IUCrJ

Michaelis-like complex of SARS-CoV-2 main protease visualized by room-temperature X-ray crystallography

Daniel W. Kneller,[a,b] Qiu Zhang,[a,b] Leighton Coates,[b,c] John M. Louis[d] and Andrey Kovalevsky[a,b]*

[a]Neutron Scattering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA, [b]National Virtual Biotechnology Laboratory, US Department of Energy, Washington, DC 20585, USA, [c]Second Target Station, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA, and [d]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, MD 20892-0520, USA. *Correspondence e-mail: kovalevskyay@ornl.gov

SARS-CoV-2 emerged at the end of 2019 to cause an unprecedented pandemic of the deadly respiratory disease COVID-19 that continues to date. The viral main protease (M^{pro}) is essential for SARS-CoV-2 replication and is therefore an important drug target. Understanding the catalytic mechanism of M^{pro}, a cysteine protease with a catalytic site comprising the noncanonical Cys145–His41 dyad, can help in guiding drug design. Here, a 2.0 Å resolution room-temperature X-ray crystal structure is reported of a Michaelis-like complex of M^{pro} harboring a single inactivating mutation C145A bound to the octapeptide Ac-SAVLQSGF-CONH$_2$ corresponding to the nsp4/nsp5 autocleavage site. The peptide substrate is unambiguously defined in subsites S5 to S3′ by strong electron density. Superposition of the Michaelis-like complex with the neutron structure of substrate-free M^{pro} demonstrates that the catalytic site is inherently pre-organized for catalysis prior to substrate binding. Induced fit to the substrate is driven by P1 Gln binding in the predetermined subsite S1 and rearrangement of subsite S2 to accommodate P2 Leu. The Michaelis-like complex structure is ideal for *in silico* modeling of the SARS-CoV-2 M^{pro} catalytic mechanism.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causes the deadly coronavirus disease 2019 (COVID-19) that engulfed the globe in a pandemic in early 2020 and remains a significant health threat despite the availability of several vaccines and antiviral therapies (Gavor *et al.*, 2020; Meo *et al.*, 2021; Doroftei *et al.*, 2021; Bidram *et al.*, 2021). The main protease of SARS-CoV-2 (3CL^{pro} or M^{pro}) is virally encoded and its precise function is essential for virus replication. M^{pro} cleaves two polyproteins, pp1a and pp1ab (~450 and ~790 kDa, respectively), into individual functional proteins that build the virus genome replication/transcription machinery (Wu *et al.*, 2020; Xu *et al.*, 2020). M^{pro} is a cysteine protease (Tong, 2002) that hydrolyses scissile peptide bonds at 11 specific locations, including its own release as the non-structural protein nsp5 (termed autoprocessing), in the polyprotein precursors. The substrate specificity of M^{pro} and its active-site architecture are dissimilar relative to human proteases. Thus, M^{pro} is considered an important drug target for the design and development of specific antivirals with potentially minimal nonspecific binding to human proteins (Liu *et al.*, 2020; Jin *et al.*, 2020; Ton *et al.*, 2020; Suárez & Díaz,

     973
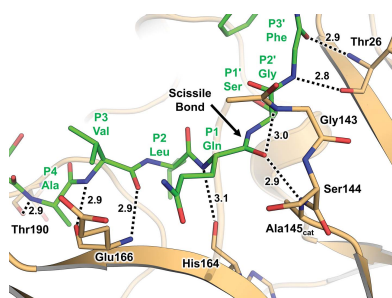
2020; Acharya *et al.*, 2020; Pathak *et al.*, 2021), including the possibility of designing pan-coronavirus antivirals (Ullrich & Nitsche, 2020).

M^{pro} is a homodimer, with each protomer composed of three distinct domains (Zhang *et al.*, 2020; Jin *et al.*, 2020; Anand *et al.*, 2002). The catalytic domains I (residues 1–101) and II (residues 102–200) have antiparallel β-barrel structures that resemble the chymotrypsin fold. The helical domain III (residues 201–306) comprises five α-helices and promotes dimerization through interactions with specific residues of domains I and II (Chou *et al.*, 2004; Barrila *et al.*, 2006; Tsai *et al.*, 2010). The catalytic site of M^{pro} is noncanonical; it has a catalytic Cys145–His41 dyad believed to be assisted by a conserved (catalytic) water molecule (Wang *et al.*, 2020), instead of the more common (Ser)Cys–His–Asp(Glu) triads found in other cysteine (and serine) proteases (Gorbalenya & Snijder, 1996). The M^{pro} cleavage sites are characterized by a preference for Gln and Leu/Phe at substrate positions P1 and P2, respectively (Fan *et al.*, 2004; Goetz *et al.*, 2007). Substrate specificity is conserved in most coronavirus M^{pro} cleavage sites (Lai *et al.*, 2006). Based on the crystal structures of the closely related SARS-CoV M^{pro} inactive mutant H41A in complex with an 11-amino-acid peptide substrate (Xue *et al.*, 2008) and SARS-CoV M^{pro} inactive mutant C145A containing a ten-residue C-terminal prosequence (Muramatsu *et al.*, 2016), the M^{pro} active-site cavity located on the protein surface can accommodate up to nine substrate residues in positions P5 through P4′ in the corresponding substrate-binding subsites S5 through S4′.

To date, most of the SARS-CoV-2 M^{pro} X-ray crystal structures deposited in the Protein Data Bank (PDB) and associated studies have focused on the design of protease inhibitors. Conversely, there is an apparent lack of structural studies into the catalytic mechanism of SARS-CoV-2 M^{pro}. Lee *et al.* (2020) have recently reported capturing a covalent acyl-enzyme intermediate using native enzyme and a product complex using the inactive C145A mutant. In both structures, the C-terminal residues of one protomer extend into the active site of the other protomer, thus mimicking autocleavage of its C-terminus. Consequently, a Michaelis-like complex structure of SARS-CoV-2 M^{pro} bound to a discrete peptide substrate remains unidentified. This paper reports a room-temperature X-ray crystal structure of SARS-CoV-2 M^{pro} C145A mutant ($M^{pro/C145A}$) in complex with the octapeptide Ac-SAVLQSGF-$CONH_2$ corresponding to the nsp4/nsp5 autocleavage site (henceforth referred to as $M^{pro/C145A}$–substrate), at 2.0 Å resolution and pH 6.5. The peptide substrate spans residues from P5 through P3′ and is chemically protected by acetylation at the N-terminus and amidation at the C-terminus. The octapeptide binds asymmetrically in the M^{pro} active-site cavity such that the five groups P5–P1 bind to the nonprime substrate-binding subsites S5 through S1 and the three groups P1′–P3′ bind to the prime substrate-binding subsites S1′ through S3′. We hypothesize that this substrate mode of binding is the reason behind our success in producing crystals of such a complex. We previously reported enzyme kinetics measurements for a FRET substrate based on the same nsp4/

nsp5 autocleavage site (k_{cat} = 0.28 s^{-1} and K_m = 170 μM; Kneller, Galanie *et al.*, 2020).

2. Materials and methods

2.1. General information

Crystallization reagents and supplies were purchased from Hampton Research, Aliso Viejo, California, USA. Crystallographic supplies for crystal mounting and X-ray diffraction data collection at room temperature were purchased from MiTeGen, Ithaca, New York, USA. The Ac-SAVLQSGF-$CONH_2$ peptide used in this study was custom-synthesized by Biomatik Corp., Kitchener, Ontario, Canada.

2.2. Expression and purification of $M^{pro/C145A}$

A DNA insert encoding a 6×His tag followed by a TEV protease cleavage site and SARS CoV-2 M^{pro} (306 amino acids) bearing an active-site C145A mutation was cloned into the pJ414 vector (ATUM) and transformed into *Escherichia coli* BL21(DE3) cells (Agilent). The last residue in the TEV protease recognition sequence (ENLYFQG/S) was chosen to be a serine, such that upon cleavage the $M^{pro/C145A}$ construct starts with the native N-terminal Ser1 residue. Cells were grown in Luria–Bertani medium, and the $M^{pro/C145A}$ fusion protein was expressed using established protocols. The fusion protein was purified under native conditions by lysis and Ni–NTA affinity chromatography in 20 mM Tris–HCl pH 8, 150 mM NaCl, 1 mM TCEP, 20 mM imidazole. This step was followed by size-exclusion chromatography on a Superose 12 column (16 × 60 cm, Cytiva) to remove minor high-molecular-weight contaminants and to exchange the buffer to 25 mM Tris–HCl pH 8, 100 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 20 mM imidazole, which is suitable for cleavage by TEV protease at a molar ratio of 75:1 fusion protein:TEV protease. TEV protease was prepared as described by Lucast *et al.* (2001). Upon completion of the cleavage and adjustment of the EDTA concentration to 0.25 mM, $M^{pro/C145A}$ was collected in the flowthrough from the Ni–NTA column, concentrated and then subjected again to isocratic fractionation on Superose 12 in 25 mM Tris pH 7.6, 150 mM NaCl, 1 mM TCEP. Peak fractions were combined, concentrated and aliquots of the protein were stored at − 20 C. Purity was verified both by SDS–PAGE and electrospray ionization mass spectrometry.

2.3. Crystallography of the $M^{pro/C145A}$–substrate complex

We have previously published detailed methods for crystallizing high-quality M^{pro} crystals using the sitting-drop vapor-diffusion methodology (Kneller, Galanie *et al.*, 2020; Kneller, Phillips, Kovalevsky *et al.*, 2020). Crystallization conditions for the $M^{pro/C145A}$ mutant were discovered by automated high-throughput screening at the Hauptman–Woodward Medical Research Institute (Luft *et al.*, 2003). The protein was concentrated to 6 mg ml^{-1}. Initially, flower-like crystal aggregates unsuitable for room-temperature X-ray diffraction grew using a sample of 270 μl protein solution mixed with 1.6 mg lyophilized peptide in 20% PEG 3350,

0.1 M bis-Tris pH 7.0. Subsequently, the protein was mixed in a 1:5 molar ratio with peptide from a 50 mM stock in 100% DMSO and used for crystallization with 20% PEG 3350, 0.1 M bis-Tris pH 6.5. This condition produced a final pH in the crystallization drop of 6.5 as measured by a microelectrode. Drops were seeded using a cat whisker and flower-shaped aggregates grew with thicker 'petal' substituents after one week of incubation at 14°C. A single crystal from the aggregate was harvested and mounted on loops and capillary setups for room-temperature diffraction.

2.4. Room-temperature X-ray diffraction data collection and structure refinement

Room-temperature X-ray diffraction data were collected on a Rigaku HighFlux HomeLab instrument equipped with a MicroMax-007 HF X-ray generator, Osmic VariMax optics and a Dectris EIGER R 4M hybrid photon-counting detector. Data were integrated using the *CrysAlis*^{Pro} software suite (Rigaku, The Woodlands, Texas, USA) and were reduced and scaled using *AIMLESS* (Evans & Murshudov, 2013) from the *CCP*4 suite (Winn *et al.*, 2011). The M$^{pro/C145A}$–substrate structure was solved by molecular replacement using PDB entry 6wqf (Kneller, Phillips, O'Neill *et al.*, 2020) with *Phaser* (McCoy *et al.*, 2007) from *CCP*4. The model was iteratively refined using *phenix.refine* from the *Phenix* suite (Liebschner *et al.*, 2019) and *Coot* (Emsley *et al.*, 2010; Casañal *et al.*, 2020) aided by *MolProbity* (Chen, Arendall *et al.*, 2010) for geometry validation. Final data-collection and refinement statistics are given in Supplementary Table S1.

3. Results and discussion

3.1. Binding of peptide substrate

In M$^{pro/C145A}$–substrate, a substrate occupies each of the two active-site grooves in the M^{pro} homodimer structure (Fig. 1). There is unambiguous electron density for all residues except for the P3′ Phe side chain. Interactions between M^{pro} and substrate are nearly identical in both protomers; hence, subsequent details of our analysis will focus on protomer *A*. Main-chain interactions between the peptide and M^{pro} are consistent with adding a β-strand to the antiparallel β-sheets flanking the scissile bond [Fig. 2(*a*)]. There are eight hydrogen bonds between the main chains of the peptide substrate and M^{pro}, which involve residues Thr190, Glu166, His164, Ser144, Gly143 and Thr26. In the catalytic site, P1 Gln is driven into the S1 subsite via a 2.7 Å hydrogen bond to the His163 imidazole and a water-mediated interaction with Asn142, whereas its main-chain carbonyl O atom is centered into the characteristic oxyanion hole [Fig. 2(*b*)]. The bulky hydrophobic P2 Leu side chain fills its respective hydrophobic S2 subsite, whereas the smaller P4 Ala juts into the shallow S4 pocket capped by the flexible S5 loop. P2′ Gly contributes to substrate binding via backbone hydrogen-bond interactions with Thr26, whereas P1′ Ser makes no hydrogen bonds to M^{pro}, as is the case for P5 Ser. The P5 Ser, P3 Val, P2′ Ser and P3′ Phe side chains face the solvent space between crystallo-

graphic symmetry-related molecules, with the P3′ Phe phenyl ring being dynamic and poorly defined in the electron-density maps.

The C-terminal residues of M^{pro} flank the last helix in domain III. Here, residues 302–306 constitute a flexible loop (Kneller, Phillips, O'Neill *et al.*, 2020) and are typically disordered in crystal structures of SARS-CoV-2 M^{pro} to the
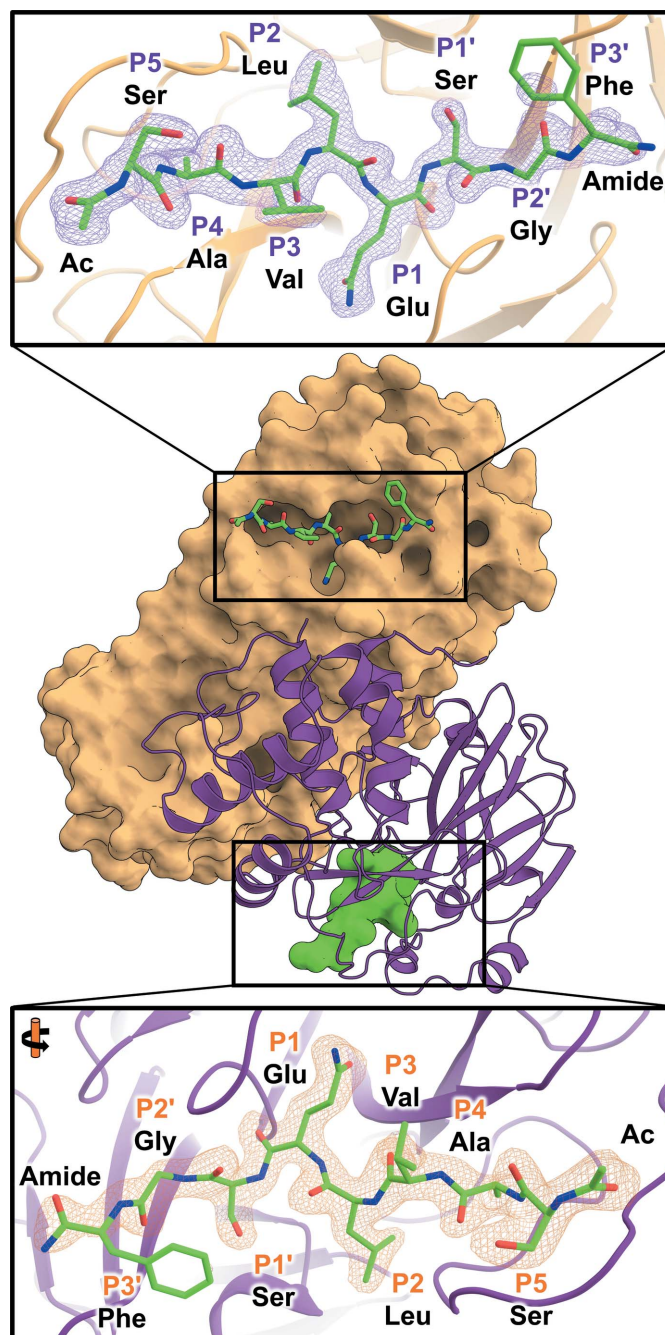

Figure 1
M$^{pro/C145A}$–substrate complex. M$^{pro/C145A}$ (protomer *A* is shown as a purple cartoon and protomer *B* as an orange surface) was co-crystallized with an octapeptide (green sticks and surface) corresponding to the nsp4/nsp5 autocleavage site. The top inset shows the peptide in protomer *B* in conventional orientation, while the bottom inset shows protomer *A* rotated by 180°. $F_o - F_c$ omit difference electron-density maps for each peptide are contoured at 3σ.

extent that some crystallographers choose not to model the C-terminus (Dai *et al.*, 2020; Zhang *et al.*, 2020). In this context, it is noteworthy that electron density for the C-terminal residues up to the very last Gln306 amine is clearly observed in $M^{pro/C145A}$–substrate (Supplementary Fig. S1). When the C-terminal residues of $M^{pro/C145A}$–substrate and the product complex reported by Lee *et al.* (2020) are superimposed, the main-chain atoms align with an r.m.s.d. of 0.7 Å for residues 303–306 (Supplementary Fig. S2), demonstrating that the C-terminus of M^{pro} is pre-organized to bind into the active site during M^{pro} autoprocessing.

3.2. Comparison with substrate-free M^{pro}

Superposition of $M^{pro/C145A}$–substrate and our recently determined substrate-free neutron crystallographic structure

of wild-type SARS-CoV-2 M^{pro} (PDB entry 7jun; Kneller, Phillips, Weiss *et al.*, 2020) offers unique insights into catalysis because the Cys145–His41 dyad was observed as a zwitterionic species with the Cys145 thiolate S^{γ} atom 3.8 Å away from His41 $N^{\epsilon 2}$, already primed for proteolysis [Fig. 2(*c*)]. When the two structures are superimposed, the Cys145 thiolate is positioned 1.8 Å away from the scissile-bond carbonyl C atom with an ideal distance and orientation for a nucleophilic attack. The scissile-bond amide N atom is positioned 2.7 Å away from the $N^{\epsilon 2}$ hydrogen on the catalytic His41, which does not significantly change its position in the complex, poised for proton transfer to the amino-terminus that is created in the cleavage reaction. The $M^{pro/C145A}$–substrate structure therefore underscores the suggestion that the zwitterionic catalytic dyad is present prior to substrate binding (Kneller, Phillips, Weiss *et al.*, 2020) instead of requiring reorganization of the catalytic
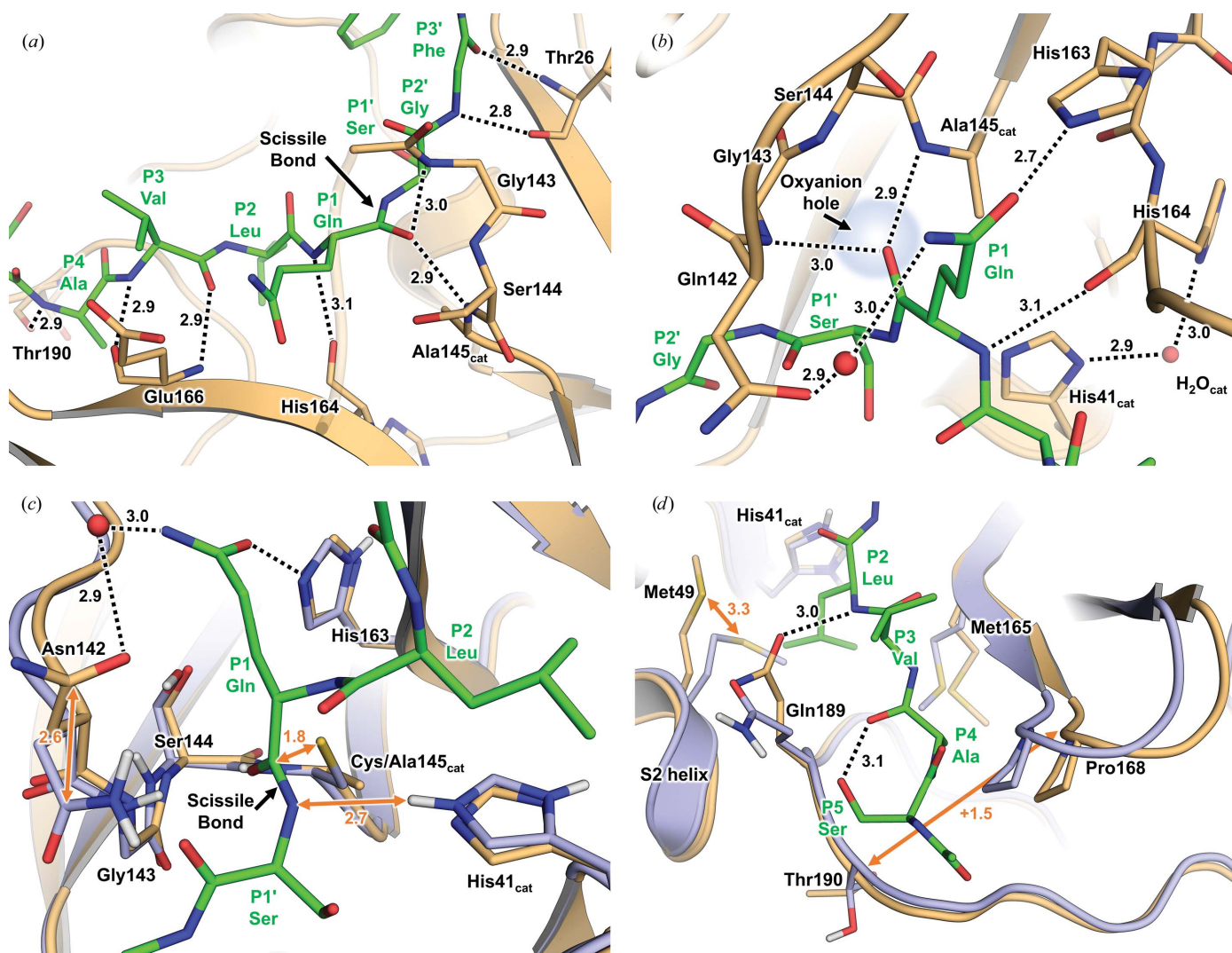

Figure 2
$M^{pro/C145A}$–substrate interactions for induced fit of substrate. (*a*) Main-chain hydrogen-bond interactions between M^{pro} and the peptide substrate (shown as orange and green C atoms, respectively). Hydrogen bonds are shown as dotted lines. Distances are in Å. (*b*) Hydrogen-bond interactions between peptide P1 Gln with the S1 subsite and the catalytic site. Water molecules are shown as red spheres. (*c*) Superposition of $M^{pro/C145A}$–substrate with the substrate-free M^{pro} neutron crystal structure (light blue; PDB entry 7jun) showing the scissile bond in relation to the catalytic Cys145–His41 zwitterion. Superposition was by a least-squares fit to C^{α} atoms of protomer *A*. (*d*) Induced fit of $M^{pro/C145A}$ to the peptide substrate by the rearrangements of side chains in the S2 and S4 subsite.

site to facilitate proton transfer from Cys145 to His41 in the catalytically resting (nonreactive) state (*i.e.* a protonated Cys145 thiol and singly protonated His41 imidazole) to create the reactive zwitterionic species (Paasche *et al.*, 2014; Ramos-Guzmán *et al.*, 2020; Díaz & Suárez, 2021).

In subsite S1, Glu166 accommodates P1 Gln by moving ~0.9 Å away from the substrate and simultaneously rotating to form a 2.8 Å hydrogen bond to His172 that is absent in substrate-free M^{pro}. This conformational change of the Glu166 side chain is due to the binding of a ligand to the active site and is consistent with the changes observed in M^{pro}–inhibitor complexes (Kneller, Galanie *et al.*, 2020; Dai *et al.*, 2020). Meanwhile, the Asn142 side chain flips ~180° and shifts 2.6 Å from its position in the substrate-free structure to produce water-mediated contacts with P1 Gln. This comparison indicates that the catalytic dyad is already pre-organized and primed for catalysis in the substrate-free structure, requiring minimal rearrangements to initiate the peptide-hydrolysis reaction. Induced fit of the peptide substrate to M^{pro} is achieved primarily through the continuation of β-sheet structures and key rearrangements of the large side chains Met165, Met49 and Gln189 from their positions in substrate-free M^{pro} to create subsites S2 and S4 [Fig. 2(*d*)]. The substrate P2 Leu forms the S2 subsite by a >3 Å steric swing of the Met49 side chain located on the short S2 helix (residues 46–50). P2 Leu also acts in consortium with the methyl C atom of P4 Ala to produce the shallow S4 subsite by initiating a new rotational isomer for Met165 and shifting the position of the S4 β-hairpin loop by as much as 1.5 Å. The peptide residue arrangement is stabilized by Gln189 rotating to hydrogen-bond to the main-chain NH of P3 Val. P5 Ser does not interact directly with M^{pro}. However, the hydroxymethyl side chain might act to nullify repulsion of the Gln189 backbone by

intramolecular hydrogen bonding to the carbonyl of P3 Val. The active site around P4 and P5 expands by >1.5 Å from the substrate-free structure to accommodate the substrate, as measured by the C$^\alpha$ distances between Pro168/168' and Thr190/190'. Interestingly, the S2 subsite helix of the complex shows only minor distal expansion, unlike in M^{pro}–inhibitor complexes where the S2 subsite helix can shift by >1.5 Å (Jin *et al.*, 2020; Kneller, Galanie *et al.*, 2020). Similarly, the S5 loop (residues 189–194) is only minimally disturbed by the substrate, shifting by <1.0 Å.

3.3. Comparison with SARS-CoV M$^{pro/H41A}$

It is instructive to compare our room-temperature M$^{pro/C145A}$–substrate structure from SARS-CoV-2 with a low-temperature structure of the SARS-CoV main protease H41A mutant, M$^{pro/H41A}$, soaked with an 11-amino-acid peptide substrate spanning the same nsp4/nsp5 autocleavage site (Xue *et al.*, 2008; Fig. 3; PDB entry 2q6g). The peptide substrate in both structures bows into the active site by conforming to the antiparallel β-sheets flanking the catalytic site. Overall, the orientations of the substrate residues in both protomers are in agreement or can be explained by a lack of electron density, with one exception. The C$^\gamma$ atoms of P2 Leu deviate by ~1 Å, which is coordinated with the rotational isomer differences observed for the Met49 and Met165 side chains. In the SARS-CoV M$^{pro/H41A}$–substrate complex, the substrate pushes the catalytic Cys145 side chain away from the scissile bond towards the vacated space left by the H41A mutation to a position where the S atom is 4.5 Å from the scissile carbonyl C atom. M^{pro} possesses 96% sequence identity between SARS-CoV and SARS-CoV-2. The nearest mutation to the active site
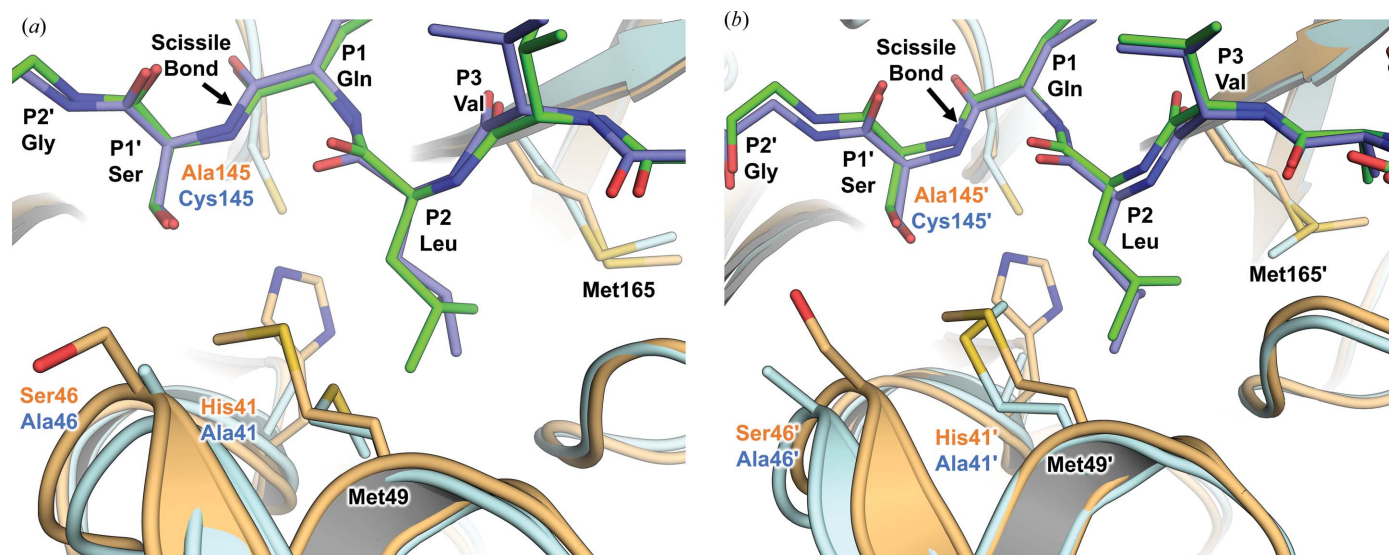

Figure 3
Comparison of substrate-bound M$^{pro/C145A}$ from SARS-CoV-2 (PDB entry 7n89) with M$^{pro/H41A}$ from SARS-CoV (PDB entry 2q6g). (*a*) Protomer *A* in M$^{pro/C145A}$–substrate from SARS-CoV-2 (orange protein, green substrate) superimposed on the SARS-CoV M$^{pro/H41A}$–substrate complex (cyan protein, slate substrate). Superposition was by a least-squares fit of C$^\alpha$ atoms of protomer *B* in PDB entry 2q6g to protomer *A* in PDB entry 7n89. (*b*) Protomer *B* in M$^{pro/C145A}$–substrate from SARS-CoV-2 (orange protein, green substrate) superimposed on the SARS-CoV M$^{pro/H41A}$–substrate complex. Superposition was by a least-squares fit of C$^\alpha$ atoms of protomer *A* in PDB entry 2q6g to protomer *B* in PDB entry 7n89.

in SARS-CoV-2 M^pro, A46S, does not introduce appreciable differences despite minor differences in the S2 subsite helix.

4. Conclusion

We describe details of the co-crystal structure of SARS-CoV-2 M^pro bearing an active-site C145A mutation in complex with an octapeptide substrate corresponding to the natural nsp4/nsp5 junction sequence which M^pro autoprocesses prior to its C-terminal junction (nsp5/nsp6 site) during its maturation from the polyprotein precursor (Hsu *et al.*, 2005; Chen, Jonas *et al.*, 2010; Muramatsu *et al.*, 2013). Substrate affinity is driven primarily by P1 Gln hydrogen-bonding interactions and P2 Leu induced-fit rearrangement, with assistance from secondary-structure main-chain interactions. Superposition of the substrate complex with the substrate-free M^pro neutron structure reinforces the suggestion that the substrate-free M^pro structure is innately pre-organized to attack the scissile bond upon a substrate-binding event. We also suggest that our structure is an ideal starting point for *in silico* modeling and simulations looking into understanding M^pro–substrate relationships due to specificity and the absence of cryo-artifacts. These results bring timely insights into the nature of catalysis to expedite structure/mechanism-based drug design.

References

Acharya, A., Agarwal, R., Baker, M. B., Baudry, J., Bhowmik, D., Boehm, S., Byler, K. G., Chen, S. Y., Coates, L., Cooper, C. J., Demerdash, O., Daidone, I., Eblen, J. D., Ellingson, S., Forli, S., Glaser, J., Gumbart, J. C., Gunnels, J., Hernandez, O., Irle, S., Kneller, D. W., Kovalevsky, A., Larkin, J., Lawrence, T. J., LeGrand, S., Liu, S.-H., Mitchell, J. C., Park, G., Parks, J. M., Pavlova, A., Petridis, L., Poole, D., Pouchard, L., Ramanathan, A., Rogers, D. M., Santos-Martins, D., Scheinberg, A., Sedova, A., Shen, Y., Smith, J. C., Smith, M. D., Soto, C., Tsaris, A., Thavappiragasam, M., Tillack, A. F., Vermaas, J. V., Vuong, V. Q., Yin, J., Yoo, S., Zahran, M. & Zanetti-Polzi, L. (2020). *J. Chem. Inf. Model.* **60**, 5832–5852.

Anand, K., Palm, G. J., Mesters, J. R., Siddell, S., Ziebuhr, J. & Hilgenfeld, R. (2002). *EMBO J.* **21**, 3213–3224.

Barrila, J., Bacha, U. & Freire, E. (2006). *Biochemistry*, **45**, 14908–14916.

Bidram, E., Esmaeili, Y., Amini, A., Sartorius, R., Tay, F. R., Shariati, L. & Makvandi, P. (2021). *ACS Biomater. Sci. Eng.* **7**, 2150–2176.

Casañal, A., Lohkamp, B. & Emsley, P. (2020). *Protein Sci.* **29**, 1069–1078.

Chen, S., Jonas, F., Shen, C., Hilgenfeld, R. & Higenfeld, R. (2010). *Protein Cell*, **1**, 59–74.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Chou, C.-Y., Chang, H.-C., Hsu, W.-C., Lin, T.-Z., Lin, C.-H. & Chang, G.-G. (2004). *Biochemistry*, **43**, 14958–14970.

Dai, W., Zhang, B., Jiang, X.-M. 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. K., Xu, Y., Yang, H. & Liu, H. (2020). *Science*, **368**, 1331–1335.

Díaz, N. & Suárez, D. (2021). *Chem. Commun.* **57**, 5314–5317.

Doroftei, B., Ciobica, A., Ilie, O.-D., Maftei, R. & Ilea, C. (2021). *Diagnostics*, **11**, 579.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.

Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J. & Lai, L. (2004). *J. Biol. Chem.* **279**, 1637–1642.

Gavor, E., Choong, Y. K., Er, S. Y., Sivaraman, H. & Sivaraman, J. (2020). *Trends Immunol.* **41**, 1006–1022.

Goetz, D. H., Choe, Y., Hansell, E., Chen, Y. T., McDowell, M., Jonsson, C. B., Roush, W. R., McKerrow, J. & Craik, C. S. (2007). *Biochemistry*, **46**, 8744–8752.

Gorbalenya, A. E. & Snijder, E. J. (1996). *Perspect. Drug Discov. Des.* **6**, 64–86.

Hsu, M.-F., Kuo, C.-J., Chang, K.-T., Chang, H.-C., Chou, C.-C., Ko, T.-P., Shr, H.-L., Chang, G.-G., Wang, A. H.-J. & Liang, P.-H. (2005). *J. Biol. Chem.* **280**, 31257–31266.

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. (2020). *Nature*, **582**, 289–293.

Kneller, D. W., Galanie, S., Phillips, G., O'Neill, H. M., Coates, L. & Kovalevsky, A. (2020). *Structure*, **28**, 1313–1320.

Kneller, D. W., Phillips, G., Kovalevsky, A. & Coates, L. (2020). *Acta Cryst.* F**76**, 483–487.

Kneller, D. W., Phillips, G., O'Neill, H. M., Jedrzejczak, R., Stols, L., Langan, P., Joachimiak, A., Coates, L. & Kovalevsky, A. (2020). *Nat. Commun.* **11**, 3202.

Kneller, D. W., Phillips, G., Weiss, K. L., Pant, S., Zhang, Q., O'Neill, H. M., Coates, L. & Kovalevsky, A. (2020). *J. Biol. Chem.* **295**, 17365–17373.

Lai, L., Han, X., Chen, H., Wei, P., Huang, C., Liu, S., Fan, K., Zhou, L., Liu, Z., Pei, J. & Liu, Y. (2006). *Curr. Pharm. Des.* **12**, 4555–4564.

Lee, J., Worrall, L. J., Vuckovic, M., Rosell, F. I., Gentile, F., Ton, A., Caveney, N. A., Ban, F., Cherkasov, A., Paetzel, M. & Strynadka, N. C. J. (2020). *Nat. Commun.* **11**, 5877.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.

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. (2020). *Eur. J. Med. Chem.* **206**, 112711.

Lucast, L. J., Batey, R. T. & Doudna, J. A. (2001). *Biotechniques*, **30**, 544–554.

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.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Meo, S. A., Bukhari, I. A., Akram, J., Meo, A. S. & Klonoff, D. C. (2021). *Eur. Rev. Med. Pharmacol. Sci.* **25**, 1663–1669.

Muramatsu, T., Kim, Y.-T., Nishii, W., Terada, T., Shirouzu, M. & Yokoyama, S. (2013). *FEBS J.* **280**, 2002–2013.

Muramatsu, T., Takemoto, C., Kim, Y.-T., Wang, H., Nishii, W., Terada, T., Shirouzu, M. & Yokoyama, S. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 12997–13002.

Paasche, A., Zipper, A., Schäfer, S., Ziebuhr, J., Schirmeister, T. & Engels, B. (2014). *Biochemistry*, **53**, 5930–5946.

Pathak, N., Chen, Y.-T., Hsu, Y.-C., Hsu, N.-Y., Kuo, C.-J., Tsai, H. P., Kang, J.-J., Huang, C.-H., Chang, S.-Y., Chang, Y.-H., Liang, P.-H. & Yang, J.-M. (2021). *ACS Nano*, **15**, 857–872.

Ramos-Guzmán, C. A., Ruiz-Pernía, J. J. & Tuñón, I. (2020). *ACS Catal.* **10**, 12544–12554.

Suárez, D. & Díaz, N. (2020). *J. Chem. Inf. Model.* **60**, 5815–5831.

Ton, A.-T., Gentile, F., Hsing, M., Ban, F. & Cherkasov, A. (2020). *Mol. Inf.* **39**, 2000028.

Tong, L. (2002). *Chem. Rev.* **102**, 4609–4626.

Tsai, M.-Y., Chang, W.-H., Liang, J.-Y., Lin, L.-L., Chang, G.-G. & Chang, H.-P. (2010). *J. Biochem.* **148**, 349–358.

Ullrich, S. & Nitsche, C. (2020). *Bioorg. Med. Chem. Lett.* **30**, 127377.

Wang, H., He, S., Deng, W., Zhang, Y., Li, G., Sun, J., Zhao, W., Guo, Y., Yin, Z., Li, D. & Shang, L. (2020). *ACS Catal.* **10**, 5871–5890.

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. (2011). *Acta Cryst.* D**67**, 235–242.

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. (2020). *Nature*, **579**, 265–269.

Xu, J., Zhao, S., Teng, T., Abdalla, A. E., Zhu, W., Xie, L., Wang, Y. & Guo, X. (2020). *Viruses*, **12**, 244.

Xue, X., Yu, H., Yang, H., Xue, F., Wu, Z., Shen, W., Li, J., Zhou, Z., Ding, Y., Zhao, Q., Zhang, X. C., Liao, M., Bartlam, M. & Rao, Z. (2008). *J. Virol.* **82**, 2515–2527.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K. & Hilgenfeld, R. (2020). *Science*, **368**, 409–412.