Both LFA-1 forms with a 6×His tag were loaded onto the beads and washed three times. Then, SARS-CoV-2 or SARS-CoV ORF7a was added and incubated with the beads for 1 hour at 4°C. To exclude nonspecific binding, 20 mM imidazole was added to the wash buffer. The beads were washed three times and then harvested, followed by SDS-PAGE analysis with Coomassie brilliant blue staining.

Biolayer interferometry (BLI) assay

The BLI experiments were performed on an Octet RED384 System (ForteBio, Bohemia, NY, USA) at 25°C. PBS containing 0.1% BSA and 1 mM $MgSO_4$ was used as the solution buffer. The active and inactive forms of the LFA-1 α_L I-domain with a 6×His tag were loaded on Ni-NTA biosensors (ForteBio) in a 50 μM protein solution. Saturation response levels between 1 nm and 2 nm were achieved for all samples in 10 min. After the baselines were established, the association-dissociation process was executed using 50 μM SARS-CoV-2 or SARS-CoV ORF7a solutions.

Size exclusive chromatography (SEC) analysis

SARS-CoV-2 or SARS-CoV ORF7a and the active or inactive form of the LFA-1 α_L I-domain were mixed at a molar ratio of 1.5:1 and incubated at 4°C for 1 hour. Two milliliters of the mixtures were loaded onto a HiLoad 16/600 Superdex 75 pg column equilibrated in PBS with 1 mM $MgSO_4$. The resulting chromatography plots overlapped with those of the purified SARS-CoV-2 or SARS-CoV ORF7a and both forms of LFA-1 α_L I-domain.

Supplemental References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213-221.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486-501.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658-674.

Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. Acta Crystallogr. D Biol. Crystallogr. *62*, 859-866.

Nelson, C.A., Pekosz, A., Lee, C.A., Diamond, M.S., and Fremont, D.H. (2005). Structure and intracellular targeting of the SARS-coronavirus Orf7a accessory protein. Structure *13*, 75-85.

Shimaoka, M., Xiao, T., Liu, J.H., Yang, Y., Dong, Y., Jun, C.D., McCormack, A., Zhang, R., Joachimiak, A., Takagi, J.*, et al.* (2003). Structures of the alpha L I domain and its complex with ICAM-1 reveal a shape-shifting pathway for integrin regulation. Cell *112*, 99-111.