Supporting information

Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease

Chunlong Ma,[1] Michael D. Sacco,[2] Brett Hurst,[3,4] Julia A. Townsend,[5] Yanmei Hu,[1] Tommy Szeto,[1] Xiujun Zhang,[2] Bart Tarbet, [3,4] Michael T. Marty,[5] Yu Chen,[2,*] Jun Wang[1,*]

[1]Department of Pharmacology and Toxicology, College of Pharmacy, The University of Arizona, Tucson, AZ, 85721, United States

[2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, 33612, United States

[3]Institute for Antiviral Research, Utah State University, Logan, UT, 84322, USA

[4]Department of Animal, Dairy and Veterinary Sciences, Utah State University, Logan, UT, 84322, USA

[5]Department of Chemistry and Biochemistry, The University of Arizona, Tucson, AZ, 85721, United States

*Corresponding author:

Jun Wang, Tel: 520-626-1366, Fax: 520-626-0749, email: junwang@pharmacy.arizona.edu

MATERIALS AND METHODS

Cell lines and viruses. Human rhabdomyosarcoma (RD); A549, MDCK, Caco-2, and Vero cells were maintained in Dulbecco's modified Eagle's medium (DMEM), BEAS2B and HCT-8 cells were maintained in RPMI 1640 medium. Both medium was supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin antibiotics. Cells were kept at 37°C in a 5% CO2 atmosphere. The USA_WA1/2020 strain of SARS-CoV-2 obtained from the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA).

Protein expression and purification. SARS CoV-2 main protease (M^{pro} or 3CL) gene from strain BetaCoV/Wuhan/WIV04/2019 was ordered from GenScript (Piscataway, NJ) in the pET29a(+) vector with E. coli codon optimization. pET29a(+) plasmids with SARS CoV-2 main protease was transformed into competent E. coli BL21(DE3) cells, and a single colony was picked and used to inoculate 10 ml of LB supplemented with 50 g/ml kanamycin at 37°C and 250 rpm. The 10-ml inoculum was added to 1 liter of LB with 50 g/ml kanamycin and grown to an optical density at 600 nm of 0.8, then induced using 1.0 mM IPTG. Induced cultures were incubated at 37 °C for an additional 3 h and then harvested, resuspended in lysis buffer (25 mM Tris [pH 7.5], 750 mM NaCl, 2 mM dithiothreitol [DTT] with 0.5 mg/ml lysozyme, 0.5 mM phenylmethylsulfonyl fluoride [PMSF], 0.02 mg/ml DNase I), and lysed with alternating sonication and French press cycles. The cell debris were removed by centrifugation at 12,000 g for 45 min (20% amplitude, 1 s on/1 s off). The supernatant was incubated with Ni-NTA resin for over 2 h at 4°C on a rotator. The Ni-NTA resin was thoroughly washed with 30 mM imidazole in wash buffer (50 mM Tris [pH 7.0], 150 mM NaCl, 2 mM DTT); and eluted with 100 mM imidazole in 50 mM Tris [pH 7.0], 150 mM NaCl, 2 mM DTT. The imidazole was removed via dialysis or on a 10,000-molecular-weight-

cutoff centrifugal concentrator spin column. The purity of the protein was confirmed with SDS-PAGE. The protein concentration was determined via 260nM absorbance with ε 32890. EV-A71 2Apro and 3Cpro were expressed in the pET28b(+) vector as previously described (*1-3*).

Peptide synthesis. The SARS-CoV-2 M^{pro} FRET substrate Dabcyl-KTSAVLQ/SGFRKME(Edans) was synthesized by solid-phase synthesis through iterative cycles of coupling and deprotection using the previously optimized procedure.(*4*) Specifically, chemmatrix rink-amide resin was used. Typical coupling condition was 5 equiv of amino acid, 5 equiv of HATU, and 10 equiv of DIEA in DMF for 5 minutes at 80 ^{o}C. For deprotection, 5% piperazine plus 0.1 M HOBt were used and the mixture was heated at 80^{o}C for 5 minutes. The peptide was cleaved from the resin using 95% TFA, 2.5% Tris, 2.5% H_2O and the crude peptide was precipitated from ether after removal of TFA. The final peptide was purified by preparative HPLC. The purify and identify of the peptide were confirmed by analytical HPLC (> 98% purity) and mass spectrometry. $[M+3]^{3+}$ calculated 694.15, detected 694.90; $[M+4]^{4+}$ calculated 520.86, detected 521.35;

Native Mass Spectrometry. Prior to analysis, the protein was buffer exchanged into 0.2 M ammonium acetate (pH 6.8) and diluted to 10 μM. DTT was dissolved in water and prepared at a 400 mM stock. Each ligand was dissolved in ethanol and diluted to 10X stock concentrations. The final mixture was prepared by adding 4 μL protein, 0.5 μL DTT stock, and 0.5 μL ligand stock for final concentration of 4 mM DTT and 8 μM protein. Final ligand concentrations were used as annotated. The mixtures were then incubated for 10 minutes at room temperature prior to analysis. Each sample was mixed and analyzed in triplicate.

Native mass spectrometry (MS) was performed using a Q-Exactive HF quadrupole-Orbitrap mass spectrometer with the Ultra-High Mass Range research modifications (Thermo Fisher

Scientific). Samples were ionized using nano-electrospray ionization in positive ion mode using 1.0 kV capillary voltage at a 150 °C capillary temperature. The samples were all analyzed with a 1,000–25,000 m/z range, the resolution set to 30,000, and a trapping gas pressure set to 3. Between 10 and 50 V of source fragmentation was applied to all samples to aid in desolvation. Data were deconvolved and analyzed with UniDec.(*5*)

Enzymatic assays. For reaction condition optimization, 200 µM SARS CoV-2 Main protease was used. pH6.0 buffer contains 20 mM MES pH6.0, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol; pH6.5 buffer contains 20 mM HEPES pH6.5, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol, pH7.0 buffer contains 20 mM HEPES pH7.0, 120 mM NaCl, 0.4 mM EDTA, 4 mM DTT and 20% glycerol. Upon addition of 20 µM FRET substrate, the reaction progress was monitored for 1 hr. The first 15 min of reaction was used to calculate initial velocity (V_i) via linear regression in prism 5. Main protease displays highest proteolytic activity in pH6.5 buffer. All the following enzymatic assays were carried in pH6.5 buffer.

For the measurements of K_m/V_{max}, screening the protease inhibitor library, as well as IC_{50} measurements, proteolytic reaction with 100 nM Main protease in 100 µl pH6.5 reaction buffer was carried out at 30 °C in a Cytation 5 imaging reader (Thermo Fisher Scientific) with filters for excitation at 360/40 nm and emission at 460/40 nm. Reactions were monitored every 90 s. For K_m/V_{max} measurements, a FRET substrate concentration ranging from 0 to 200 µM was applied. The initial velocity of the proteolytic activity was calculated by linear regression for the first 15 min of the kinetic progress curves. The initial velocity was plotted against the FRET concentration with the classic Michaelis-Menten equation in Prism 5 software. For the screening protease inhibitor library and IC_{50} measurements, 100 nM Main protease was incubated with protease inhibitor at 30°C for 30 min in reaction buffer, and then the reaction was initiated by adding 10

μM FRET substrate, the reaction was monitored for 1 h, and the initial velocity was calculated for the first 15 min by linear regression. The IC_{50} was calculated by plotting the initial velocity against various concentrations of protease inhibitors by use of a dose-response curve in Prism 5 software. Proteolytic reaction progress curve kinetics measurements with GC376, MG132, Boceprevir, Calpain inhibitor II, and Calpain inhibitor XII used for curve fitting, were carried out as follows: 5 nM Main protease protein was added to 20 μM FRET substrate with various concentrations of testing inhibitor in 200 μl of reaction buffer at 30 °C to initiate the proteolytic reaction. The reaction was monitored for 4 hrs. The progress curves were fit to a slow binding Morrison equation (equation 3) as described previously (*1, 6*):

$$E + I \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} EI \underset{k_{-2}}{\overset{k_2}{\Leftrightarrow}} EI* \quad (1)$$

$$K_I = k_{-1}/k_1 \quad (2)$$

$$P(t) = P_0 + V_s t - (Vs - V_0)(1 - e^{-kt})/k \quad (3)$$

$$k = k_2[I]/(K_I + [I]) \quad (4)$$

where P(t) is the fluorescence signal at time t, P_0 is the background signal at time zero, V_0, V_s, and and k represent, respectively, the initial velocity, the final steady-state velocity and the apparent first-order rate constant for the establishment of the equilibrium between EI and EI* (*6*). k_2/K_I is commonly used to evaluate the efficacy for covalent inhibitor. We observed substrate depletion when proteolytic reactions progress longer than 90 min, therefore only first 90 min of the progress curves were used in the curve fitting (Figure 6 middle column). In this study, we could not accurately determine the k_2 for the protease inhibitors: Calpain inhibitor II, MG132, Boceprevir, and Calpain inhibitor XII, due to the very slow k_2 in these case: significant substrate depletion before the establishment of the equilibrium between EI and EI*. In these

cases, K_I was determined with Morrison equation in Prism 5.

Differential scanning fluorimetry (DSF). The binding of protease inhibitors on Main protease protein was monitored by differential scanning fluorimetry (DSF) using a Thermal Fisher QuantStudio™ 5 Real-Time PCR System. TSA plates were prepared by mixing Main protease protein (final concentration of 3 μM) with inhibitors, and incubated at 30 °C for 30 min. 1× SYPRO orange (Thermal Fisher) were added and the fluorescence of the plates were taken under a temperature gradient ranging from 20 to 90 °C (incremental steps of 0.05 °C/s). The melting temperature (T_m) was calculated as the mid-log of the transition phase from the native to the denatured protein using a Boltzmann model (Protein Thermal Shift Software v1.3). Thermal shift which was represented as ΔT_m was calculated by subtracting reference melting temperature of proteins in DMSO from the T_m in the presence of compound.

Cytotoxicity measurement. A549, MDCK, HCT-8, Caco-2, Vero, and BEAS2B cells for cytotoxicity CPE assays were seeded and grown overnight at 37 °C in a 5% CO2 atmosphere to ~90% confluence on the next day. Cells were washed with PBS buffer and 200 μl DMEM with 2% FBS and 1% penicillin−streptomycin, and various concentration of protease inhibitors was added to each well. 48 hrs after addition the protease inhibitors, cells were stained with 66 μg/ mL neutral red for 2 h, and neutral red uptake was measured at an absorbance at 540 nm using a Multiskan FC microplate photometer (Thermo Fisher Scientific). The CC_{50} values were calculated from best-fit dose−response curves using GraphPad Prism 5 software.

SARS-CoV-2 CPE assay. Antiviral activities of test compounds were determined in nearly confluent cultures of Vero 76 cells. The assays were performed in 96-well Corning microplates. Cells were infected with approximately 60 cell culture infectious doses ($CCID_{50}$) of SARS-CoV-2 and 50% effective concentrations (EC_{50}) were calculated based on virus-induced cytopathic

effects (CPE) quantified by neutral red dye uptake after 5 days of incubation. Three microwells at each concentration of compound were infected. Two uninfected microwells served as toxicity controls. Cells were stained for viability for 2 h with neutral red (0.11% final concentration). Excess dye was rinsed from the cells with phosphate-buffered saline (PBS). The absorbed dye was eluted from the cells with 0.1 ml of 50% Sörensen's citrate buffer (pH 4.2)-50% ethanol. Plates were read for optical density determination at 540 nm. Readings were converted to the percentage of the results for the uninfected control using an Excel spreadsheet developed for this purpose. EC_{50} values were determined by plotting percent CPE versus $\log_{10}$ inhibitor concentration. Toxicity at each concentration was determined in uninfected wells in the same microplates by measuring dye uptake.

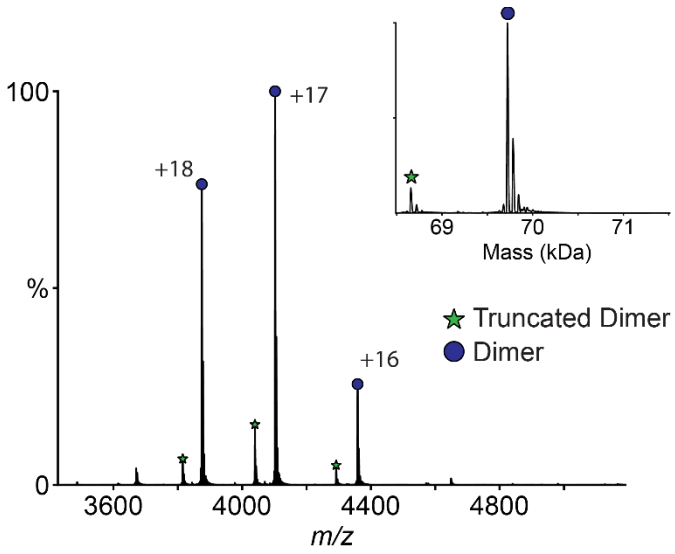
SARS-CoV-2 VYR assay. Virus yield reduction (VYR) assays were conducted by first replicating the viruses in the presence of test compound. Supernatant was harvested 3 days post-infection from each concentration of test compound and the virus yield was determined by endpoint dilution method. Briefly, supernatant virus was serially diluted in $\log_{10}$ increments then plated onto quadruplicate wells of 96-well plates seeded with Vero 76 cells. The presence or absence of CPE for determining a viral endpoint was evaluated by microscopic examination of cells 6 days after infection. From these data, 90% virus inhibitory concentrations (EC_{90}) were determined by regression analysis.

Influenza A virus A/California/07/2009 (H1N1) plaque reduction assay. The plaque assay was performed according to previously published procedures.(*7*)

M^{pro} crystallization and structure determination. 10 mg / mL of SARS-CoV-2 Mpro was incubated with 2 mM GC376 at 4˚ C O/N. The protein was diluted to 2.5 mg / mL the following

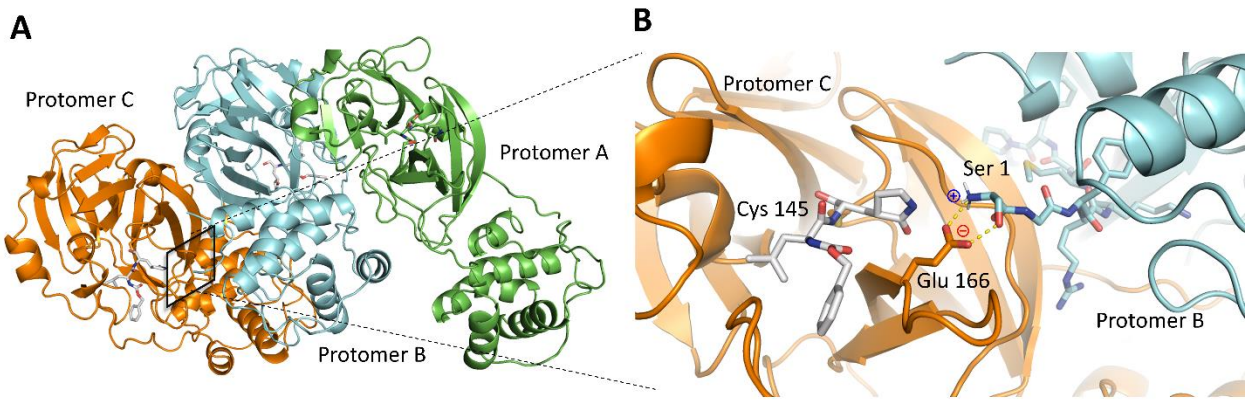
day in protein buffer (50 mM Tris pH 7.0, 150 mM NaCl, 4 mM DTT). Since GC3760 is water soluble, no precipitation was observed, and centrifugation was not necessary. Crystals were grown by mixing 2 uL of the protein solution with 1 ul of the precipitant solution (15 % PEG 2K, 10% 1,6-hexanediol, and 0.2 M NaCl) in a hanging-drop vapor-diffusion apparatus. Crystals were cryoprotected by transferring to a cryoprotectant solution (20% PEG 2K, 10% 1,6-hexanediol, 20% glycerol) and flash-frozen in liquid nitrogen.

X-ray diffraction data for the SARS-CoV2-M^{pro} GC376 complex structure was collected on the SBC 19-ID beamline at the Advanced Photon Source (APS) in Argonne, IL, and processed with the HKL3000 software suite(*8*). The CCP4 versions of MOLREP were used for molecular replacement using a previously solved SARS-CoV-2 M^{pro} (PDB ID 5RGG) as a reference model(*9*). Rigid and restrained refinements were performed using REFMAC and model building with COOT(*10, 11*). Protein structure figures were made using PyMOL (Schrödinger, LLC).



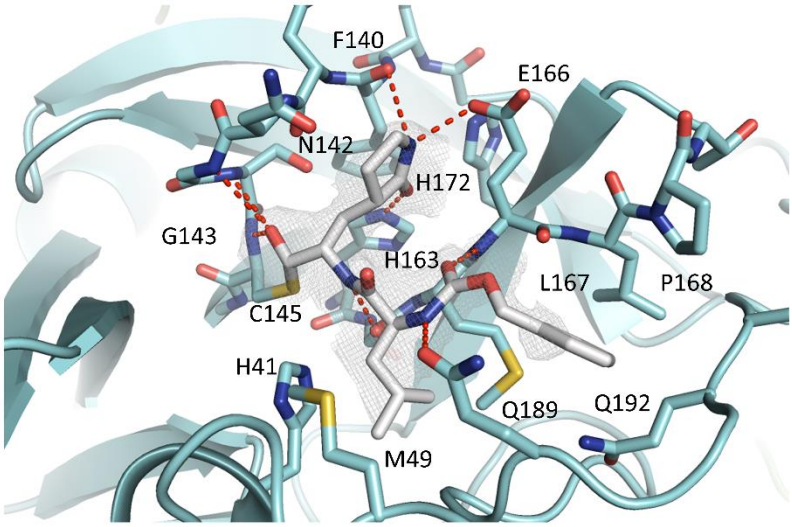
Supplementary Figure 1. Native mass spectrum of SARS-CoV-2 M^{pro}.

Native mass spectrum of SARS-CoV-2 M^{pro} with 4 mM DTT shows a dimer (*blue circle*) with a small amount of truncated dimer where one subunit has lost the C-terminal His tag (*green star*). The primary charge states are labeled, and the inset shows the deconvolved zero-charge mass distribution.

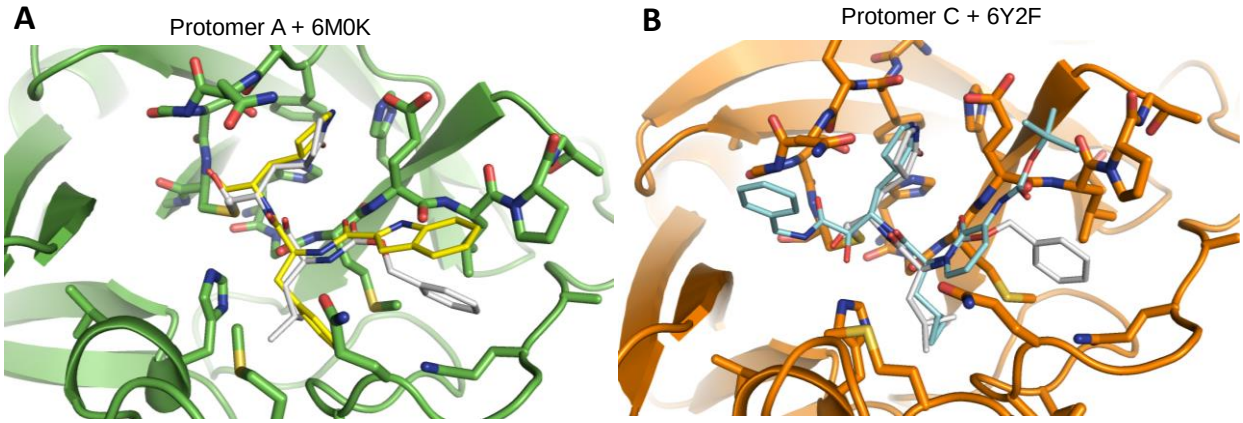


Supplementary Figure 2. Overall structure of SARS-CoV-2 M^{pro}.

(A) The three protomers in the asymmetric unit. Protomers B and C form a biological dimer. Protomer A dimerizes with a protomer from an adjacent asymmetric unit (not depicted). (B) The N-finger, or the N-terminal eight residues interact with Glu166 of the adjacent protomer, an important feature for catalytic activity.



Supplementary Figure 3. Complex structure of SARS-CoV-2 M^{pro} protomer B with GC-376 (64). Unbiased F_o-F_c map, shown in grey, is contoured at 2 σ. Hydrogen bonds are shown as red dashed lines.



Supplementary Figure 4. Overlay structures of current X-ray crystal structure with previously solved structures. (A) Overlay structures of protomer A with compound **13b** at the active site (PDB: 6M0K). (B) Overlay structures of protomer C with compound **N3** at the active site (PDB: 6Y2F). Unbiased F_o-F_c map, shown in grey, is contoured at 2 σ. Hydrogen bonds are shown as red dashed lines.

Supplementary Table 1: Cytotoxicity of SARS-CoV-2 M^{pro} inhibitors on various cell lines[a] and the counter screening against influenza virus.

	GC-376 (**64**)	Boceprevir (**28**)	MG-132 (**43**)	Calpain inhibitor II (**61**)	Calpain inhibitor XII (**62**)
MDCK	>100	>100	0.34 ± 0.02	>100	60.36 ± 2.28
Vero	>100	>100	0.45 ± 0.02	>100	>100
HCT-8	>100	>100	0.47 ± 0.02	>100	73.29 ± 11.80
A549	>100	>100	10.71 ± 3.50	>100	>100
Caco-2	>100	>100	<0.15	>100	82.02 ± 0.37
BEAS2B	>100	>100	0.14 ± 0.03	>100	78.91 ± 13.70
A/California/07/2009 (H1N1) antiviral activity[b] (μM)	> 20	> 20	N.T.	> 20	> 20

[a]Cytotoxicity was evaluated by measuring CC_{50} values (50% cytotoxic concentration) with CPE assay described in the method section. CC_{50} = mean ± S.E. of 3 independent experiments.

[b]Antiviral activity against influenza virus was tested in plaque assay.

Supplementary Table 2: Table of Crystallization Statistics

Data Collection	**PDB ID 6WTT**
Structure	SARS-CoV-2 M^{pro} + GC-376
Space Group	P 3_221
Cell Dimension	
a, b, c (Å)	101.83, 101.83, 160.02
α, β, γ (°)	90.00, 90.00, 120.00
Resolution (Å)	50.00 - 2.15
	(2.19 - 2.15)
R_{merge} (%)	0.107 (0.885)
<I>/σ<I>	5.9 (2.16)
Completeness (%)	100 (99.7)
Redundancy	9.12 (7.1)
Refinement	
Resolution (Å)	45.64 - 2.15
	(2.27 - 2.15)
No. reflections/free	52836 / 2711
R_{work}/R_{free}	0.227 /0.299
No. Heavy Atoms	7429
Protein	6968
Ligand/Ion	92
Water	369
B-Factors ($Å^2$)	
Protein	35.60
Ligand/Ion	33.07
Solvent	32.69
RMS Deviations	
Bond Lengths (Å)	0.015
Bond Angles (°)	1.84
Ramachandran Favored (%)	94.22
Ramachandran Allowed (%)	5.78
Ramachandran Outliers (%)	0.00
Rotameric Outliers (%)	2.08

***** Numbers in parentheses represent the highest resolution shell.

References:

1. R. Musharrafieh *et al.*, Validating Enterovirus D68-2A(pro) as an Antiviral Drug Target and the Discovery of Telaprevir as a Potent D68-2A(pro) Inhibitor. *Journal of virology* **93**, (2019).
2. L. Shang *et al.*, Biochemical characterization of recombinant Enterovirus 71 3C protease with fluorogenic model peptide substrates and development of a biochemical assay. *Antimicrob Agents Chemother* **59**, 1827-1836 (2015).
3. Q. Cai *et al.*, Conformational plasticity of the 2A proteinase from enterovirus 71. *Journal of virology* **87**, 7348-7356 (2013).
4. S. D. Cady, J. Wang, Y. Wu, W. F. DeGrado, M. Hong, Specific binding of adamantane drugs and direction of their polar amines in the pore of the influenza M2 transmembrane domain in lipid bilayers and dodecylphosphocholine micelles determined by NMR spectroscopy. *Journal of the American Chemical Society* **133**, 4274-4284 (2011).
5. M. T. Marty *et al.*, Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Analytical chemistry* **87**, 4370-4376 (2015).
6. J. F. Morrison, C. T. Walsh, The behavior and significance of slow-binding enzyme inhibitors. *Adv Enzymol Relat Areas Mol Biol* **61**, 201-301 (1988).
7. C. Ma, J. Zhang, J. Wang, Pharmacological Characterization of the Spectrum of Antiviral Activity and Genetic Barrier to Drug Resistance of M2-S31N Channel Blockers. *Mol Pharmacol* **90**, 188-198 (2016).
8. W. Minor, M. Cymborowski, Z. Otwinowski, M. J. A. C. S. D. B. C. Chruszcz, HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. **62**, 859-866 (2006).

9. A. Vagin, A. Teplyakov, Molecular replacement with MOLREP. *Acta crystallographica. Section D, Biological crystallography* **66**, 22-25 (2010).

10. G. N. Murshudov *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography* **67**, 355-367 (2011).

11. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-2132 (2004).