Supporting Information

Challenges of short substrate analogues as SARS-CoV-2 main protease inhibitors

Sven Ullrich, Vishnu M. Sasi, Mithun C. Mahawaththa, Kasuni B. Ekanayake, Richard Morewood, Josemon George, Laura Shuttleworth, Xiaobai Zhang, Cassidy Whitefield, Gottfried Otting, Colin Jackson, Christoph Nitsche*

Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

* christoph.nitsche@anu.edu.au

Abbreviations

FA	Formic acid
Fmoc	Fluorenylmethyloxycarbonyl
FRET	Foerster resonance electron transfer
H_2O	Water
HRMS	High resolution mass spectrometry
IFD	Induced fit docking
LRMS	Low resolution mass spectrometry
MD	Molecular dynamics
MDS	Molecular dynamics simulation
MeCN	Acetonitrile
MeOH	Methanol
MM	Molecular modelling
M^{pro}	Main protease
PDB	Protein Data Bank
RMSD	Root mean square deviation
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SEA	Simulation event analysis
SID	Simulation interaction diagram
SPPS	Solid phase peptide synthesis
TFA	Trifluoroacetic acid
UV	Ultraviolet

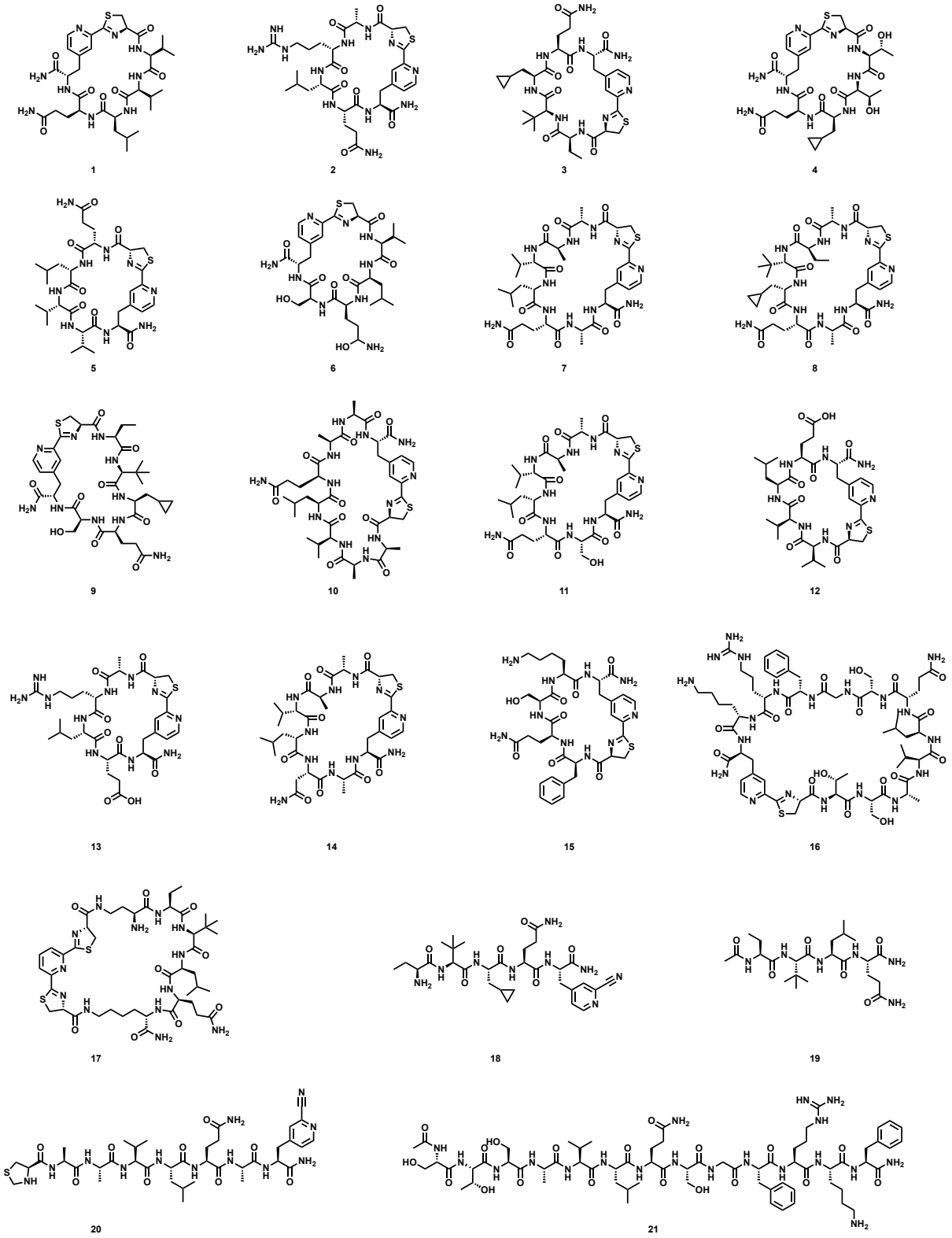

Figure S1. Structures of cyclic, stapled, and linear compounds **1**–**21**.

Instrumentation and materials

Low resolution electrospray ionization mass spectrometry analysis was performed on a Waters LCT Premier orthogonal acceleration time-of-flight mass spectrometer. High resolution electrospray ionization mass spectrometry analysis was performed on a Thermo Scientific Orbitrap Elite mass spectrometer. Analytical liquid chromatography-mass spectrometry was performed on an Agilent HPLC-MS (1260/6120) equipped with a reverse-phase column (Poroshell 120 EC-C_{18}, 2.7 μm, 3.0 x 50 mm) held at 30°C or on a Waters Acquity UPLC-MS system equipped with a reverse phase column (Acquity UPLC BEH-C_{18}, 1.7 μm, 2.1 x 150 mm, 130 Å). A flow rate of 0.3 ml/min was utilized and elution was monitored by UV absorbance. Peptides were purified by preparative HPLC using a Waters 600 controller equipped with a reverse-phase column (SymmetryPrep C18, 100 Å, 7 μm, 19 x 150 mm), autosampler (717 plus), diode array detector (2996), and a Waters Fraction Collector III. Solid phase peptide synthesis was performed using Rink amide resin with a capacity of 0.67 mmol/g (Auspep, Australia) and Fmoc protected amino acids purchased from GL Biochem (China), AK Scientific (USA) or prepared as per reported procedure.[1, 2] Compounds **1**–**15**, **17**–**20** were prepared manually using polypropylene syringes with filter (Torviq, USA). Compounds **16** and **21** were synthesized with the microwave peptide synthesizer Biotage Initiator+ Alstra. The inhibition assay was monitored by a fluorophotometer (Spectramax M2e plate reader, Molecular Devices; Infinite 200 PRO M Plex, Tecan).

Plasmid construction

The gene of SARS-CoV-2 main protease (M^{pro})[3] was cloned in between the *Nde*I and *Xho*I sites of the T7 vector pET-47b (+). The construct contains the M^{pro} autocleavage site (SAVLQ↓SGFRK; arrow indicating the cleavage site) at the N-terminus. At the C-terminus, the construct contains a modified PreScission cleavage site (SGVTFQ↓GP) connected to a His_6-tag. The M^{pro} R298A gene was cloned in between the *Nde*I and *Eco*RI sites of the T7 vector pETMCSI.[4] The R298A mutation is known to prevent the dimerization, which favors the NMR assignment as it increases the sensitivity of NMR experiments.[5] This construct contained an N-terminal His_6-tag followed by a TEV cleavage site. Plasmid constructions and mutagenesis were conducted with a QuikChange protocol using mutant T4 DNA polymerase.[6]

Protein expression

Wildtype and R298A mutant of SARS-CoV-2 M^{pro} were expressed in *E. coli* BL21 DE3 transformed with the desired plasmid. Protein expression was conducted in a bioreactor Labfors 5 (INFORS HT, Switzerland). ^{15}N and ^{13}C labeling was achieved by following the modified protocol developed for the Labfors 5 biorector.[7] Initially, cells were inoculated in 50 mL minimal fermenter media (6.8 g/L KH_2PO_4, 7.1 g/L Na_2HPO_4, 0.71 g/L Na_2SO_4, 2.0 mL/L 1 M $MgCl_2$, 18 g/L glucose, 2.6 g/L $^{15}NH_4Cl$, 0.2 mL/L trace metal mix) and grown overnight at 37 °C in a 220 rpm shaker (10 g/L ^{13}C-labeled glucose was used when expressing ^{15}N and ^{13}C double-labeled protein). The overnight culture was inoculated in 450 mL of ^{15}N minimal fermenter medium in the bioreactor. After the OD reached 12–13, 9 g of glucose and 1.3 g of $^{15}NH_4Cl$ were added and induced with IPTG at a final concentration of 1 mM. After induction, the culture was grown under at 18 °C overnight for protein overexpression. Cells were harvested by centrifugation at 5,000 *g* for 15 minutes and lysed by passing twice through a Emulsiflex-C5 homogenizer (Avestin, Canada). The lysate was centrifuged at 13,000 *g* for 1 h and the filtered supernatant was loaded onto a 5 mL Ni-NTA column (GE Healthcare, USA) equilibrated with binding buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 5% glycerol). The protein was eluted with elution buffer (binding buffer containing, in addition, 300 mM imidazole) and the fractions were analyzed by 12% SDS-PAGE. PreScission cleavage and TEV cleavage were conducted in binding buffer in the presence of 1 mM DTT with a protein-to-protease ratio of 100:1. Following cleavage of the His_6-tag, the buffer was exchanged to NMR buffer (20 mM HEPES-KOH pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA). All samples were analysed by mass spectrometry using an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific, USA) coupled with an UltiMate S4 3000 UHPLC (Thermo Scientific, USA). 7.5 pmol of sample were injected to the mass analyzer via an Agilent ZORBAX SB-C3 Rapid Resolution HT Threaded Column (Agilent, USA).

NMR experiments

All NMR spectra were recorded at 25°C, using 800 and 600 MHz Bruker Avance NMR spectrometers. Samples were prepared in 20 mM HEPES-KOH pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA buffer in 3 mm NMR tubes. 10% D_2O were added to provide a lock signal. 0.1– 0.5 mM protein samples were used for [^{15}N,^{1}H]-HSQC experiments.

General synthesis of 1-21

The Fmoc solid-phase peptide synthesis protocol was performed as described previously.[2] Compounds **1**–**16** were synthesized using the reported sidechain to tail condensation method involving an unnatural amino acid with a 2-cyanopyridine sidechain and N-terminal cysteine in aqueous buffer at neutral pH.[1] Compound **17** was synthesized using the reported peptide stapling method involving 1,2-aminothiol sidechains and 2,6-dicyanopyridine in aqueous buffer at neutral pH.[2] Compounds **18**–**21** were not chemically modified post Fmoc-SPPS.

FRET-based SARS-CoV-2 M^{pro} assays

The FRET-based SARS-CoV-2 M^{pro} inhibition assay was adapted from the literature.[3, 8] Black 96-well polypropylene plates with U-bottom were obtained from Greiner Bio-One, Austria. FRET substrate DABCYL-KTSAVLQ↓SGFRKM-E(EDANS)-NH_2 was obtained from Mimotopes, Australia. The assay buffer consisted of 20 mM Tris-HCl pH 7.3, 100 mM NaCl, 1 mM EDTA, 1 mM DTT. 25 nM of enzyme and 25 μM substrate were used. Compounds **1**–**17** were incubated for 10 min before the addition of FRET substrate. To determine kinetic parameters, an EDANS standard curve was generated as reported previously.[9] Measurements were performed in triplicate. To determine the inhibition constant of **21**, inhibitor concentrations of 10, 16, 26, 44, and 70 μM and substrate concentrations of 10, 20, 35, and 50 μM were used (measurements were performed in duplicate). GraphPad Prism 9.0 was used for data analysis and presentation.

Analytical HPLC-MS method A

This method employed a gradient system of MeOH in H_2O. After injection, the gradient was started with 5% MeOH at 1 min., followed by a gradual increase of MeOH to 90% at 10 min. In both solvents 0.1% TFA was used as an additive.

Analytical UPLC-MS method B

This method employed a gradient system of MeCN in H_2O. After injection, the gradient was started with 5% MeCN at 1 min., followed by a gradual increase of MeCN to 90% at 10 min. In both solvents 0.1% FA was used as an additive.

Preparative HPLC method C

This method employed a gradient system of MeOH in H_2O. After injection, the gradient was started with 5% MeOH, followed by a gradual increase of MeOH to 90% at 10 min. In both solvents 0.1% FA was used as an additive.

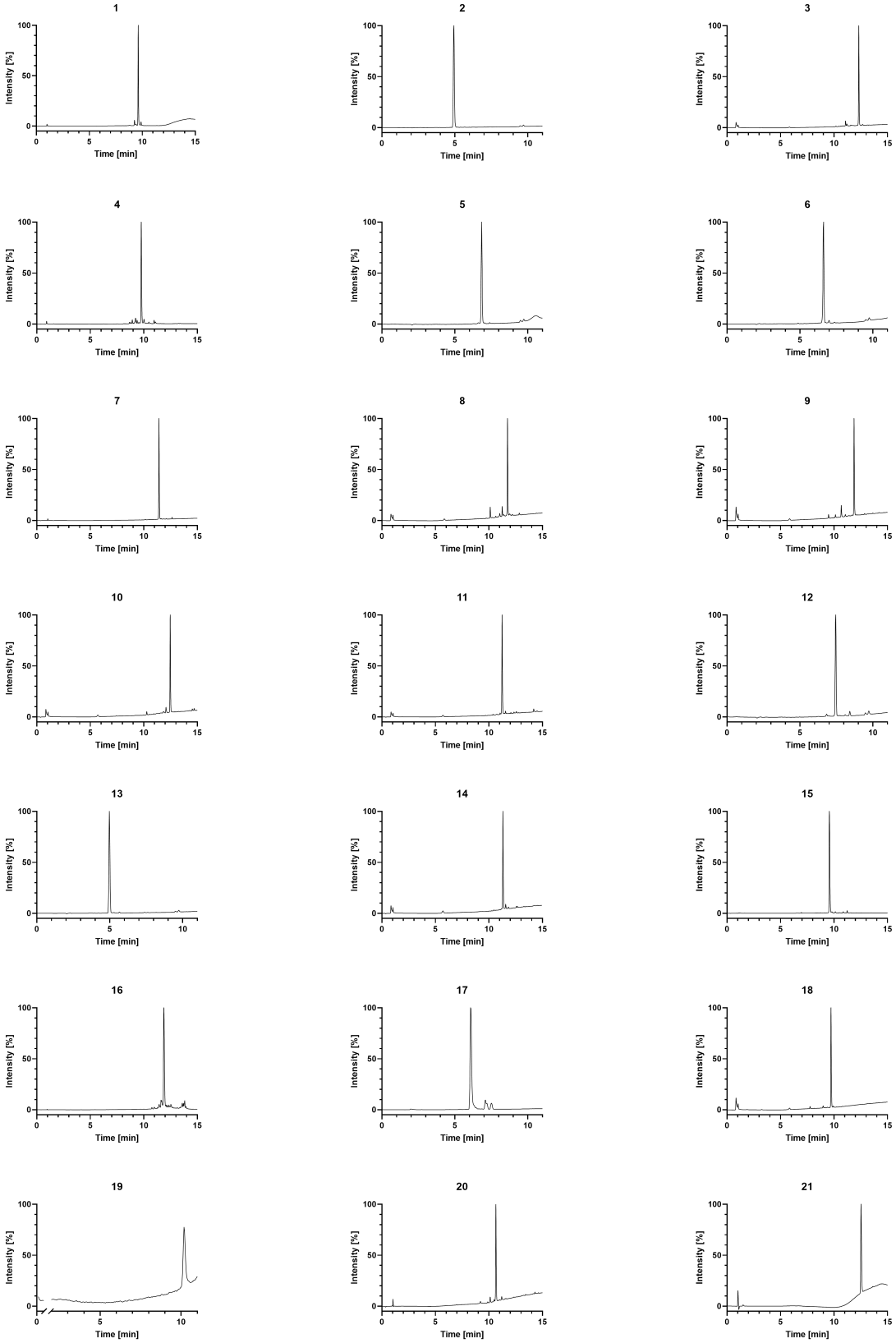

Figure. S2. Chromatograms of **1**–**21**. **1**–**18** and **20**–**21** monitored at 254 nm. Due to the low UV activity of compound **19**, an excerpt of its total ion current chromatogram is shown instead. Method A: **1**, **3**–**4**, **7**–**11**, **14**–**16**, **18**–**20**; method B: **2**, **5**–**6**, **12**–**13**, **17**.

Table S1. Mass spectrometry characterization of compounds **1**–**21**.

Compound	Molecular formula	Ion	Calculated mass	Observed mass
1	$C_{33}H_{49}N_9O_7S$	$[M+H]^+$	716.3554	716.3552
2	$C_{32}H_{48}N_{12}O_7S$	$[M+H]^+$	745.3568	745.3589
3	$C_{33}H_{47}N_9O_7S$	$[M+H]^+$	714.3397	714.3392
4	$C_{31}H_{43}N_9O_9S$	$[M+Na]^+$	740.2802	740.2802
5	$C_{33}H_{49}N_9O_7S$	$[M+Na]^+$	738.3373	738.3370
6	$C_{31}H_{47}N_9O_8S$	$[M+Na]^+$	726.3009	726.3013
7	$C_{37}H_{55}N_{11}O_9S$	$[M+Na]^+$	852.3803	852.3797
8	$C_{39}H_{57}N_{11}O_9S$	$[M+Na]^+$	878.3959	878.3967
9	$C_{36}H_{52}N_{10}O_9S$	$[M+Na]^+$	823.3537	823.3544
10	$C_{40}H_{60}N_{12}O_{10}S$	$[M+H]^+$	901.4354	901.4355
11	$C_{37}H_{55}N_{11}O_{10}S$	$[M+Na]^+$	868.3752	868.3759
12	$C_{33}H_{48}N_8O_8S$	$[M+H]^+$	717.3394	717.3394
13	$C_{32}H_{45}N_{10}O_9S$	$[M+H]^+$	746.3408	746.3405
14	$C_{36}H_{53}N_{11}O_9S$	$[M+Na]^+$	838.3646	838.3649
15	$C_{35}H_{46}N_{10}O_8S$	$[M+H]^+$	767.3299	767.3304
16	$C_{64}H_{98}N_{20}O_{17}S$	$[M+H]^+$	1451.7218	1451.7217
17	$C_{44}H_{67}N_{13}O_9S_2$	$[M+H]^+$	986.4704	986.4706
18	$C_{30}H_{45}N_9O_6$	$[M+H]^+$	628.3571	628.3583
19	$C_{23}H_{42}N_6O_6$	$[M+H]^+$	499.3244	499.3249
20	$C_{38}H_{58}N_{12}O_9S$	$[M+Na]^+$	881.4068	881.4067
21	$C_{66}H_{105}N_{19}O_{19}$	$[M+H]^+$	1468.7912	1468.7911

Molecular modelling studies

Three M^{pro} dimer structures (PDB ID: 6XHU, 6WTM, and 6XQT)[10-12] were chosen for the molecular docking studies.

***In silico* protein and ligand preparation**

The PDB structures were exported to Maestro and processed using the protein preparation wizard (Schrödinger Release 2019-1: Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2019). The initial sequence of steps involved the assignment of bond orders, addition of missing hydrogens, creating disulfide bonds, converting selenomethionines to methionines, removing solvent molecules, and generating het states using the Epik program at pH 7.0 ± 2.0, for each structure, respectively. The final refinement step involved optimization of the H-bond networks using the PROPKA program at pH 7.0 followed by restrained minimization, where the heavy atoms converged to an RMSD of 0.3 Å. The ligand was removed from the processed 6XQT file for the subsequent molecular modelling (MM) studies.

The structures of ligands **1**, **6**, and **20** were drawn using ChemDraw (version 20.0.0.41) and the 3D structure was generated using the Ligprep program in Maestro (Schrödinger Release 2019-1: LigPrep, Schrödinger, LLC, New York, NY, 2019), which involved setting ionization states using Epik at pH 7.0, and selecting options for generating tautomers and stereoisomers, while retaining the original chiral centers and minimizing the structure using OPLS3e forcefield.

Molecular docking

The induced fit docking (IFD) program in Maestro (Schrödinger Release 2019-1: Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2019; Prime, Schrödinger, LLC, New York, NY, 2019) was utilized for docking compound **1** against the four structures, respectively. The ligands **6** and **20** were docked against the 6XQT structure. The standard IFD protocol was followed, applying the OPLS3e forcefield for the docking. The receptor box for each structure was generated by manually setting the centroid of the residues Thr25–Leu27, His41, Ser46, Met49, Tyr54, Phe140–Gly143, Cys145, His163–Pro168, Phe185, Asp187, and Thr190–Gln192. The box size was adjusted to accommodate ligands with length ≤ 25 Å. The ligand ring conformations were sampled while skipping the conformations with energy > 2.5 kcal/mol. Glide docking was performed with the default settings for receptor and ligand van der Waals scaling factor of 0.50 to report a maximum of 20 poses, while the residues within 5 Å distance of the ligand were refined, along with sidechain optimization, using Prime. For

the final Glide redocking step, the standard precision protocol was followed with default settings. The final poses were inspected in PyMOL (PyMOL Molecular Graphics System Version 2.0, Schrödinger, LLC). The binding free energies for the ligands **1**, **6**, and **20** against M^{pro} dimer (PDB: 6XQT) were calculated using the Prime/MM-GBSA method.[13]

Molecular dynamics simulation for the selected structure

The best pose from the docked compound **1**–M^{pro} dimer (PDB ID: 6XQT) complex was selected for molecular dynamics (MD) studies using Desmond (Schrödinger Release 2019-1: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2019) to predict a plausible mechanism of inhibitor binding to the receptor surface.

The system was prepared using the TIP3P explicit solvent model inside an orthorhombic box, using the buffer method at a distance of (25x25x25) Å for the compound **1**–M^{pro} dimer complex. The corresponding counterions were added along with 0.15 M NaCl and the OPLS3e force field was applied to the system. Minimization was performed using the default parameters with a total simulation time of 100 ps. The molecular dynamics simulations were performed in three replicates starting from different random seeds, where the total simulation time for each experiment was set to 300 ns with a time step of 2.0 fs. The default cut-off method was selected to define the short-range Coulombic interactions, with a cut-off radius of 9 Å. The NPT ensemble was selected for carrying out the simulation with an initial relaxation simulation, performed using the NVT ensemble, at a temperature of 300 K and 1.01325 bar pressure. The Nose-Hoover Chain thermostat and Martyna-Tobias-Klein barostat were selected to control the temperature and pressure, respectively, throughout the simulation.

The trajectories for all the replicate simulations for the compound **1**–M^{pro} dimer were analyzed using the Simulation Event Analysis (SEA) module of Desmond, and the corresponding protein and ligand RMSD plots were obtained. The simulation interaction diagram (SID) program was utilized to view the ligand–protein interactions occurring throughout the simulation run time.

Analysis of docking results

The docked conformations generated for each IFD screen were visualized in PyMOL and the representative structures were selected based on the Glide GScore value as well as the binding pose, and interactions observed for the ligand P_1 and P_2 groups in the S_1 and S_2 subsites of M^{pro}, respectively. The M^{pro} monomer structure with bound ligand (PDB ID: 6LU7)[14] was used as the reference to select the best pose from each IFD screen.

Ligand **1** primarily established polar interactions via H-bond formation with the residues Thr25, His41, Leu141, Asn142, and Glu166 as shown in **Figure S3A**. The Gln sidechain of the ligand formed a H-bond with the residues Leu141 and Glu166, respectively. The backbone carbonyl of the Gln side chain amide group formed a H-bond with Asn142. The backbone amide carrying the acetamide sidechain of the ligand formed a H-bond with Asp142 and the acetamide sidechain formed a H-bond with His41. The pyridine group formed a H-bond with Thr25.

Ligand **6** formed H-bonds with the residues Thr26, Leu141, His164, and Gln189 as shown in **Figure S3B**. The Gln sidechain of the ligand forms a H-bond with Leu141 and the backbone carbonyl formed a H-bond with His164. The backbone carbonyl of the Gln side chain amide in the ligand formed a H-bond with Asn142. The side chain of the Gln residue of the ligand was also involved in the formation of an aromatic H-bond with His163. The sidechain hydroxyl group of the Ser residue of the ligand formed a H-bond with Thr26 and the backbone carbonyl of the ligand Leu formed a H-bond with Gln189.

Compound **20** interacted with the residues Asn142, Gly143, His164, Glu166, and Gln189 via H-bonds as shown in **Figure S3C**. The Gln sidechain of the ligand formed a H-bond with Glu166, and the backbone amide shared a H-bond with His164 and Asn142, respectively. The backbone carbonyl of the Leu residue of the ligand formed a H-bond with Asn142. The sidechain amide of the Gln residue of ligand **20** formed an aromatic H-bond each with the sidechain imidazole group of His163 and His172, respectively. The pyridine group formed an aromatic H-bond with the backbone carbonyl of Glu166 and the acetamide sidechain formed a H-bond with Gln189. The backbone carbonyl of the Ala residue of the ligand formed a H-bond with Gly143.

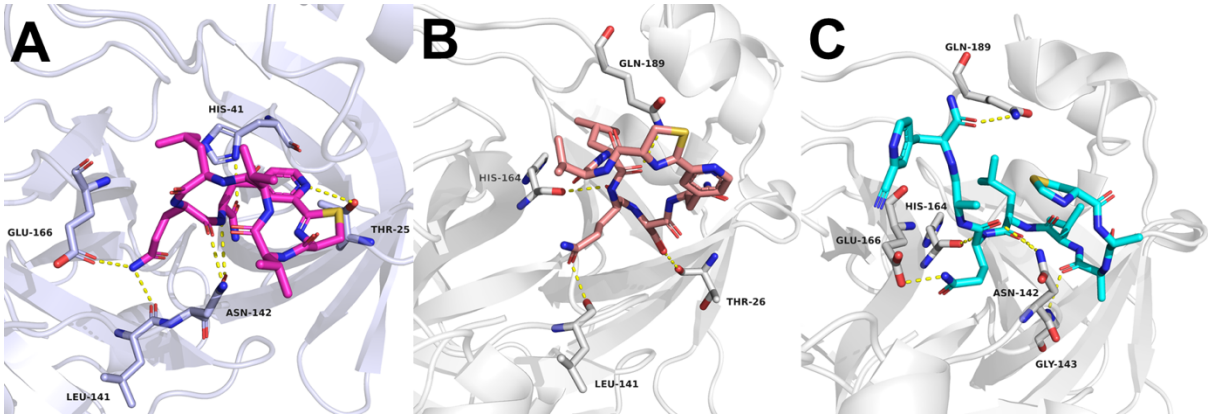

Figure S3. Binding interactions of compounds **1**, **6**, and **20** with the M^{pro} dimer active site. (A) Ligand **1** (pink) interacting with the residues Thr25, His41, Leu141, Asn142, and Glu166 via H-bonds (yellow dashed line). (B) Ligand **6** (salmon) forming H-bonds with Thr26, Leu141, His164, and Gln189. (C) Ligand **20** (cyan) forming H-bonds with residues Asn142, Gly143, His164, Glu166, and Gln189.

Compound 1–M^{pro} dimer MDS studies

The protein backbone RMSD converged rapidly within the first 50 ns with minor fluctuations and remained well within 3 Å until the end of the simulation, as shown in **Figure S4A**. The ligand RMSD fitted to the protein revealed that the ligand RMSD converged rapidly after 50 ns and major conformational changes occurred in the third simulation in the time interval from 248 ns until 256 ns as depicted in **Figure S4B**, indicating ligand diffusion out of the binding pocket. Visualization of the frames in these intervals confirmed that the ligand indeed twice diffused out of the binding pocket to return after 256 ns to occupy the respective sub-pockets, as shown in **Figure S5**, where it stayed until the end of the simulation with minor fluctuations.

The ligand RMSD fitted to the protein for the first and second simulation converged rapidly within 50 ns and the ligand was stabilized after 70 ns to 150 ns, with the P_1 group of the ligand mostly occupying near to the S_1 sub-pocket and major variations in the occupancy of the S_2 sub-pocket by the P_2 Leu residue. The ligand RMSD with respect to the reference ligand pose showed that the ligand stabilized between the interval from 60 ns to 185 ns and from 225 ns to 300 ns as shown in **Figure S4C**.

The representative structures were derived using the trajectory clustering program of Desmond, and SID was used to visualize the prevalent binding interactions between the ligand and the protein.

Schrödinger simulation interactions diagram (SID) analysis of the triplicate simulations (interaction diagrams shown in **Figure S6A-C**) revealed that the ligand Gln sidechain engaged in H-bond interactions with the residues Gly143, Ser144, Cys145 and His164 for more than 30% of the total simulation time in both the first and second simulation, while for the third simulation the Gln group interacted with only Gly143 and Glu166. SID results also revealed that the ligand P_1 group did not form satisfactory interactions (lasting more than 30% of total simulation time) with the major S_1 subsite residues (Phe140, Leu141, Asn142, His163, His172)[14] in all the replicate simulations. The ligand–protein binding site of the representative structure derived from the replicates were visualized in PyMOL, which revealed for the first replicate the Gln amide sidechain of the ligand formed one H-bond with the backbone amide groups of Gly143 and Cys145, and the sidechain hydroxyl group of Ser144, respectively. The backbone amide of the P_1 group was also involved in H-bond formation with Asn142 as shown in **Figure S6D**. Gln189 formed a H-bond with the backbone carbonyl of the Val residue in the

ligand. The ligand acetamide group formed a H-bond with the backbone carbonyl of Thr26 and the imidazole sidechain of His41, respectively.

The representative structure for the second simulation showed the Gln sidechain formed one H-bond with the backbone amide of Cys145, one H-bond with the backbone amide and the sidechain hydroxyl group of Ser144, respectively, and one H-bond with Gly143 as shown in **Figure S6E**. The backbone amide of ligand P_1 group formed one H-bond with Asn142. The acetamide sidechain group of the ligand formed one H-bond with the imidazole sidechain of His41, while the backbone carbonyl of ligand Val was involved in one H-bond with Gln189.

For the third replicate, the representative structure revealed that the ligand Gln amide sidechain formed one H-bond with the backbone amide groups of Gly143 and Cys145, as shown in **Figure S6F**. The ligand acetamide group formed one H-bond with backbone carbonyl of Thr26 and the imidazole sidechain of His41, respectively. The backbone amide of the ligand Val group formed one H-bond with the sidechain hydroxyl group of Ser46.

In all the representative structures, the position of the ligand Leu group was noticeably different from the docked conformation and did not extend well into the S_2 subpocket.

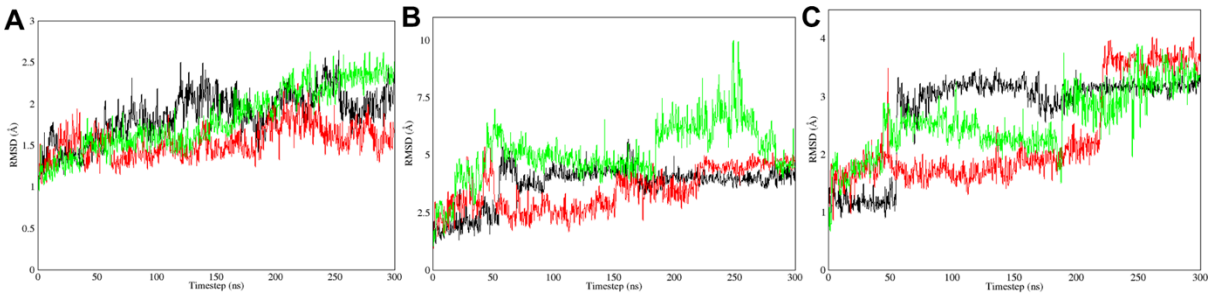

Figure S4. RMSD plots obtained for the triplicate simulations for the protein backbone (A), ligand fitted to protein (B), and ligand fitted to ligand (C). The replicate simulations 1, 2, and 3 are depicted in black, red, and green, respectively.

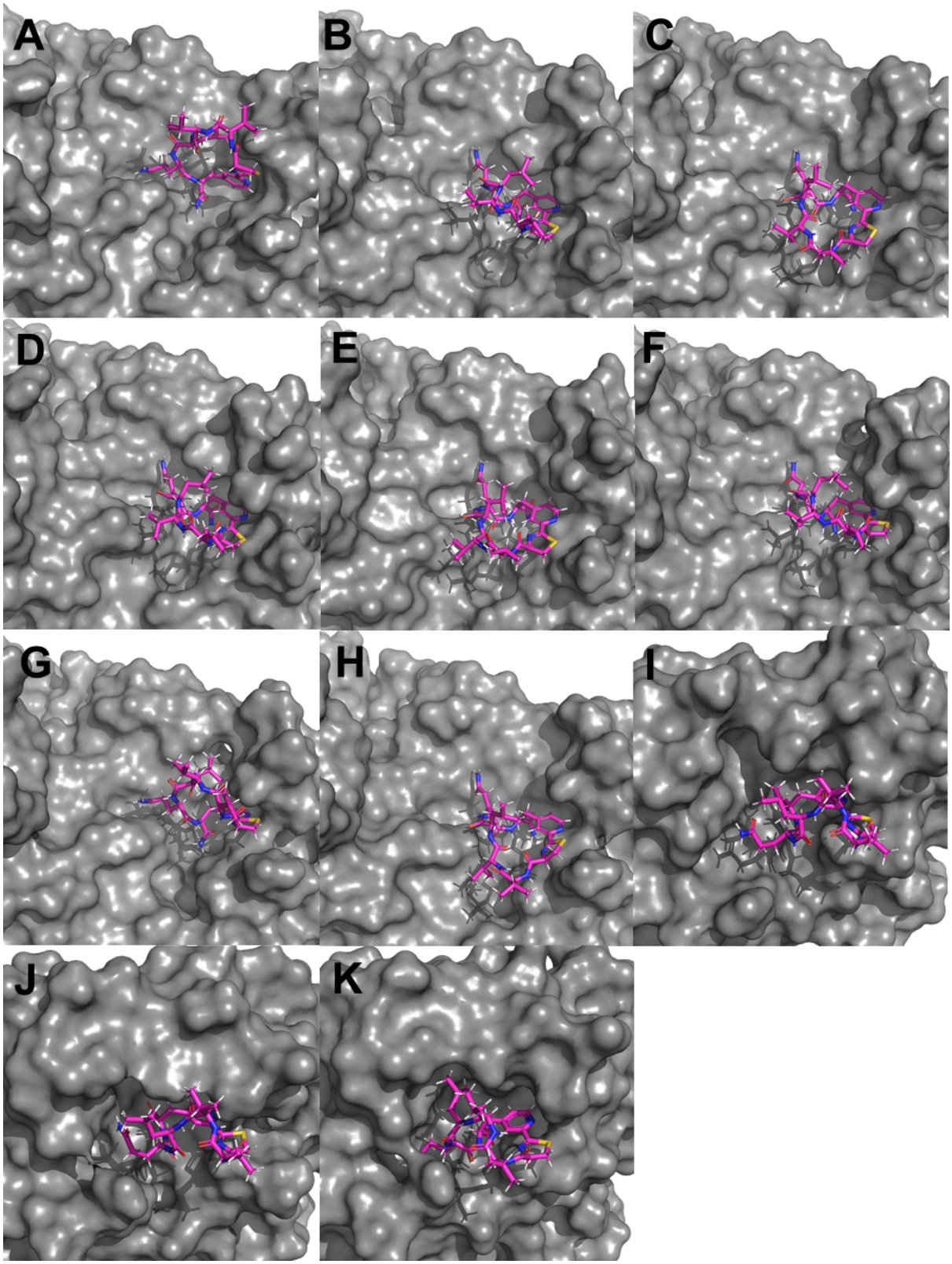

Figure S5. Frames captured between 248 ns and 256 ns revealing the orientation of ligand **1** in the binding site as observed in the third simulation. 0 ns (A), 248.1 ns (B), 248.4 ns (C), 248.7 ns (D), 249.0 ns (E), 249.3 ns (F), 252.6 ns (G), 252.9 ns (H), 253.2 ns (I), 254.4 ns (J), 255.6 ns (K).

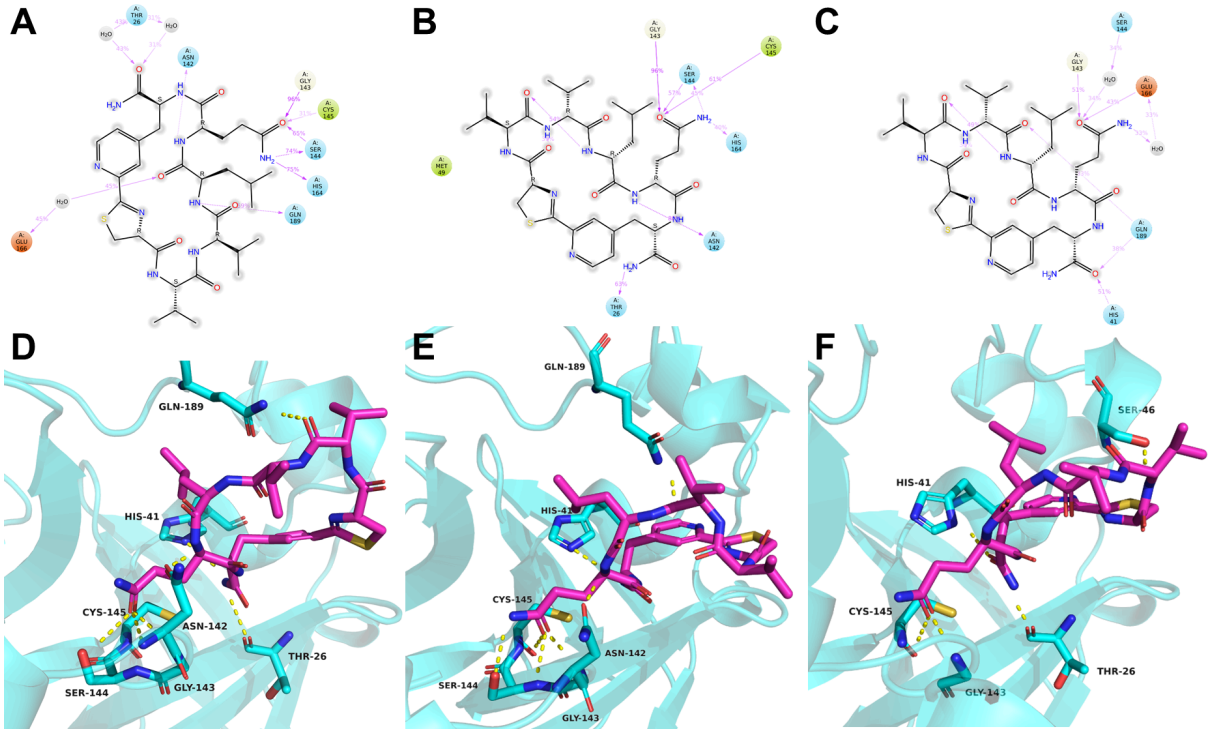

Figure S6. SID generated for the first (A), second (B), and third replicate MDS (C), depicting the ligand–protein interactions occurring for more than 30% of the total simulation time. The polar contacts in the form of H-bonds (yellow-dashed lines) between **1** (pink) and the active site residues, were observed in a representative structure from the first (D), second (E) and third simulation (F), respectively.

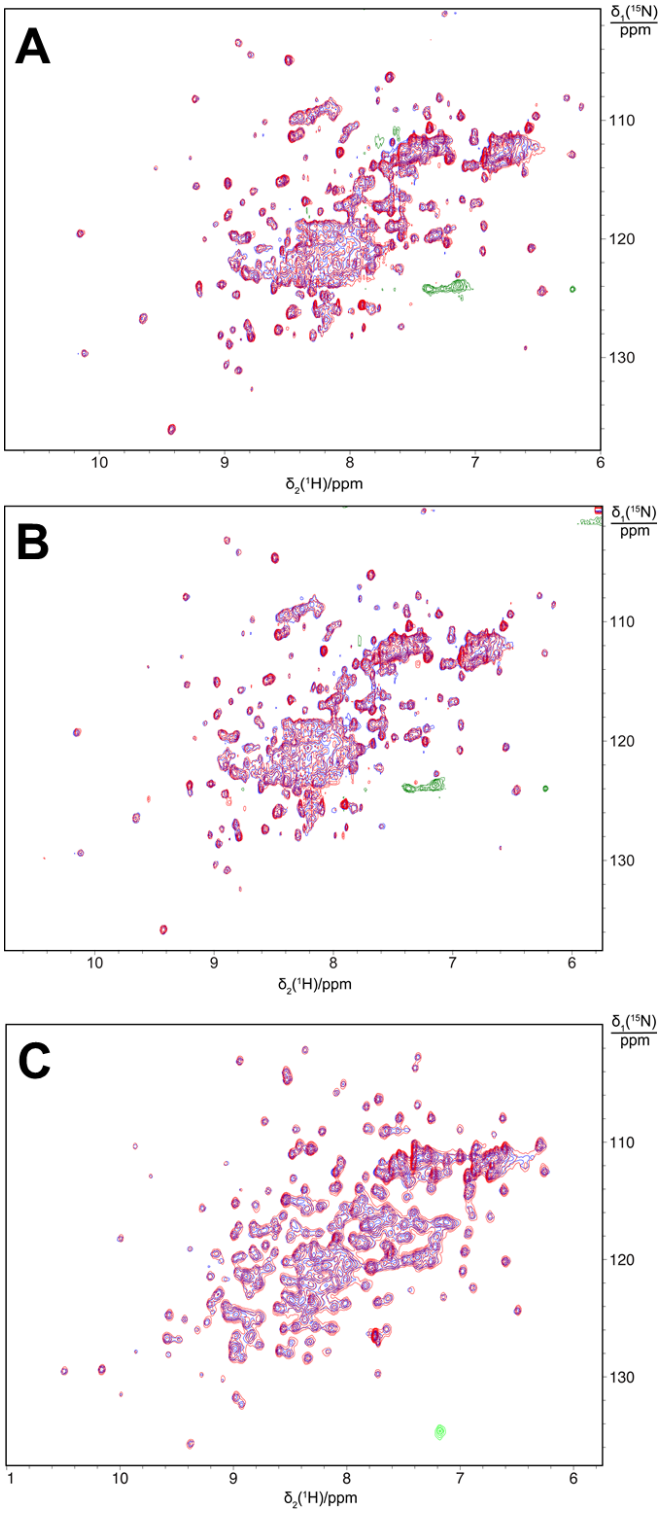

Figure S7. Superimposition of [^{15}N-^{1}H]-HSQC spectra of M^{pro} in the absence (blue spectrum) and presence (red spectrum) of equimolar quantities of compounds **1** and **7**. (A) 0.1 mM solution of ^{15}N/^{2}H-labeled wild-type M^{pro}. (B) Same as (A), but with compound **7**. (C) 0.3 mM solution of ^{15}N/^{13}C-labeled M^{pro} R298A with and without compound **7**. The spectra in the presence of compound superimpose almost fully with the spectra recorded without compound, indicating that the compounds fail to bind and therefore fail to cause spectral changes.

References

1. Nitsche C, Onagi H, Quek J-P, Otting G, Luo D, Huber T. Biocompatible macrocyclization between cysteine and 2-cyanopyridine generates stable peptide inhibitors. *Org Lett.* 2019;21(12):4709-4712. https://doi.org/10.1021/acs.orglett.9b01545.
2. Morewood R, Nitsche C. A biocompatible stapling reaction for in situ generation of constrained peptides. *Chem Sci.* 2021;12(2):669-674. https://doi.org/10.1039/d0sc05125j.
3. Zhang L, Lin D, Sun X, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science.* 2020;368(6489):409-412. https://doi.org/10.1126/science.abb3405.
4. Neylon C, Brown SE, Kralicek AV, Miles CS, Love CA, Dixon NE. Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry.* 2000;39(39):11989-11999. https://doi.org/10.1021/bi001174w.
5. Shi J, Sivaraman J, Song J. Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease. *J Virol.* 2008;82(9):4620-4629. https://doi.org/10.1128/jvi.02680-07.
6. Kalendar R, Qi R, Otting G. Mutant T4 DNA polymerase for easy cloning and mutagenesis. *Plos One.* 2019;14(1):e0211065. https://doi.org/10.1371/journal.pone.0211065.
7. Klopp J, Winterhalter A, Gébleux R, Scherer-Becker D, Ostermeier C, Gossert AD. Cost-effective large-scale expression of proteins for NMR studies. *J Biomol NMR.* 2018;71(4):247-262. https://doi.org/10.1007/s10858-018-0179-0.
8. Zhu L, George S, Schmidt MF, Al-Gharabli SI, Rademann J, Hilgenfeld R. Peptide aldehyde inhibitors challenge the substrate specificity of the SARS-coronavirus main protease. *Antiviral Res.* 2011;92(2):204-212. https://doi.org/10.1016/j.antiviral.2011.08.001.
9. Ma C, Sacco MD, Hurst B, et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* 2020;30(8):678-692. https://doi.org/10.1038/s41422-020-0356-z.

10. Kneller DW, Galanie S, Phillips G, O'Neill HM, Coates L, Kovalevsky A. Malleability of the SARS-CoV-2 3CL M^{pro} active-site cavity facilitates binding of clinical antivirals. *Structure.* 2020;28(12):1313-1320. https://doi.org/10.1016/j.str.2020.10.007.
11. Kneller DW, Phillips G, O'Neill HM, et al. Room-temperature X-ray crystallography reveals the oxidation and reactivity of cysteine residues in SARS-CoV-2 3CL M^{pro}: insights into enzyme mechanism and drug design. *IUCrJ.* 2020;7(6). https://doi.org/10.1107/S2052252520012634.
12. Vuong W, Khan MB, Fischer C, et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat Commun.* 2020;11(1):4282. https://doi.org/10.1038/s41467-020-18096-2.
13. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins.* 2011;79(10):2794-2812. https://doi.org/10.1002/prot.23106.
14. Jin Z, Du X, Xu Y, et al. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature.* 2020;582(7811):289-293. https://doi.org/10.1038/s41586-020-2223-y.