Inhibitor Binding Modulates Protonation States in the Active Site of SARS-CoV-2 Main Protease

Daniel W. Kneller,[1,2] Gwyndalyn Phillips,[1,2] Kevin L. Weiss,[1,2] Qiu Zhang,[1,2] Leighton Coates,[1,2,3]* and Andrey Kovalevsky[1,2]*

[1]*Neutron Scattering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, 37831, USA*

[2]*National Virtual Biotechnology Laboratory, US Department of Energy*

[3]*Second Target Station, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, 37831, USA*

*Corresponding Authors: Leighton Coates: coatesl@ornl.gov, Andrey Kovalevsky: kovalevskyay@ornl.gov

Notice: This manuscript has been authored by UT-Battelle LLC under DOE Contract No. DE-AC05-00OR22725.The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. http://energy.gov/downloads/doe-public-access-plan

ABSTRACT

The main protease (3CL M^{pro}) from SARS-CoV-2, the virus that causes COVID-19, is an essential enzyme for viral replication with no human counterpart, making it an attractive drug target. Although drugs have been developed to inhibit the proteases from HIV, hepatitis C and other viruses, no such therapeutic is available to inhibit the main protease of SARS-CoV-2. To directly observe the protonation states in SARS-CoV-2 M^{pro} and to elucidate their importance in inhibitor binding, we determined the structure of the enzyme in complex with the α-ketoamide inhibitor telaprevir using neutron protein crystallography at near-physiological temperature. We compared protonation states in the inhibitor complex with those determined for a ligand-free neutron structure of M^{pro}. This comparison revealed that three active-site histidine residues (His41, His163 and His164) adapt to ligand binding, altering their protonation states to accommodate binding of telaprevir. We suggest that binding of other α-ketoamide inhibitors can lead to the same protonation state changes of the active site histidine residues. Thus, by studying the role of active site protonation changes induced by inhibitors we provide crucial insights to help guide rational drug design, allowing precise tailoring of inhibitors to manipulate the electrostatic environment of SARS-CoV-2 M^{pro}.

Keywords: joint neutron/X-ray crystallography, room temperature, SARS-CoV-2, COVID-19, main protease, 3CL M^{pro}, protonation state, hydrogen bond, drug design

INTRODUCTION

The number of confirmed COVID-19 cases worldwide is relentlessly marching towards one hundred million, while the number of deaths is approaching a grim milestone of two million. Sadly, this deadly disease caused by the novel coronavirus SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2)[1-4] has become one of the leading causes of death on the planet in 2020, according to the World Health Organization (www.who.int). Although several vaccines have been developed[5-7] to slow the spread of SARS-CoV-2, the need for therapeutic intervention options, including small-molecule drugs that inhibit essential steps in the viral replication cycle, cannot be overstated.[8-12] Small-molecule drugs have shown tremendous success in treating people infected with HIV,[13,14] hepatitis C[15,16,] and influenza[17,18] viruses, and an RNA polymerase inhibitor remdesivir has been recently approved for treatment of COVID-19 by the US Food and Drug Administration.[19]

SARS-CoV-2, a single-stranded, positive-sense RNA virus with a genome comprised of ~ 30k nucleotides, belongs to the β-coronavirus genus of the *Coronaviridae* family.[20] A vital step in the viral replication cycle is the cleavage of two polyproteins, pp1a and pp1ab, encoded in the viral replicase gene into individual functional viral proteins.[20,21] Each polyprotein is mainly processed, or hydrolyzed, by a chymotrypsin-like protease, 3CL M^{pro} or main protease, that belongs to the class of cysteine protease enzymes.[22,23] The functional main protease (hereafter referred to as M^{pro}) is essential for SARS-CoV-2 proliferation as the production of infectious virions depends entirely on the enzymatic activity of M^{pro}. Hence, SARS-CoV-2 M^{pro} is undeniably a crucial target for designing specific small-molecule protease inhibitors[24-29]and for potential repurposing of known clinical drugs.[30-35] Though no clinical drugs are available for use against SARS-CoV-2 M^{pro}, several protease inhibitors have been designed to inhibit the very closely related SARS-CoV M^{pro}[36-39] that shares 96% of amino acid sequence identity with the SARS-CoV-2 enzyme and has a similar catalytic efficiency and almost identical three-dimensional structure.[25,27,40,41]

Two identical protomers of SARS-CoV-2 M^{pro}, each with a molecular mass of ~34 kDa, create the catalytically active homodimeric enzyme through non-covalent interactions (Fig 1A). Each protomer consists of three structural/functional domains – catalytic domains I (residues 8-

101) and II (residues 102-184), and the α-helical domain III (residues 201-303) crucial for protein dimerization (Fig. 1B). Previous studies have shown that the monomeric enzyme is catalytically inactive, as was demonstrated for SARS-CoV M^{pro}.[42,43] The active site cavity is a shallow cleft located on the protein surface between domains I and II. There are six substrate binding subsites, named S1′ through S5, that can bind either substrate residues or chemical groups of inhibitors in positions P1′ through P5. Peptide bond cleavage is carried out at the base of the well-defined subsite S1, where the non-canonical catalytic dyad composed of Cys145 and His41 is located. Catalysis is believed to be assisted by a water molecule positioned at the protein interior side of subsite S2 and hydrogen bonded to the catalytic His41, His164, and Asp187.[25,27,34,40] Scissile peptide bond cleavage begins through a nucleophilic attack by the Cys145 thiolate on the substrate carbonyl carbon. The negatively charged oxygen of the resultant hemithioketal intermediate is stabilized by a canonical oxyanion hole formed by the main chain amide NH groups of Gly143, Ser144, and Cys145.[44] Interestingly, subsites S2 and S4 need to be carved out by the substrate or inhibitor groups P2 and P4, respectively, that push protein residues away from their positions in the ligand-free enzyme.[45] Conversely, subsites S1′, S3, and S5 are fully exposed to the bulk solvent.

Our recent neutron crystallographic study of the ligand-free SARS-CoV-2 M^{pro} provided direct visualization of hydrogen (H) atom locations and hydrogen bonding interactions throughout the enzyme.[46] The catalytic dyad was observed in a zwitterionic form in the enzyme without substrate or inhibitor. The zwitterion is comprised of a deprotonated, negatively charged Cys145 thiolate and a doubly protonated, positively charged His41 imidazolium situated ~ 3.8 Å apart from each other. Neutrons are unique probes of light atoms such as H as they are scattered by atomic nuclei instead of electron clouds, which interact with X-rays. Therefore, neutron crystallography can be used to accurately determine the positions of H and its heavier isotope deuterium (D) atoms in biomacromolecules.[47,48] Deuterium has a potent neutron scattering length of 6.671 fm (https://www.ncnr.nist.gov/resources/n-lengths/), allowing its detection in protein structures even at moderate resolutions of 2.5-2.6 Å.[49-51] Also, cold neutrons used in neutron crystallographic experiments with wavelengths in the range of 2-5 Å cause no radiation damage to biological samples; therefore, neutron diffraction data can be collected at near-physiological (room) temperature avoiding possible artifacts produced during cryo-cooling and cryo-protection of protein crystals necessary for the mitigation of radiation damage in X-ray cryo-crystallography.[52]

Structure-guided and computer-assisted drug design approaches require a detailed atomic picture of a target biomacromolecule, which is generally obtained using X-ray cryo-crystallography. However, half of the atoms, namely H atoms, are often overlooked in X-ray structures because they usually cannot be located with confidence unless X-ray data extend to sub-Å resolution. Even then, the most interesting and functionally relevant H atoms are typically invisible in electron density maps at ultra-high resolutions because they are mobile and electron-poor, often participating in highly polarized chemical bonds.[53-55] Detailed structures of the ligand-free and ligand-bound drug targets are essential to steer drug design efforts in the right direction. Knowledge of where H atoms relocate due to inhibitor binding can provide critical information on how protonation states and thus electric charges are modulated in the protein active site cavity, improving rational drug design. For example, molecular dynamics simulations have demonstrated that protonation states in the SARS-CoV-2 M^{pro} active site cavity can be altered when an inhibitor binds,[56] but experimental evidence has been lacking.

We recently obtained room-temperature X-ray structures of the SARS-CoV-2 M^{pro} in complex with hepatitis C clinical protease inhibitors boceprevir, narlaprevir, and telaprevir, establishing their mode of binding and mechanism of action.[57] Using *in vitro* enzyme inhibition assays, we determined that these clinical drugs designed against hepatitis C NS3/4A protease inhibited the SARS-CoV-2 M^{pro} activity in the micromolar range. Here, we report a room-temperature neutron structure of the SARS-CoV-2 M^{pro} in complex with telaprevir (M^{pro}-Telaprevir) determined at 2.4 Å resolution and refined jointly with a room-temperature X-ray dataset collected from the same crystal to 2.0 Å resolution, thereby improving the accuracy of the experimental model (Table S1).[58] We chose this complex because it produces crystals of morphology and size amenable for neutron diffraction (Figure S1). Telaprevir represents a promising class of covalent protease inhibitors called α-ketoamides, while also possessing chemical groups in positions P1' through P5 (Fig. 1C,D).[27] Locations of D atoms in M^{pro} were accurately determined and compared to those observed in our previous neutron structure[46] of the ligand-free enzyme, which was crystallized at the same pD value. We discovered that protonation states of key histidine residues (His41, His163, and His164) in the SARS-CoV-2 M^{pro} active site cavity are altered after telaprevir binds, revealing different electric charges between the ligand-free and ligand-bound states of the enzyme. Remarkably, the net electric charge of +1 in the active site cavity is conserved between the two structures. These results are crucial for continued efforts in structure-guided and computational drug design, emphasizing the essential nature of a full atomic-level visualization of protein structure, function, and inhibition.

RESULTS

Protonation states in the catalytic site. The reactive carbonyl warhead of the telaprevir α-ketoamide group is attacked by the Cys145 thiolate nucleophile to generate a hemithioketal. Like other α-ketoamide inhibitors,[27,34,57] the nucleophilic attack is stereospecific, generating only the *S*-enantiomer, in which the newly formed hydroxyl group of the hemithioketal faces towards His41 (Fig. 2A). The distance between the hemithioketal oxygen and His41 Nε2 is 2.6 Å, implicating a strong hydrogen bond. In our neutron structure of the ligand-free SARS-CoV-2 M^{pro}, we observed that the Cys145-His41 catalytic dyad adopts a charge-separated state.[46] Hence, one might suggest that the hemithioketal-His41 pair could maintain charge separation, keeping the positive charge on His41 and having the deprotonated negatively charged hydroxyl on hemithioketal in M^{pro}-Telaprevir. Instead, the nuclear density clearly shows that the hemithioketal hydroxyl is protonated (Fig. 2B), whereas His41 is neutral. A reasonable explanation of this observation is that the hemithioketal oxygen is protonated by His41 through a direct proton transfer either in concert with the Cys145 nucleophilic attack on the carbonyl warhead of telaprevir or following the hemithioketal S-C bond formation. Interestingly, the hydroxyl D atom is not positioned on the straight line connecting the hemithioketal oxygen and His41 Nε2 atoms, but instead the O-D…Nε2 angle is 129° indicating that this hydrogen bond may not be strong (Fig 2C).

In M^{pro}-Telaprevir, the His41 imidazole Nδ1 forms a 2.7 Å hydrogen bond with the catalytic water molecule (Nδ1-D…O$_{D2O}$ distance of 1.7 Å) (Fig. 2C), which is nearly identical to that observed in the ligand-free M^{pro}. The catalytic water is also hydrogen bonded with His164 and Asp187, but it is 3.5 Å away from the main chain ND of His41 compared to 3.2 Å in the ligand-free M^{pro}. Thus, the latter hydrogen bond is not present when telaprevir binds. Moreover, in the inhibitor complex, His164 is neutral, having lost the Nε2 D atom that was observed in ligand-free M^{pro}. This leads to an increase in $N\varepsilon2_{His164}…O\gamma1_{Thr175}$ distance of 0.2 Å. Hence, the hydrogen bond between His164 and Thr175 is effectively abolished upon telaprevir binding. The hydroxyl of Thr175 remains rotated towards and donates its D in a hydrogen bond with the main chain carbonyl of Asp176. Consequently, in our M^{pro}-Telaprevir structure, the catalytic water molecule is no longer surrounded by positively charged histidine side chains, significantly altering the electrostatics around the catalytic site.

Protonation states in subsite S1. Important ionizable residues of subsite S1 are shown in Figure 3A. Subsite S1 is bordered by residues 140-144 spanning the oxyanion hole, His163, Glu166, and the N-terminal Ser1′ of the other enzyme protomer. Subsite S1 is selective for Gln at substrate position P1. Inhibitors possessing sub-μM affinity to M^{pro} typically mimic Gln by introducing a lactam functionality as their P1 substituent in order to engage His163 in a hydrogen bond.[39] Telaprevir features a hydrophobic norvaline substituent in P1 position that cannot make hydrogen bonds in subsite S1. Compared to the ligand-free structure, inhibitor binding leads to the recruitment of a water molecule into subsite S1. This water molecule is positioned between the telaprevir's P1 norvaline and His163 and is hydrogen bonded with the His163 imidazole. Both the electron and nuclear densities for this water molecule are weak, indicating its mobility. Despite the unfavorable norvaline P1 moiety, His163's protonation state has changed such that the imidazole is doubly protonated and positively charged in M^{pro}-Telaprevir compared to that in the ligand-free enzyme, where the Nδ1 not facing the S1 subsite was found to be protonated.[46] This is significant because His163 protonation explains the apparent ability of the P1 lactam carbonyl to act as a hydrogen bond acceptor with the His163 imidazole observed in several previous X-ray structures of other M^{pro} inhibitors.[26,27,34] Notably, superposition of M^{pro}-Telaprevir with the X-ray structure of inhibitor 13b complex (PDB ID 6Y2F)[27] reveals that the water molecule position observed in our neutron structure mimics the carbonyl oxygen of inhibitor 13b (Figure S2).

Similar to the ligand-free enzyme, the N-terminal amine is found to be protonated and positively charged (ND_3^+) in M^{pro}-Telaprevir, where D atoms form a hydrogen bond with the Phe140 main chain carbonyl, the Glu166 carboxylate, and a D_2O molecule. However, the Glu166 carboxylate side chain is rotated by almost 90° relative to its conformation in the ligand-free enzyme, moving it 0.5 Å towards His172 (Figure 3B). Thus, the $O\varepsilon1_{Glu166}$…$N\varepsilon2_{His172}$ distance is reduced from 3.3 Å to 2.8 Å, shortening the D…O distance from 2.5 Å to 2.0 Å upon telaprevir binding and considerably strengthening this hydrogen bond. Furthermore, $O\varepsilon1_{Glu166}$ in the telaprevir complex is 1.2 Å closer to $N\varepsilon2_{His163}$ than in ligand-free M^{pro}, in agreement with His163 becoming protonated in the M^{pro}-Telaprevir inhibitor complex.

DISCUSSION

Crystallographic structures are essential for making well-educated predictions when designing inhibitors utilizing computer-assisted and structure-guided drug design approaches. It is crucial to determine the locations of heavy atoms and the lighter hydrogen atoms that govern the protonation states of amino acid residues and inhibitor substituents.[55] The presence or absence of hydrogen atoms on ionizable chemical groups, *e.g.* imidazole of histidine, amine of lysine or N-terminus, or carboxylate of aspartate or glutamate, governs their electric charges and thereby regulates the electrostatics in the vicinity of these groups. Detecting hydrogen atoms in a drug target and a bound ligand is undoubtedly essential to establish protonation states and hydrogen bonding interactions and decipher how protonation states may change due to ligand binding. The potent scattering of neutrons from hydrogen and its isotope deuterium enables neutron crystallography to directly and accurately map hydrogen atoms in biological macromolecules and bound ligands.[46,49-51,59,60] In this work, we have successfully determined a room-temperature neutron structure of SARS-CoV-2 M^{pro} in complex with the hepatitis C clinical protease inhibitor, telaprevir, from a 0.5 mm^3 crystal using a partially deuterated enzyme. This structure has permitted us to perform a detailed analysis of D atom positions and compare the protonation states in the active site cavity between the ligand-bound and ligand-free states of the M^{pro} with broad implications for drug design.

In the SARS-CoV-2 M^{pro}-Telaprevir neutron structure, we observed protonation of the hemithioketal hydroxyl, likely occurring through a proton transfer from His41 that was found doubly protonated in our previous neutron structure of the ligand-free enzyme. The hemithioketal hydroxyl makes a short, but probably weak, hydrogen bond with the neutral His41 in M^{pro}-Telaprevir that would be required to enhance inhibitor binding affinity. Stronger hydrogen bonds would possibly form when the hemithioketal or hemithioacetal oxygen from a warhead carbonyl is directed into the oxyanion hole upon covalent attack on the inhibitor, as was observed in the structures of some M^{pro} inhibitors.[26,34] The neutral His41 position is stabilized by an additional hydrogen bond with the catalytic water molecule, whose position is held by hydrogen bonds with Asp187 and His164. His164, positioned deeper in the protein interior and observed as doubly protonated in the ligand-free enzyme, has lost its Nε2 D atom resulting in the imidazole ring moving slightly away from its hydrogen bond partner Thr175 upon telaprevir binding. Intriguingly, His163 located in the subsite S1 has gained a proton on its Nε2 atom that faces the bulk solvent in the ligand-free enzyme and P1 norvaline in the telaprevir-bound complex. Moreover, His163 protonation, evidently driven by inhibitor binding, induces a conformational change of the Glu166 side chain that brings it closer to the His163 imidazolium while significantly shortening its hydrogen bond with His172 and retaining the hydrogen bond with the N-terminal ammonium of the second protomer. This observation agrees with His163 becoming positively charged, attracting the negatively charged Glu166. The establishment of hydrogen bond interactions between His163 and the inhibitor lactam ring in several X-ray structures of other inhibitor complexes implies that His163 is protonated in these complexes as well.[25-27,34,61]

We have observed that the binding of a covalent inhibitor can modulate the protonation states of histidine residues in the active site cavity of SARS-CoV-2 M^{pro}. His41, His163, and His164 have different protonation states relative to those in the ligand-free enzyme (Table 1). His41 and His164 lose hydrogens to become uncharged and neutral, whereas His163 gains a proton to become positively charged when telaprevir binds. Cys145 is also uncharged as it is covalently bonded to telaprevir. Nevertheless, the overall electric charge of the active site cavity of +1 is maintained upon inhibitor binding, even though the electrostatics of the ligand-binding cavity differ significantly between the ligand-free and ligand-bound enzyme. It is not unreasonable to suggest that the same protonation state changes of the active site histidine residues may occur upon binding of other α-ketoamide inhibitors. Furthermore, we note that the exact protonation states of the ionizable residues we observed in the active site cavities of ligand-free[46] and inhibitor-bound SARS-CoV-2 M^{pro} have not been predicted by molecular simulations,[56] emphasizing the critical importance of experimentally determining the locations of H atoms. As a result, the design of covalent inhibitors against SARS-CoV-2 M^{pro} should consider the observed protonation state changes of the ionizable residues in the active site cavity triggered by inhibitor binding.

CONCLUSION

Enzymes rely upon the transfer and movement of hydrogen atoms to carry out their function. We have shown that the active site cleft of SARS-CoV-2 M^{pro} contains three histidine residues that can change their protonation states upon an inhibitor binding, revealing how M^{pro} can tune its electrostatics to accommodate inhibitor and substrate binding into the active site cleft. We observed that inhibitor binding results in a cascade of protonation/deprotonation (Table 2) and conformational events that remodel the SARS-CoV-2 M^{pro} active site cavity as follows:

1) Upon hemithioketal formation a proton is transferred from doubly protonated His41 to the hemithioketal hydroxyl to retain the overall neutral charge of the catalytic site.
2) His164 loses a proton from the Nε2 atom, with no charge, and moves away from Thr175 side chain hydroxyl, which in effect eliminates the hydrogen bond between these two residues.
3) His163 obtains a proton to become doubly protonated, recruiting a loosely bound water molecule to Nε2. The positively charged imidazolium cation attracts the negatively charged carboxylate of Glu166 that changes its conformation to move 1.2 Å closer to His163 and concomitantly strengthens its hydrogen bond with His172.

4) His172 remains singly protonated on Nε2, with no charge, because Nδ1 is locked in a hydrogen bond with the main chain NH of Gly138.
5) The overall +1 electric charge of the active site cavity is retained.

By considering how the protonation states and electric charges of the ionizable histidine residues in the active site can be altered by inhibitor binding in structure-assisted and computational drug design, protease inhibitors can be improved to specifically target the M^{pro} enzyme from SARS-CoV-2.

METHODS

General Information. Protein purification columns were purchased from Cytiva (Piscataway, New Jersey, USA). Crystallization reagents were purchased from Hampton Research (Aliso Viejo, California, USA). Crystallographic supplies were purchased from MiTeGen (Ithaca, New York, USA) and Vitrocom (Mountain Lakes, New Jersey, USA). Telaprevir was purchased from BioVision Inc. (Milpitas, CA, USA). A detailed protocol for hydrogenated enzyme expression, purification, and crystallization of M^{pro} to grow neutron diffraction quality crystals has been published elsewhere.[62]

Cloning, expression, and purification of partially deuterated SARS-CoV-2 M^{pro}. The codon-optimized sequence of M^{pro} from SARS-CoV-2 was cloned into the pD451-SR vector harboring kanamycin resistance (ATUM, Newark, CA) and transformed into chemically-competent *E. coli* BL21 (DE3) cells. Before producing deuterated M^{pro}, a frozen glycerol stock was first revived in H_2O minimal medium. Unlabeled glucose (0.5% w/v) and kanamycin (100 mg/mL) was used in this and all other media. After initial growth in H_2O minimal medium, the cells were adapted stepwise to minimal medium containing increasing percentages (50,75, and 100%) of D_2O. The final D_2O-adapted preculture was used to inoculate the bioreactor vessel to an initial volume of 2.8 L. Following inoculation, the BioFlo 310 bioreactor controller console (Eppendorf, Enfield, CT) was set to maintain the temperature (30 C) and the dissolved oxygen level (>30%). The pD was kept above 7.3 by the controlled addition of sodium deuteroxide solution in D_2O (10% w/w). Once the initial glycerol was exhausted, the culture was fed with a solution containing 20% (w/v) unlabeled glucose and 0.2% (w/v) $MgSO_4$. At an OD_{600} of ~8.8, isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM in order to induce protein expression. The cells were collected ~13.5 h later by centrifugation at 6,000g for 40 min. After removing the supernatant, the wet cell paste was harvested and stored at -80 C until further use. The protein was purified according to the published procedure.[45] Upstream of the M^{pro} N-terminus codes for a maltose binding protein (MBP) followed by the protease autoprocessing site SAVLQ↓SGFRK (arrow indicates the autocleavage site) which corresponds to the cleavage position between NSP4 and NSP5 in the viral polyprotein. Downstream to the M^{pro} C-terminus codes for the human rhinovirus 3C (HRV-3C) protease cleavage site (SGVTFQ↓GP), which is connected to a 6xHis tag. The N-terminal flanking sequence is autoprocessed *in vivo* during the expression, whereas the C-terminal flanking sequence is removed upon *in vitro* treatment with HRV-3C protease (Millipore Sigma, St. Louis, MO).

Crystallization. Initial protein crystallization conditions were discovered by screening conducted at the Hauptman-Woodward Medical Research Institute (HWI).[63] Crystal aggregates were reproduced using the sitting drop vapor diffusion method using 25% PEG3350, 0.1 M Bis-Tris pH 6.5 in 20μL drops with 1:1 ratio of the protein:well solution. Aggregates were transformed into microseeds using Hampton Research Seed Beads™. For partially deuterated M^{pro}-Telaprevir co-crystallization, freshly purified deuterated M^{pro} in 20 mM Tris, 150 mM NaCl, 1 mM TCEP, pH 8.0 was concentrated to ~10.3 mg/mL and mixed with telaprevir, from 60 mM stocks in 100% DMSO, at a 1:5 molar ratio. After room-temperature incubation for 30 minutes, precipitation in the sample was removed via centrifugation and filtration through a 0.2-micron centrifugal filter. Crystallization was achieved in a Hampton 9-well plate and sandwich box set-up with 50 μL drops of protein mixed with 18% PEG3350, 0.1 M Bis-Tris pH 6.5 at 1:1 ratio seeded with 0.2 μL of microseeds at a 1:200 dilution. After 28 days incubation at 14°C, the crystal used for neutron diffraction data collection grew to final dimensions of ~1.5x0.7x.05 mm (~0.5 mm^3) (Figure S1). The crystal was mounted in a fused quartz capillary accompanied with 20% PEG3350 prepared with 100% D_2O and allowed to H/D exchange for two weeks before starting the neutron data collection. The pH in the crystallization drop at the time of crystal mounting was measured by microelectrode to be 6.6, corresponding to a final pD of 7.0 (pD = pH + 0.4); these are identical conditions to those previously used to determine the neutron structure of the ligand free enzyme.[46]

Neutron diffraction data collection. The M^{pro}-Telaprevir crystal was screened for diffraction quality using a broad-bandpass Laue configuration using neutrons from 2.8 to 10 Å at the IMAGINE instrument at the High Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory.[64-67] The full neutron diffraction dataset was then collected using the Macromolecular Neutron Diffractometer (MaNDi) instrument at the Spallation Neutron Source (SNS).[67-69] The crystal was held stationary at room temperature, and diffraction data were collected for 20 hours using all neutrons between 2-4.16 Å. The next 20-hour diffraction image was collected after crystal rotation by $\Delta\varphi = 10°$. A total of twenty-one data frames were collected in the final neutron dataset. Diffraction data were reduced using the Mantid package, with integration carried out using three-dimensional TOF profile fitting.[70] Wavelength normalization of the Laue data was performed using the Lauenorm program from the Lauegen suite.[71,72] The neutron data collection statics are shown in Table S1.

X-ray diffraction data collection. The room-temperature X-ray diffraction dataset was collected from the same crystal following the neutron data collection on a Rigaku HighFlux HomeLab instrument equipped with a MicroMax-007 HF X-ray generator and Osmic VariMax optics. The diffraction images were collected using an Eiger R 4M hybrid photon counting detector. Diffraction data were integrated using the CrysAlis Pro software suite (Rigaku Inc., The Woodlands, TX). Diffraction data were then reduced and scaled using the Aimless[73] program from the CCP4 suite[74]; molecular replacement using PDB code 6XQS[57] was then performed with Molrep from the CCP4 program suite. The protein structure was first refined against the X-ray using *Phenix.refine* from the Phenix[75] suite of programs to obtain an accurate model for the subsequent X-ray/neutron joint refinement. The X-ray data collection statics are shown in Table S1.

Joint X-ray/neutron refinement. The joint X-ray/neutron refinement of ligand-free 3CL M^{pro} was performed using *nCNS*[76], and the structure was manipulated in *Coot*.[77] After initial rigid-body refinement, several cycles of positional, atomic displacement parameter, and occupancy refinement were performed. The structure was checked for the correctness of side-chain conformations, hydrogen bonding, and orientations of D_2O water molecules built based on the mF_O-DF_C difference neutron scattering length density maps. The $2mF_O$-DF_C and mF_O-DF_C neutron scattering length density maps were then examined to determine the correct orientations of hydroxyl (Ser, Thr, Tyr), thiol (Cys) and ammonium (Lys) groups, and protonation states of the enzyme residues. The protonation states of some disordered side chains could not be obtained directly and remained ambiguous. All water molecules were refined as D_2O. Initially, water oxygen atoms were positioned according to their electron density peaks and then were shifted slightly in accordance with the neutron scattering length density maps. Because M^{pro} for this study was partially deuterated, all H positions in the protein were modeled as D atoms. In telaprevir only the labile H atoms were modeled as Ds. The occupancies of D atoms were refined individually within the range of −0.56 (pure H) to 1.00 (pure D) because the neutron scattering length of H is –0.56 times that of D. Before depositing the neutron structure to the PDB, a script was run that converts a record for the coordinates of a D atom into two records corresponding to an H and a D partially occupying the same site, both with positive partial occupancies that add up to unity. The percent D at a specific site is calculated according to the following formula: %D = {Occupancy(D) + 0.56}/1.56.

Data availability

The coordinates and structure factors for the SARS-CoV-2 3CL M^{pro}-Telaprevir complex have been deposited in the PDB with the accession code 7LB7. Any other relevant data are available from the corresponding authors upon reasonable request.

Acknowledgments

This research was supported by the DOE Office of Science through the National Virtual Biotechnology Laboratory (NVBL), a consortium of DOE national laboratories focused on response to COVID-19, with funding provided by the Coronavirus CARES Act. This research used resources at the Spallation Neutron Source and the High Flux Isotope Reactor, which are DOE Office of Science User Facilities operated by the Oak Ridge National Laboratory. The Office of

Biological and Environmental Research supported research at ORNL's Center for Structural Molecular Biology (CSMB), a DOE Office of Science User Facility. This research used resources at the Second Target Station, which is a DOE Office of Science User Facilities Construction Project at Oak Ridge National Laboratory. We thank Dr. Hugh M. O’Neill from ORNL for assistance during expression of the partially deuterated protein

Author contributions

L.C and A.K. conceived the study. A.K., G.P. and Q.Z. designed and cloned the gene. G.P., K.L.W. and Q.Z. performed expression of the partially deuterated protein. D.W.K. and A.K. crystallized the protein. A.K. and D.W.K. collected the X-ray diffraction data. L.C. collected and reduced the neutron diffraction data. A.K., D.W.K. and L.C. refined the structure. D.W.K., L.C., and A.K. wrote the paper with help from all co-authors.

Competing interests

The authors declare no competing interests.

REFERENCES

1. Hussain, Md. A., Yadav, S., Hadda, V., Suri, T. M., Tiwari, P., Mittal, S., Madan, K., Mohan, A. Covid-19: a comprehensive review of a formidable foe and the road ahead. Expert Rev. Resp. Med. 2020, 14, 869-879.

2. Rastogi, Y. R., Sharma, A., Nagraik, R., Aygun, A., Sen, F. The novel coronavirus 2019-nCoV: its evolution and transmission into humans causing global COVID-19 pandemic. Int. J. Environ. Sci. Technol. 2020, 17, 4381-4388.

3. Lotfi, M., Hamblin, M. R., Rezaei, N. COVID-19: transmission, prevention, and potential therapeutic opportunities. Clin. Chim. Acta 2020, 508, 254-266.

4. Helmy, Y. A., Fawzy, M., Elaswad, A., Sobieh, A., Kenney, S. P., Shehata, A. A. TheCOVID-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control. J. Clin. Med. 2020, 9, 1225.

5. Callaway, E. What Pfizer's landmark COVID vaccine results mean for the pandemic. Nature 2020, *doi:https://doi.org/10.1038/d41586-020-03166-8*

6. Voysey, M., Clemens, S. A. C., Madhi, S. A., Weckx, L. Y. et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomized controlled trails in Brazil, South Africa, and the UK. Lancet 2020, https://doi.org/10.1016/S0140-6736(20)32661-1

7. Callaway, E. COVID vaccine excitement builds as Moderna reports third positive result. Nature 2020, 587, 337-338.

8. Liu, C., Zhou, Q., Li, Yingzhu, Garner, L. V., Watkins, S. P., Carter, L. J., Smoot, J., Gregg, A. C., Daniels, A. D., Jervey, S., Albaiu, D. Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. ACS Central Sci. 2020, 6, 315-331.

9. Pillaiyar, T., Meenakshisundaram, S., Manickam, M. Recent discovery and development of inhibitors targeting coronaviruses. Drug Discov. Today 2020, 25, 668-688.

10. Sarkar, C., Mondal, M., Islam, M. T., Martorell, M., Docea, A. O., Maroyi, A., Sharifi-Rad, J., Calina, D. Potential therapeutic options for COVID-19: current status, challenges, and future perspectives. Frontiers Pharmacol. 2020, 11, 572870.

11. Zumla, A., Chan, J. F. W., Azhar, E. I., Hui, D. S. C., Yuen, K.-Y. Coronaviruses – drug discovery and therapeutic options. Nat. Rev. 2016, 15, 327-347.

12. Gil, C., Ginex, T., Maestro, I., Nozal, V., Barrado-Gil, L., Cuesta-Geijo, M. A., Urquiza, J., Ramirez, D., Alonso, C., Campillo, N. E., Martinez, A. COVID-19: drug targets and potential treatments. J. Med. Chem. 2020, 63, 12359-12386.

13. Engelman, A., Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat. Rev. 2012, 10, 279-290.

14. Lu, D.-Y., Wu, H.-Y., Yarla, N. S., Xu, B., Ding, J., Lu, T.-R. HAART in HIV/AIDS treatments: future trends. Infect. Disorders Drug Targets 2018, 18, 15-22.

15. Soriano, V., Vispo, E., Poveda, E., Labarga, P., Martin-Carbonero, L., Fernandez-Montero, J. V., Barreiro, P. Directly acting antivirals against hepatitis C virus. J. Antimicrob. Chemother. 2011, 66, 1673-1686.

16. Shahid, I., AlMalki, W. H., Hassan, S., Hafeez, M. H. Real-world challenges for hepatitis C virus medications: a critical review. Critical Rev. Microbiol. 2017, 44, 143-160.

17. Webster, R. G., Govorkova, E. A. Continuing challenges in influenza. Annals New York Acad. Sci. 2014, 1323, 115-139.

18. Toots, M., Plemper, R. K. Next-generation direct-acting influenza therapeutics. Translational Res. 2020, 220, 33-42.

19. Eastman, R. T., Roth, J. S., Brimacombe, K. R., Simeonov, A., Shen, M., Patnaik, S., Hall, M. D. Remdesivir: a review of its discovery and development leading to emergency use authorization for treatment of COVID-19. ACS Centr. Sci. 2020, 6, 672-683.

20. Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes,

E. C.; Zhang, Y. Z., A new coronavirus associated with human respiratory disease in China. *Nature* 2020, *579* (7798), 265-269.

21. Xu, J., Zhao, S., Teng, T., Abdalla, A. E., Zhu, W., Xie, L., Wang, Y., Guo, X. Systematic comparison of two animal-to-human transmitted human coronaviruses: SARS-CoV-2 and SARS-CoV. Viruses 2020, 12, 244.

22. Gorbalenya, A. E., Snijder, E. J. Viral cysteine proteases. Perspect. Drug Discov. Des. 1996, 6, 64-86.

23. Tong, L. Viral proteases. Chem. Rev. 2002, 102, 4609-4626.

24. Douangamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. Nat. Commun. 2020, 11, 5047.

25. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M^{pro} from COVID-19 virus and discovery of its inhibitors. *Nature* 2020, 582, 289-293.

26. Dai, W., Zhang, B., Jiang, X.-M., Su, H., Li, J., Zhao, Y., Xie, X., Jin, Z., Peng, J., Liu, F., Li, C., Li, Y., Bai, F., Wang, H., Cheng, X., Cen, X., Hu, S., Yang, X., Wang, J., Liu, X., Xiao, G., Jiang, H., Rao, Z., Zhang, L.-K., Xu, Y., Yang, H., Liu, H. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science 2020, 368, 1331-1335.

27. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* 2020, 368, 409-412.

28. Hoffman, R. L., Kania, R. S., Brothers, M. A., Davies, J. F., Ferre, R. A. et al. Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. J. Med. Chem. 2020, 63, 12725-12747.

29. Yang, K. S.; Ma, X. R.; Ma, Y.; Alugubelli, Y. R.; Scott, D. A.; Vatansever, E. C.; Drelich, A. K.; Sankaran, B.; Geng, Z. Z.; Blankenship, L. R.; Ward, H. E.; Sheng, Y. J.; Hsu, J. C.; Kratch, K. C.; Zhao, B.; Hayatshahi, H. S.; Liu, J.; Li, P.; Fierke, C. A.; Tseng, C. K.; Xu, S.; Liu, W. R. A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors. *ChemMedChem* **2020**. https://doi.org/10.1002/cmdc.202000924.

30. Oerlemans, R.; Ruiz-Moreno, A. J.; Cong, Y.; Dinesh Kumar, N.; Velasco-Velazquez, M. A.; Neochoritis, C. G.; Smith, J.; Reggiori, F.; Groves, M. R.; Dömling, A. Repurposing the HCV NS3–4A Protease Drug Boceprevir as COVID-19 Therapeutics. *RSC Med. Chem.* **2021**. https://doi.org/10.1039/D0MD00367K.

31. Ma, C., Sacco, M. D., Hurst, B., Townsend, J. A., Hu, Y., Szeto, T., Zhang, X., Tarbet, B., Marty, M. T., Chen, Y., Wang, J. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. Cell Res. 2020, 30, 678-692.

32. Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N. et al. Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. Nature 2020, 586, 113-119.

33. Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., Zhu, Y., Zhu, C., Hu, T., Du, X., Duan, Y., Yu, J., Yang, X., Yang, X., Yang, K., Liu, X., Guddat, L.W., Xiao, G., Zhang, L., Yang, H., Rao, Z. Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. Nat. Struct. Mol. Biol. 2020, 27, 529-532.

34. Fu, L., Ye, F., Feng, Y., Yu, F., Wang, Q., Wu, Y., Zhao, C., Sun, H., Huang, B., Niu, P., Song, H., Shi, Y., Li, X., Tan, W., Qi, J., Gao, G. F. Both boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. Nat. Commun. 2020, 11, 4417.

35. Ma, C., Hu, Y., Townsend, J. A., Lagarias, P. I., Marty, M. T., Kolocouris, A., Wang, J. Ebselen, disulfiram, carmofur, PX-12, tideglusib, and shikonin are nonspecific promiscuous SARS-CoV-2 main protease inhibitors. ACS Pharmacol. Translat. Sci. 2020, *https://dx.doi.org/10.1021/acsptsci.0c00130*

36. Konno, S., Thanigaimalai, P., Yamamoto, T., Nakada, K., Kakiuchi, R., Takayama, K., Yamazaki, Y., Yakushiji, F., Akaji, K., Kiso, Y., Kawasaki, Y., Chen, S.-E., Freire, E., Hayashi, Y. Design and

synthesis of new tripeptide-type SARS-CoV 3CL protease inhibitors containing an electrophilic arylketone moiety. Bioorg. Med. Chem. 2013, 21, 412-424.

37. Jacobs, J., Grum-Tokars, V., Zhou, Y, Turlington, M., Saldanha, S. A. et al. Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridine-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. J. Med. Chem. 2013, 56, 534-546.

38. Pillaiyar, T., Manickam, Namasivayam, V., Hayashi, Y., Jung, S.-H. An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. J. Med. Chem. 2016, 59, 6595-6628.

39. Liu, Y., Liang, C., Xin, L., Ren, X., Tian, L., Ju, X., Li, H., Wang, Y., Zhao, Q., Liu, H., Cao, W., Xie, X., Zhang, D., Wang, Y., Jian, Y. The development of coronavirus 3C-like protease ($3CL^{pro}$) inhibitors from 2010 to 2020. Eur. J. Med. Chem. 2020, 206, 112711.

40. Anand, K., Ziebuhr, Wadhwani, P., Mesters, J. R., Hilgenfeld, R. Coronavirus main protease ($3CL^{pro}$) structure: basis for design of anti-SARS drugs. Science 2003, 300, 1763-1767.

41. Solowiej, J., Thomson, J. A., Ryan, K., Luo, C., He, M., Lou, J., Murray, B. W. Steady-state and pre-steady-state kinetic evaluation of severe acute respiratory syndrome coronavirus (SARS-CoV) 3CLpro cysteine protease: development of an ion-pair model for catalysis. Biochemistry 2008, 47, 2617-2630.

42. Hsu, W.-C., Chang, H.-C., Chou, C.-Y., Tsai, P.-J., Lin, P.-I., Chang, G.-G. Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. J. Biol. Chem. 2005, 280, 22741-22748.

43. Lin, P.-Y., Chou, C.-Y., Chang, H.-C., Hsu, W.-C., Chang, G.-G. Correlation between dissociation and catalysis of SARS-CoV main protease. Arch. Biochem. Biophys. 2008, 472, 34-42.

44. Menard, R., Storer, A. C. Oxyanion hole interactions in serine and cysteine protease. Biol. Chem. Hoppe-Sayler 1992, 373, 393-400.

45. D.W. Kneller, G. Phillips, H.M. O'Neill, R. Jedrzejczak, L. Stols, P. Langan, A. Joachimiak, L. Coates, A. Kovalevsky Structural plasticity of SARS-CoV-2 3CL M^{pro} active site cavity revealed by room temperature X-ray crystallography. *Nat. Commun.* 2020, **11**, 3202.

46. D.W. Kneller, G. Phillips, K.L. Weiss, S. Pant, Q. Zhang, H.M. O'Neill, L. Coates, A. Kovalevsky. Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *J. Biol. Chem.* 2020, 295, 17365-17373.

47. Niimura, N. and Podjarny, A. (2011) Neutron protein crystallography. Oxford University Press. 250 pp.

48. Neutron Crystallography in Structural Biology, Methods in Enzymology 2020, Ed. P.C.E. Moody, Academic Press, 634, 389pp.

49. O. Gerlits, K.L. Weiss, M.P. Blakeley, G. Veglia, S.S. Taylor, A. Kovalevsky (2019) Zooming in on protons: neutron structure of protein kinase A trapped in a product complex. *Science Adv.* **5**, eaav0482.

50. P.S. Langan, D.W. Close, L. Coates, R.C. Rocha, K. Ghosh, C. Kiss, G. Waldo, J. Freyer, A. Kovalevsky, A.R.M. Bradbury (2016) Evolution and characterization of a new reversibly photoswitching chromogenic protein, Dathail. *J. Mol. Biol.* **428**, 1776-1789.

51. M.T. Banco, V. Mishra, A. Ostermann, T.E. Schrader, G.B. Evans, A.Y. Kovalevsky, D.R. Ronning (2016) Neutron structures of the Helicobacter pylori 5'-methylthioadenosine nucleosidase highlight proton sharing and protonation states. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13756-13761.

52. Garman, E. F. Radiation damage in macromolecular crystallography: what is it and why should we care? Acta Cryst. 2010, D66, 339-351.

53. Blakeley, M. P., Mitschler, A., Hazemann, I., Meilleur, F., Myles, D. A. A., Podjarny, A. Comparison of hydrogen determination with X-ray and neutron crystallography in a human aldose reductase-inhibitor complex. Eur. Biophys. J. 2006, 35, 577-583.

54. Gardberg, A. S., Del Castillo, A. R., Weiss, K. L., Meilleur, F., Blakeley, M. P., Myles, D. A. A. Unambiguous determination of H-atom positions: comparing results from neutron and high-resolution X-ray crystallography. Acta Cryst. D 2010, 66, 558-567.

55. Bax, B., Chung, C.-w., Edge, C. Getting the chemistry right: protonation, tautomers and the importance of H atoms in biological chemistry. Acta Crystallogr. D 2017, 73, 131-140.

56. A. Pavlova, D.L. Lynch, I. Daidone, L. Zanetti-Polzi, M.D. Smith, C. Chipot, D.W. Kneller, A. Kovalevsky, L. Coates, A.A. Golosov, C.J. Dickson, C. Velez-Vega, J.S. Duca, J.V. Vermaas, Y.T. Pang, A. Acharya, J.M. Parks, J.C. Smith, J. Gumbart (2020) Inhibitor binding influences the protonation states of histidines in SARS-CoV-2 main protease. *Chem. Sci.*, In Press. (doi:10.1039/D0SC04942E)

57. D.W. Kneller, S. Galanie, G. Phillips, H.M. O'Neill, L. Coates, A. Kovalevsky (2020) Malleability of SARS-CoV-2 3CL M^{pro} active site cavity facilitates binding of clinical antivirals. *Structure* **28**, 1313-1320.

58. Adams, P. D.; Mustyakimov, M.; Afonine, P. V.; Langan, P. Generalized X-ray and Neutron Crystallographic Analysis: More Accurate and Complete Structures for Biological Macromolecules. *Acta Cryst. D* **2009**, *65*, 567-573.

59. Kumar, P., Serpersu, E. H., Cuneo, M. J. A low-barrier hydrogen bond mediates antibiotic resistance in a noncanonical catalytic triad. *Sci. Adv.* 2018, 4, eaas8667.

60. A. Kovalevsky, M. Aggarwal, H. Velazquez, M.J. Cuneo, M.P. Blakeley, K.L. Weiss, J.C. Smith, S.Z. Fisher, R. McKenna (2018) 'To be or not to be' protonated: atomic details of human carbonic anhydrase-clinical drug complexes by neutron crystallography and simulation. *Structure* **26**, 383-390.e3.

61. Rathayake, A. D., Zheng, J., Kim, Y., Perera, K. D., Mackin, S., Meyerholz, D. K., Kashipathy, M. M., Battaile, K. P., Lovell, S., Perlman, S., Groutas, W. C., Chang, K.-O. 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-infected mice. Science Transl. Med. 2020, 12, eabc5332.

62. D.W. Kneller, G. Phillips, A. Kovalevsky, L. Coates (2020) Room-temperature neutron and X-ray data collection of 3CL M^{pro} from SARS-CoV-2. *Acta Cryst.* **F76**, 483-487.

63. Luft, J. R., Collins, R. J., Fehrman, N. A., Lauricella, A. M., Veatch, C. K. & DeTitta, G. T. (2003). *J Struct Biol.* **142**, 170-179.

64. F. Meilleur, A. Kovalevsky, D.A.A. Myles (2020) IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. *Methods Enzymol.* **634**, 69-85.

65. F. Meilleur, L. Coates, M.J. Cuneo, A. Kovalevsky, D.A.A. Myles (2018) The neutron macromolecular crystallography instruments at Oak Ridge National Laboratory: Advances, challenges, and opportunities. *Crystals* **8**, 388.

66. F. Meilleur, P. Munshi, L. Robertson, A.D. Stoica, L. Crow, A. Kovalevsky, T. Koritsanszky, B.C. Chakoumakos, R. Blessing, D.A.A. Myles (2013) The IMAGINE instrument: first neutron protein structure and new capabilities for neutron macromolecular crystallography. *Acta Cryst.* **D69**, 2157-2160.

67. L. Coates, H.B. Cao, B.C. Chakoumakos, M.D. Frontzek, C. Hoffmann, A.Y. Kovalevsky, Y. Liu, F. Meilleur, A.M. dos Santos, D.A.A. Myles, X.P. Wang, F. Ye (2018) A suite-level review of the neutron single-crystal diffraction instruments at Oak Ridge National Laboratory. *Rev. Sci. Instrum.* **89**, 092802.

68. L. Coates, B. Sullivan (2020) The macromolecular neutron diffractometer at the spallation neutron source. *Methods Enzymol.* **634**, 87-99.

69. Coates, L., Cuneo, M. J., Frost, M. J., He, J., Weiss, K. L., Tomanicek, S. J., McFeeters, H., Vandavasi, V. G., Langan, P., Iversion, E. B. The macromolecular neutron diffractometer MaNDi at the Spallation Neutron Source. *J. Appl. Cryst.* 2015, 48, 1302-1306.

70. Arnold, O., Bilheux, J. C., Borreguero, J. M., Buts, A., Campbell, S. I., Chapon, L., Doucet, M., Draper, N., Leal, R. F., Gigg, M. A., Lynch, V. E., Markyardsen, A., Mikkelson, D. J., Mikkelson, R. L., Miller, R., Palmen, K., Parker, P., Passos, G., Perring, T. G., Peterson, P. F., Ren, S., Reuter, M.

A., Sayici, A. T., Taylor, J. W., Taylor, R. J., Tolchenoy, R., Zhou, W. & Zikoysky, J. (2014). Nucl Instrum Meth A 764, 156-166.

71. Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. (1998). J Appl Crystallogr 31, 496-502.

72. Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). J Appl Crystallogr 22, 483-497.

73. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallographica Section D* **2013,** *69* (7), 1204-1214.

74. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **2011,** *67*, 235-242.

75. 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.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

76. Mustyakimov, M.; Langan, P. Copyright C-06, 104 Patch for CNS; nCNS an Open Source Distribution Patch for CNS for Macromolecular Structure Refinement (Los Alamos National Security, Los Alamos, NM, USA) **2007**.

77. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.

Table 1. Protonation states and the corresponding electric charges of the ionizable residues in SARS-CoV-2 M^{pro} as observed in the neutron structures of the ligand-free and telaprevir-bound enzyme.

Residue	**M^{pro} ligand-free** (PDB ID 7JUN)		**M^{pro}-Telaprevir** (PDB ID 7LB7)	
	Charge	Species	Charge	Species
$Cys145_{cat}$	–1	Thiolate (S^-)	0	S-C-OD (hemithioketal)
$His41_{cat}$	+1	Nδ1-D, Nε2-D	0	Nδ1-D
His163	0	Nδ1-D	+1	Nδ1-D, Nε2-D
His164	+1	Nδ1-D, Nε2-D	0	Nδ1-D
His172	0	Nε2-D	0	Nε2-D
Net charge	**+1**		**+1**	

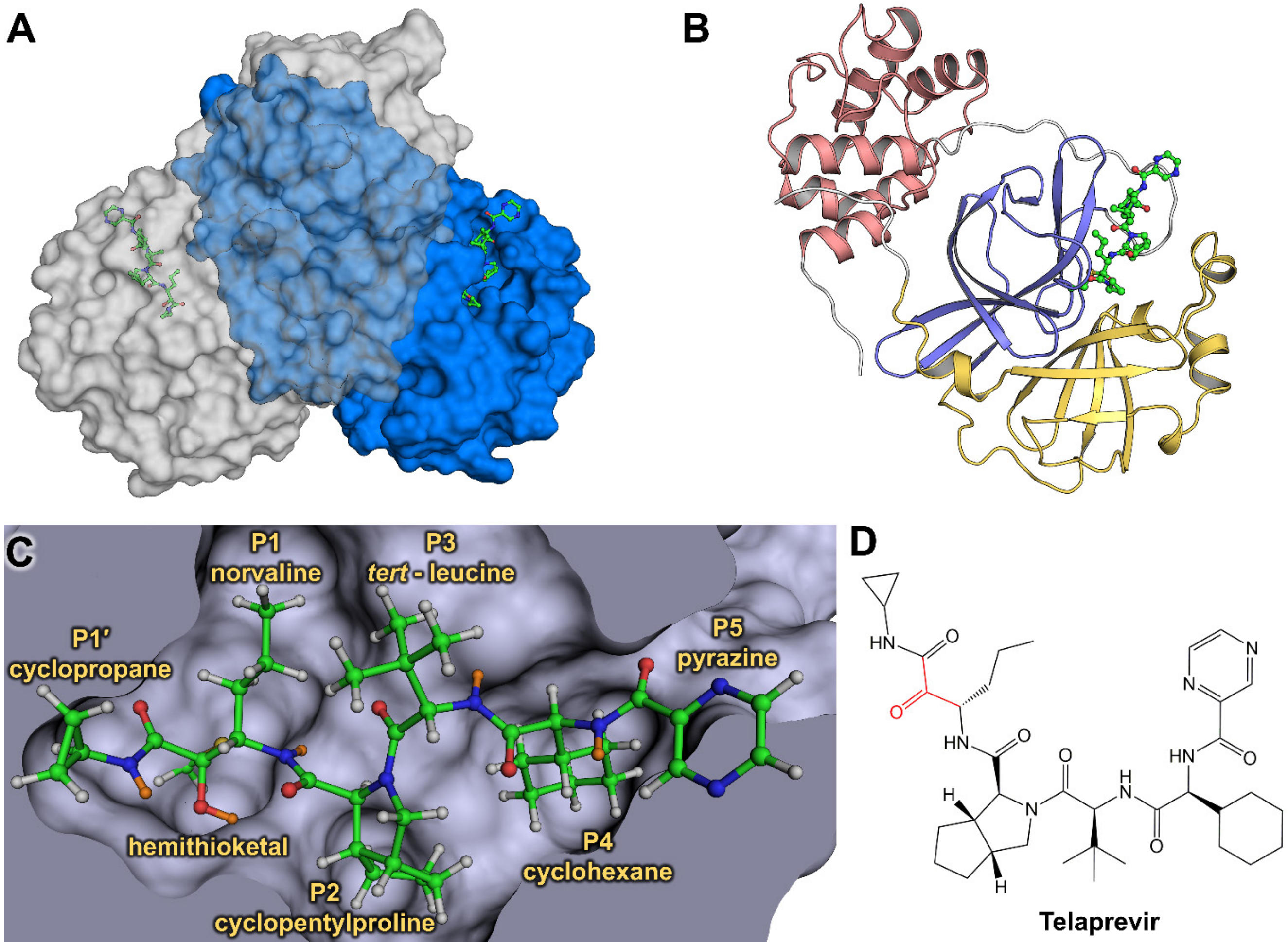
A
B
C
P1 norvaline
P3 *tert* - leucine
P1′ cyclopropane
P5 pyrazine
hemithioketal
P4 cyclohexane
P2 cyclopentylproline
D
HN
O
O
HN
O
N
N
O
HN
O
H
N
H
N
O
H
Telaprevir

Figure 1. Joint X-ray/neutron structure of SARS-CoV-2 3CL M^{pro} and binding of hepatitis C clinical protease inhibitor telaprevir. (**A**) The catalytically active dimer is shown in surface representation, with telaprevir shown in ball-and-stick representation. (**B**) One enzyme protomer is shown in cartoon representation colored according to the domain structure – domain I is yellow, domain II is blue, and domain III is dark pink. (**C**) Active site cavity with the covalently bound telaprevir. H atoms are shown in gray, whereas D atoms are colored orange. (**D**) A chemical diagram of telaprevir with the ketone warhead shown in red.

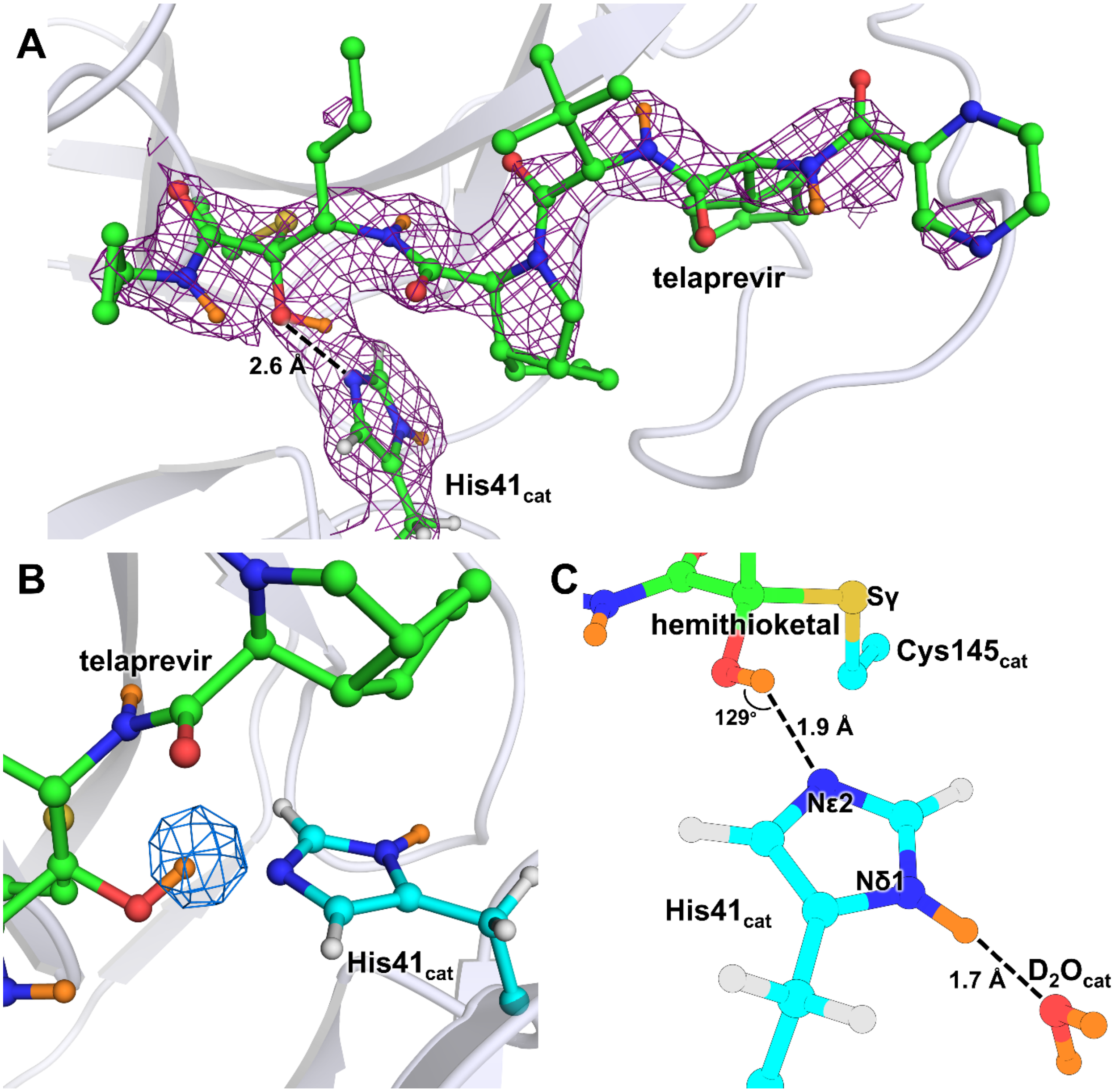
A
telaprevir
2.6 Å
$His41_{cat}$
B
telaprevir
$His41_{cat}$
C
hemithioketal
Sγ
$Cys145_{cat}$
129°
1.9 Å
Nε2
Nδ1
$His41_{cat}$
1.7 Å
D_2O_{cat}

Figure 2. The SARS-CoV-2 M^{pro} catalytic dyad forms a hemithioketal with telaprevir that possesses a protonated hydroxyl. (**A**) The $2F_O$-F_C nuclear density map of telaprevir and the catalytic His41 contoured at a level of 1.0 σ (violet mesh) in the SARS-CoV-2 M^{pro} active site cavity. (**B**) The omit nuclear density map contoured at 3.5 σ level (blue mesh) clearly indicates the hydroxyl of the hemithioketal is protonated. (**C**) The hydrogen bonding interactions involving His41 are shown in detail. H atoms on telaprevir are omitted for clarity, whereas D atoms are colored orange.

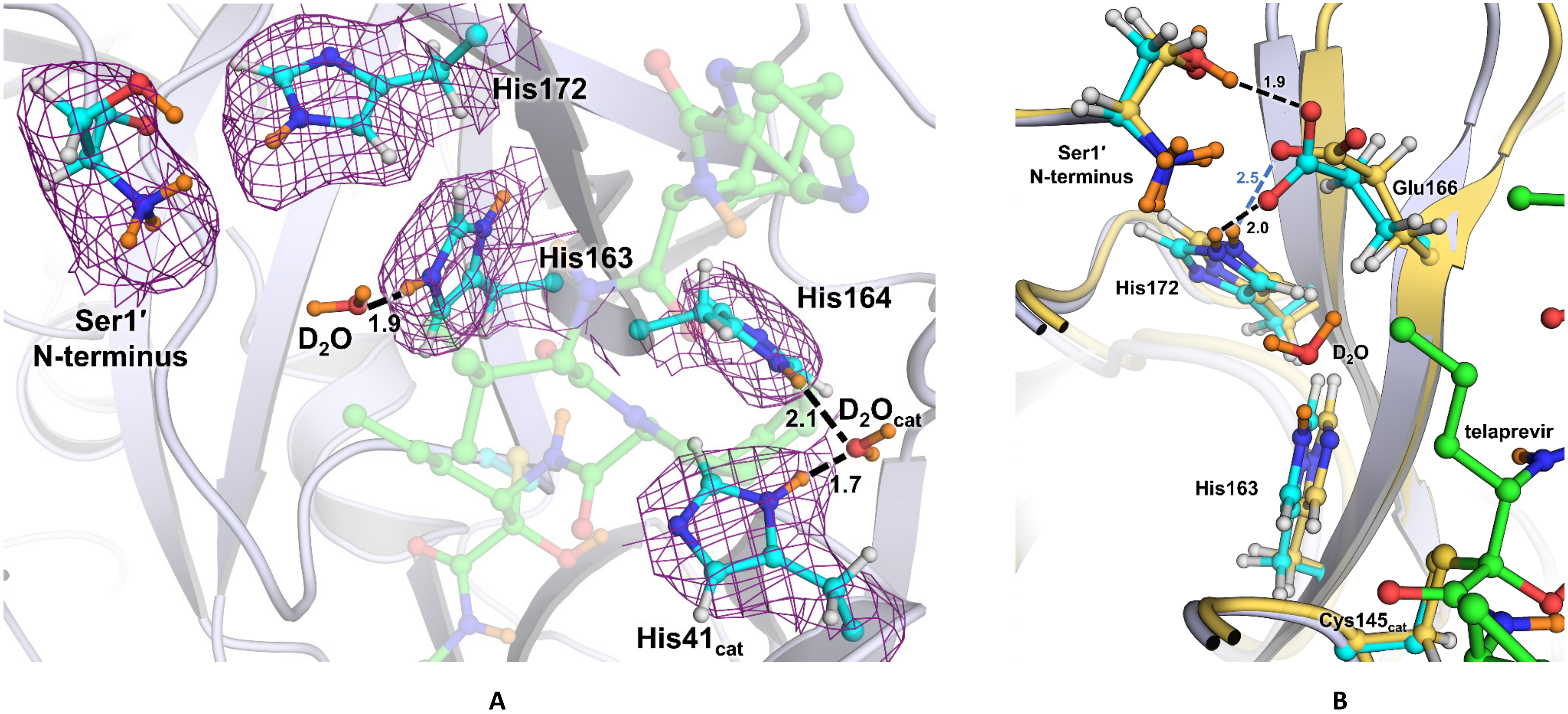
His172
Ser1′
N-terminus
His163
D$_2$O
1.9
His164
2.1
D$_2$O$_{cat}$
1.7
His41$_{cat}$
A
1.9
Ser1′
N-terminus
2.5
Glu166
2.0
His172
D$_2$O
telaprevir
His163
Cys145$_{cat}$
B

Figure 3. (**A**) Ionizable residues of the active site and S1 subsite of the M^{pro}-telaprevir complex. A $2F_O$-F_C nuclear density map of Ser1′ N-terminus and histidines 41,163, 164, and 172 is shown contoured at 1.0 σ level (violet mesh). All distances are shown in Ångstroms with telaprevir being transparent for clarity. (**B**) Superposition of M^{pro}-telaprevir and apo-M^{pro} (PDB ID 7JUN) neutron structures showing S1 subsite. Distances are in Å.