



# Room-temperature X-ray crystallography reveals the oxidation and reactivity of cysteine residues in SARS-CoV-2 3CL M<sup>Pro</sup>: insights into enzyme mechanism and drug design

Daniel W. Kneller,<sup>a</sup> Gwyndalyn Phillips,<sup>a</sup> Hugh M. O'Neill,<sup>a</sup> Kemin Tan,<sup>b,c</sup> Andrzej Joachimiak,<sup>b,c,d</sup> Leighton Coates<sup>a\*</sup> and Andrey Kovalevsky<sup>a\*</sup>

Received 3 August 2020

Accepted 16 September 2020

Edited by K. Moffat, University of Chicago, USA

**Keywords:** room-temperature X-ray crystallography; SARS-CoV-2; 3CL main protease; 3CL M<sup>Pro</sup>; cysteine oxidation; protonation state; enzyme mechanism; drug design.

**PDB references:** SARS-CoV-2 3CL M<sup>Pro</sup>, peroxysulfenic Cys145 structure (structure I), 6xb0; peroxysulfenic Cys145/*N*-ethylmaleimide Michael addition Cys156 structure (structure II), 6xb1; Cys145/Cys156 *N*-ethylmaleimide Michael addition structure (structure III), 6xb2; native structure at pH 6.0 (structure IV), 6xhu

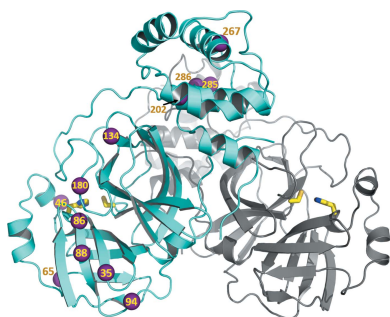
**Supporting information:** this article has supporting information at [www.iucrj.org](http://www.iucrj.org)

<sup>a</sup>Neutron Scattering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA, <sup>b</sup>Center for Structural Genomics of Infectious Diseases, Consortium for Advanced Science and Engineering, University of Chicago, Chicago, IL 60667, USA, <sup>c</sup>Structural Biology Center, X-ray Science Division, Argonne National Laboratory, Argonne, IL 60439, USA, and <sup>d</sup>Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60367, USA. \*Correspondence e-mail: [coatesl@ornl.gov](mailto:coatesl@ornl.gov), [kovalevskyay@ornl.gov](mailto:kovalevskyay@ornl.gov)

The emergence of the novel coronavirus SARS-CoV-2 has resulted in a worldwide pandemic not seen in generations. Creating treatments and vaccines to battle COVID-19, the disease caused by the virus, is of paramount importance in order to stop its spread and save lives. The viral main protease, 3CL M<sup>Pro</sup>, is indispensable for the replication of SARS-CoV-2 and is therefore an important target for the design of specific protease inhibitors. Detailed knowledge of the structure and function of 3CL M<sup>Pro</sup> is crucial to guide structure-aided and computational drug-design efforts. Here, the oxidation and reactivity of the cysteine residues of the protease are reported using room-temperature X-ray crystallography, revealing that the catalytic Cys145 can be trapped in the peroxysulfenic acid oxidation state at physiological pH, while the other surface cysteines remain reduced. Only Cys145 and Cys156 react with the alkylating agent *N*-ethylmaleimide. It is suggested that the zwitterionic Cys145–His45 catalytic dyad is the reactive species that initiates catalysis, rather than Cys145-to-His41 proton transfer via the general acid–base mechanism upon substrate binding. The structures also provide insight into the design of improved 3CL M<sup>Pro</sup> inhibitors.

## 1. Introduction

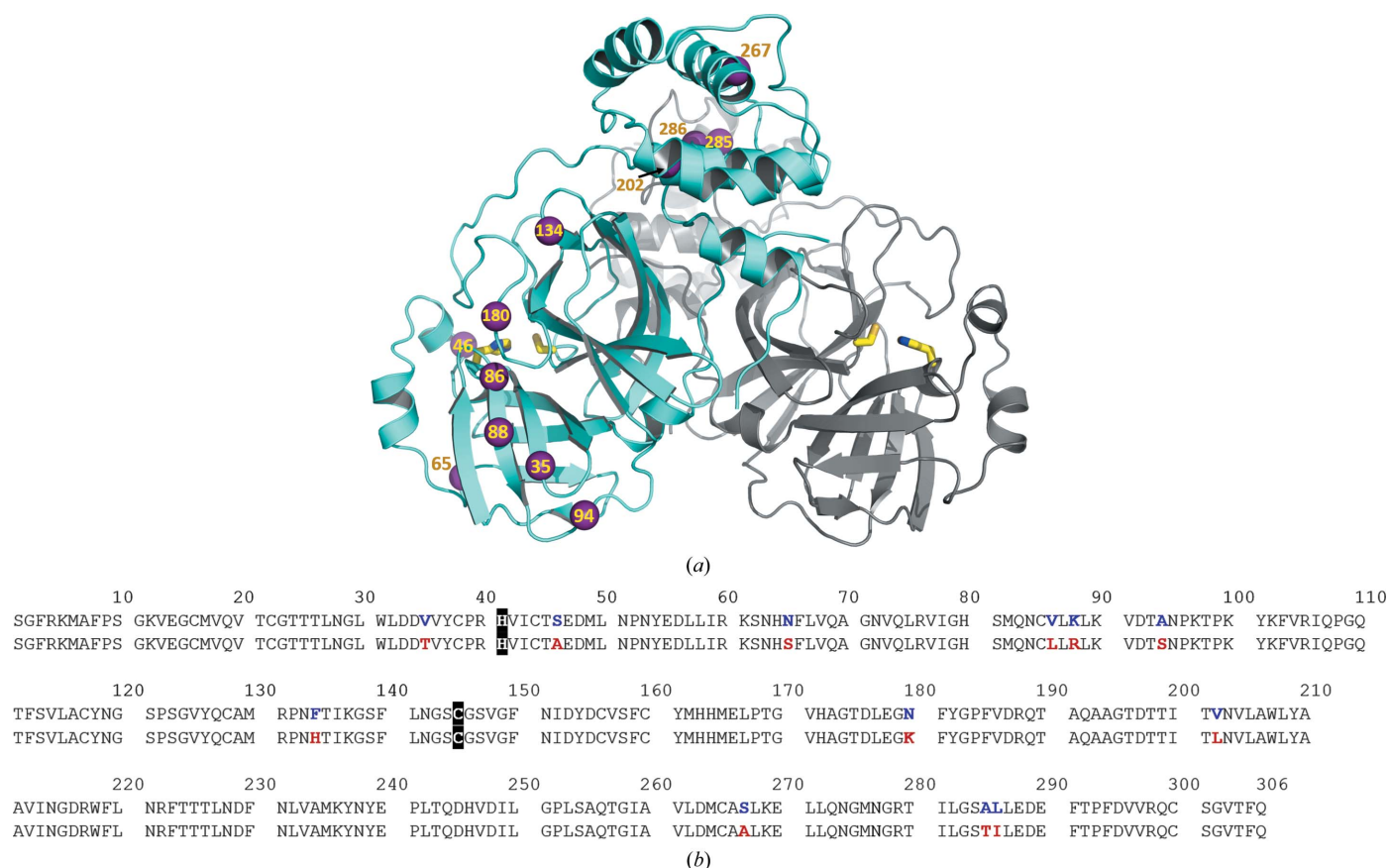
The coronavirus-induced disease COVID-19 has swiftly spread to every corner of the globe, causing a severe pandemic not seen in a century. With over 30 million confirmed cases worldwide to date and no approved treatments or vaccines, COVID-19, which is caused by the novel coronavirus SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2; Helmy *et al.*, 2020; Wu *et al.*, 2020; Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020; Liu *et al.*, 2020), has resulted in almost every country taking unprecedented social measures to limit its spread. SARS-CoV-2, a single-stranded enveloped RNA virus, is related to the common cold viruses, and its ~30 kb genome is 82% identical to that of the SARS coronavirus (SARS-CoV) that was responsible for an outbreak in 2003 (Xu *et al.*, 2020). However, SARS-CoV-2 appears to be more contagious than similar viruses. The replication of such viruses is fully dependent on the function of a 3-chymotrypsin-like protease enzyme (3CL M<sup>Pro</sup>) that hydrolyzes large polyproteins to generate essential nonstructural proteins, which play a fundamental role in transcription/replication during infection (Hilgenfeld, 2014; Gorbalenya & Snijder, 1996; Muramatsu *et*



*al.*, 2016). The SARS-CoV-2 3CL M<sup>PRO</sup> enzyme operates at no fewer than 11 cleavage sites on two large polyproteins 1a and 1ab (replicase 1a, ~450 kDa, and replicase 1ab, ~790 kDa). As 3CL M<sup>PRO</sup> enzymes are indispensable for viral proliferation and owing to the lack of homologous human proteins, these enzymes are promising targets for the design and development of specific protease inhibitors (Pillaiyar *et al.*, 2016; Wang *et al.*, 2017; Dai *et al.*, 2020; Jo *et al.*, 2020; Zhang *et al.*, 2020; Jin, Du *et al.*, 2020) and the repurposing of existing clinical drugs (Jin, Zhao *et al.*, 2020; Ma *et al.*, 2020; Riva *et al.*, 2020). The catalytically active SARS-CoV-2 3CL M<sup>PRO</sup> enzyme is a dimer. Cleavage of the large polyprotein chains by 3CL M<sup>PRO</sup> occurs at the glutamine residue in the P1 position of the substrate via a Cys145–His41 dyad, in which the cysteine thiol functions as the nucleophile in the proteolytic process (Jin, Zhao *et al.*, 2020).

SARS-CoV-2 3CL M<sup>PRO</sup> has 12 amino-acid substitutions compared with the SARS-CoV homolog that occur throughout the sequence in the catalytic domains I and II, and in the helical domain III (Fig. 1; Xu *et al.*, 2020; Jin, Du *et al.*, 2020; Anand *et al.*, 2003; Yang *et al.*, 2005). These substitutions do not appear to significantly affect the function of SARS-CoV-2 3CL M<sup>PRO</sup>, as its catalytic efficiency is very similar to that of the SARS-CoV enzyme (Zhang *et al.*, 2020; Jin, Du *et al.*,

*al.*, 2020; Xue *et al.*, 2007; Huang *et al.*, 2004; Solowiej *et al.*, 2008). Therefore, in mechanistic terms, it can be expected that both enzymes catalyze the peptide-hydrolysis reaction in a similar fashion. The exact enzyme mechanism, however, is still unknown and is poorly characterized. Specifically, it has been debated whether substrate hydrolysis is initiated by proton transfer from the thiol of Cys145 to the imidazole of His41 within the catalytic Cys–His dyad through a general acid–base mechanism (Huang *et al.*, 2004) to give a deprotonated negatively charged thiolate on Cys145 and a protonated positively charged imidazolium on His41, or whether this ion pair (or zwitterion) already exists before substrate binding (Solowiej *et al.*, 2008; Paasche *et al.*, 2014). In either case, the formed thiolate (S<sup>−</sup>) would be a much more reactive species (*i.e.* a better nucleophile) towards the C atom of the scissile peptide bond than a protonated thiol (SH) would be. Here, we used room-temperature X-ray crystallography to present experimental evidence that Cys145 in SARS-CoV-2 3CL M<sup>PRO</sup> is highly reactive and can easily be oxidized to an uncommon peroxy-sulfenic thiol adduct at neutral pH. In contrast, the other 11 Cys residues in the protein sequence are much less prone to oxidation, even though several of them, namely Cys85, Cys156, Cys160 and Cys300, are located on the protein surface and are exposed to the bulk solvent. Remarkably, in



**Figure 1**  
 (a) Structure of the 3CL M<sup>PRO</sup> dimer from SARS-CoV-2 (PDB entry 6wqf; Kneller *et al.*, 2020) in cartoon representation. The locations of the 12 amino-acid substitutions relative to the SARS-CoV enzyme are shown as violet spheres in one monomer. The catalytic Cys–His dyad is shown as sticks in each monomer. (b) Amino-acid sequence alignment of 3CL M<sup>PRO</sup> from SARS-CoV-2 (first row) and SARS-CoV (second row) indicating the positions of the amino-acid substitutions. The catalytic His41 and Cys145 are highlighted in black.

our room-temperature X-ray structure of SARS-CoV-2 3CL M<sup>Pro</sup> obtained from a crystal grown at the acidic pH of 6.0, Cys145 is not oxidized. In addition, Cys145 and Cys156 readily react with the alkylating reagent *N*-ethylmaleimide in the crystal. Our crystallographic data obtained at the near-physiological temperature of 293 K and pH values of 7.0 and 6.0 provide relevant information on the catalytic mechanism of 3CL M<sup>Pro</sup>, the function of the enzyme and the design of specific protease inhibitors to battle COVID-19.

## 2. Materials and methods

### 2.1. General information

Protein-purification supplies were purchased from GE Healthcare, Piscataway, New Jersey, USA. Crystallization reagents were purchased from Hampton Research, Aliso Viejo, California, USA.

### 2.2. Cloning, expression and purification of 3CL M<sup>Pro</sup>

The 3CL M<sup>Pro</sup> (Nsp5 M<sup>Pro</sup>) from SARS CoV-2 was cloned into the pMCSG81-Delta238 plasmid, designated pCSGID-Mpro, expressed and purified according to the published procedure (Kneller *et al.*, 2020). To generate the authentic protease, the N-terminus is flanked by the maltose-binding protein followed by the 3CL M<sup>Pro</sup> autocleavage site SAVLQ↓SGFRK (the arrow indicates the cleavage site) corresponding to cleavage between NSP4 and NSP5 in the viral polyprotein. At the C-terminus, the construct codes for the Human rhinovirus 3C PreScission protease cleavage site (SGVTFQ↓GP) connected to a His<sub>6</sub> tag. The N-terminal flanking sequence is autoprocessed by 3CL M<sup>Pro</sup> during expression, whereas the C-terminal flanking sequence is removed by treatment with PreScission protease. For crystallization, the authentic 3CL M<sup>Pro</sup> was concentrated to 5 mg ml<sup>-1</sup>.

### 2.3. Crystallization

The protein was first crystallized in 0.1 M bis-Tris pH 6.5, 22% PEG 3350. Several crystal ‘flower’ aggregates were collected from this condition and were used to generate microseeds using a Hampton Research Seed Bead kit. For room-temperature X-ray crystallography, crystals were grown in 10 µl sitting drops made by mixing the protein sample (5 mg ml<sup>-1</sup>) and reservoir solution (0.1 M bis-Tris pH 6.5, 18–20% PEG 3350) in a 1:1 ratio with 0.2 µl of microseeds (1:200 dilution). This condition was formulated to produce a pH of 7.0 in the crystallization drop, which was confirmed by direct pH measurement using a microelectrode. Single plate-like crystals grew in several days and continued to grow at 14°C. To obtain fresh crystals with a native (non-oxidized) Cys145, a batch setup was used to grow crystals in 200 µl drops containing protein at 6.6 mg ml<sup>-1</sup> in 10% PEG 3350, 0.1 M bis-Tris pH 7.0 with 0.4 µl of microseeds.

### 2.4. Generation of structures I–IV

Structure **I** containing a peroxysulfenate-modified Cys145 was obtained by soaking a native 3CL M<sup>Pro</sup> crystal transferred from a crystallization drop containing 0.5 mM TCEP in the corresponding crystallization well solution that contained no reducing agent for one hour at 14°C. Structure **II** containing a peroxysulfenate-modified Cys145 and an *N*-ethylmaleimide-modified Cys156 was obtained by first manipulating another crystal in the same fashion, producing structure **I**, but was then soaked in the dark for 72 h at 4°C in a solution containing 2 mM *N*-ethylmaleimide originally dissolved in 100% DMSO. Structure **III** containing *N*-ethylmaleimide-modified Cys145 and Cys156 was obtained by soaking a fresh native 3CL M<sup>Pro</sup> crystal in a solution containing 2 mM *N*-ethylmaleimide for one hour in the dark at room temperature. Structure **IV** in the monoclinic space group *P*2<sub>1</sub> containing the native Cys145 was generated by growing crystals in 0.1 M bis-Tris pH 6.0, 18–20% PEG 3350 using a protein sample dialyzed into the same buffer to guarantee a pH of 6.0 in the crystallization drops.

### 2.5. X-ray data collection and structure refinement

To collect room-temperature diffraction data sets, crystals of 3CL M<sup>Pro</sup> were mounted using the MiTeGen (Ithaca, New York, USA) room-temperature capillary setup. Room-temperature X-ray crystallographic data for the crystals producing structures **I**, **II**, **III** and **IV** were collected immediately following soaking using a Rigaku HighFlux HomeLab instrument equipped with a MicroMax-007 HF X-ray generator and Osmic VariMax optics. The diffraction images were obtained using an EIGER R 4M hybrid photon-counting detector. The diffraction data were integrated using the *CrysAlis<sup>Pro</sup>* software suite (Rigaku, The Woodlands, Texas, USA). The diffraction data were then reduced and scaled using *AIMLESS* (Evans & Murshudov, 2013) from the *CCP4* suite (Winn *et al.*, 2011). Molecular replacement using PDB entry 6wqf (Kneller *et al.*, 2020) was then performed with *MOLREP* (Vagin & Teplyakov, 2010) from the *CCP4* suite. Refinement of each protein structure was conducted using *phenix.refine* from the *Phenix* suite of programs (Liebschner *et al.*, 2019) and the *Coot* molecular-graphics program (Emsley *et al.*, 2010). The geometry of each final structure was then carefully checked with *MolProbity* (Chen *et al.*, 2010). The data-collection and refinement statistics are given in Supplementary Table S1.

## 3. Results

### 3.1. The Cys145 thiol is readily oxidized at pH 7

In our first room-temperature X-ray structure of ligand-free SARS-CoV-2 3CL M<sup>Pro</sup>, structure **I**, obtained at 1.80 Å resolution, we observed that Cys145 is in the rarely seen peroxysulfenic (–S–O–OH) thiol oxidation state that is presumably formed by direct reaction of the cysteine S atom with the molecular oxygen dissolved in water [Figs. 2(a) and 2(b)]. We obtained this structure using a crystal that grew within three days and was then soaked overnight in the

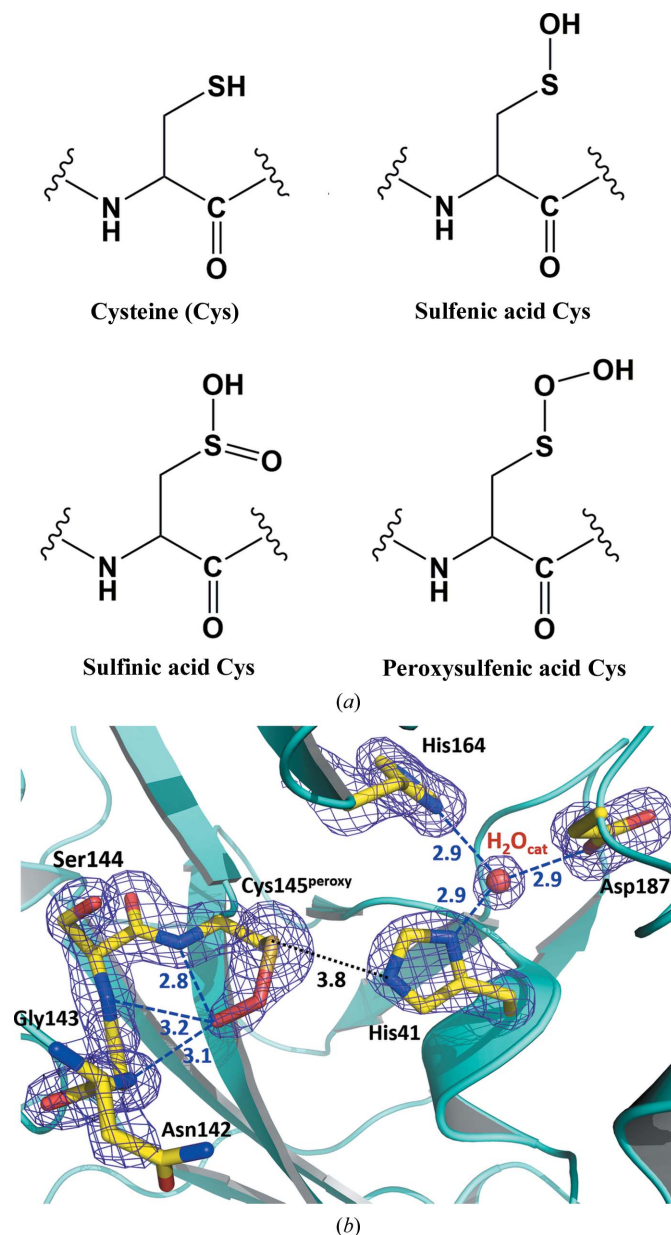
crystallization well solution that contained no reducing agents at pH 7 (Supplementary Figs. S1 and S2). Such fresh crystals normally did not exhibit an oxidized Cys145, as seen in our previously reported room-temperature structure of the ligand-free enzyme (PDB entry 6wqf; Kneller *et al.*, 2020). Importantly, we could also reproduce structure **I** by letting the enzyme crystals grow for more than one week at 14°C, thus indicating that the 0.5 mM TCEP reducing agent present in the crystallization drops does not prevent Cys145 oxidation over a prolonged duration.

It is remarkable that peroxy-sulfenic Cys145 (Cys145<sup>peroxy</sup>) is clearly detected in the SARS-CoV-2 3CL M<sup>Pro</sup> crystal instead of the sulfenic (–S–OH) or sulfinic [–S=O(OH)] acid cysteine modifications [Fig. 2(a)], which are far more prevalent thiol oxidation states (Rudyk & Eaton, 2014; Poole *et al.*, 2020); Peroxy-sulfenic cysteine has previously been observed in structures of cysteine dioxygenase containing a mononuclear nonheme iron, where this thiol oxidation state is presumably stabilized by a direct interaction of the Cys<sup>peroxy</sup> distal-to-sulfur O atom with the iron cation (Simmons *et al.*, 2008; Driggers *et al.*, 2013). In structure **I** the O–O moiety of Cys145<sup>peroxy</sup> is rotated into the oxyanion hole formed by the Ser139–Cys145 turn and forms hydrogen bonds to the main-chain amide N atoms of Gly143, Ser144 and Cys145<sup>peroxy</sup> [Fig. 2(b) and Supplementary Fig. S2]. During catalysis, the oxyanion hole (Taranto *et al.*, 2008; Wu *et al.*, 2013) stabilizes the negative charge on the scissile peptide-bond carbonyl O atom when the covalent tetrahedral intermediate forms (Ménard & Storer, 1992; Tong, 2002; Rauwerdink & Kazlauskas, 2015). It can also stabilize the negatively charged O atom of some aldehyde inhibitors when the reactive aldehyde group forms a covalent C–S bond with Cys145, as was found in some structures of SARS-CoV 3CL M<sup>Pro</sup> (Yang *et al.*, 2003; Zhu *et al.*, 2011). However, in a recent structure of SARS-CoV-2 3CL M<sup>Pro</sup> in complex with an aldehyde inhibitor the aldehyde O atom faces in the opposite direction from the oxyanion hole to make a hydrogen bond to His41 (Dai *et al.*, 2020). Because hydrolase enzymes have an oxyanion hole that is specifically designed to interact with and stabilize the negatively charged O atom (Ménard & Storer, 1992; Tong, 2002; Rauwerdink & Kazlauskas, 2015), we predict that in the Cys145<sup>peroxy</sup> moiety observed in our room-temperature structure the distal-to-sulfur O atom is likely to be deprotonated and negatively charged (–S–O–O<sup>−</sup>).

### 3.2. Cys156 reacts with *N*-ethylmaleimide when Cys145 is oxidized at pH 7

As Cys145 is the catalytic residue, the high reactivity of Cys145 towards molecular oxygen is to be expected. Interestingly, the conditions mentioned above, which resulted in the oxidation of Cys145, did not cause the oxidation of the other surface cysteine residues (Cys85, Cys156, Cys160 and Cys300) that are exposed to bulk solvent. The rest of the cysteine residues in 3CL M<sup>Pro</sup> reside in hydrophobic pockets that are shielded from the bulk solvent and are not expected to be reactive. To probe the reactivity of the surface cysteines, we

manipulated another crystal in the same fashion as the one that gave structure **I** to make sure that Cys145 is oxidized and then soaked the crystal in a solution containing *N*-ethylmaleimide (see Section 2). *N*-Ethylmaleimide is a known Michael reagent for the determination of reactive thiols in proteins (Hill *et al.*, 2009). We obtained a 1.80 Å resolution room-temperature X-ray structure from this crystal (Supplementary Fig. S1), which we call structure **II**. In this structure, we observed a product of *N*-ethylmaleimide Michael addition to Cys156 [Figs. 3(a) and 3(b) and Supplementary Fig. S2] with a refined occupancy of 80%, but no other exposed cysteines



**Figure 2**  
(a) Chemical diagrams of possible oxidation states of a cysteine side chain. (b) The catalytic site of SARS-CoV-2 3CL M<sup>Pro</sup> structure **I**. Possible hydrogen bonds are shown as blue dashed lines; the distance between Cys145 and His41, which is too long for a hydrogen bond, is shown as a black dotted line. The  $2F_o - F_c$  electron-density map contoured at the  $1.5\sigma$  level is shown as a purple mesh. All distances are given in ångströms.

reacted with this alkylating reagent. Cys145 was again found in the peroxy-sulfonate oxidation state, identical to that in structure **I** [Fig. 3(c)], but its conformation was slightly different. In structure **II**, the distal-to-sulfur O atom in Cys145<sup>peroxy</sup> rotates away from the oxyanion hole. It no longer makes reasonable hydrogen bonds to the main-chain amide N atoms of the Gly143, Ser144 and Cys145<sup>peroxy</sup> residues. The closest contact is with NH of Gly143 at a distance of 3.4 Å, suggesting that this O atom might be protonated in structure **II**, and the Cys145<sup>peroxy</sup> moiety is neutral (–S–O–OH).

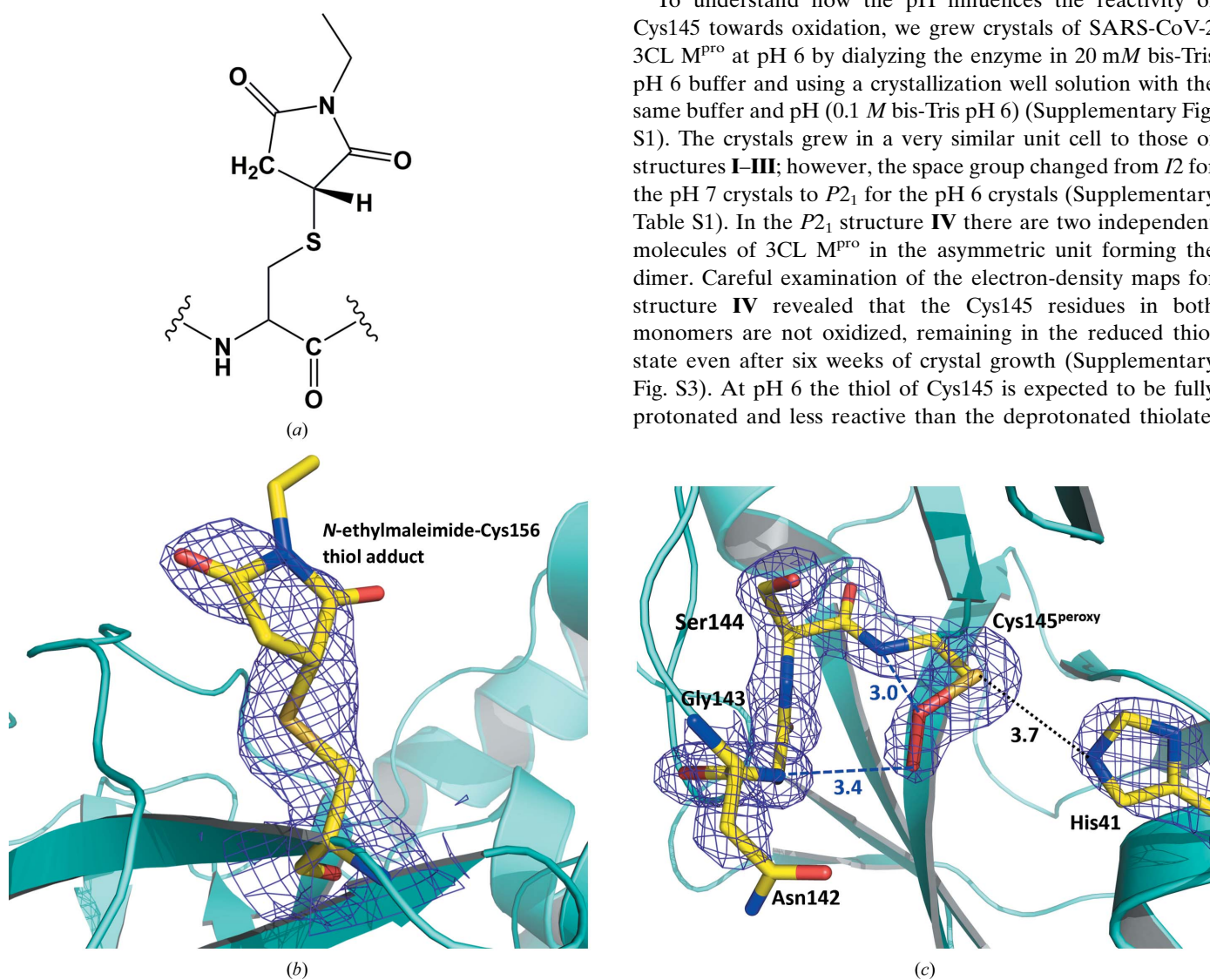
### 3.3. The reduced thiols of Cys145 and Cys156 react with *N*-ethylmaleimide at pH 7

To confirm the reactivity of Cys145 towards *N*-ethylmaleimide and to eliminate the possibility that the lack of reactivity

of other surface cysteine residues towards this alkylating agent as observed in structure **II** was owing to the presence of the oxidized catalytic Cys145, we soaked a fresh native crystal of SARS-CoV-2 3CL M<sup>Pro</sup> in a solution containing *N*-ethylmaleimide (see Section 2 and Supplementary Fig. S1). Removing the exposure of the crystal to a potentially oxidizing environment produced structure **III**, in which we observed that both Cys145 and Cys156 are conjugated with *N*-ethylmaleimide (Fig. 4 and Supplementary Fig. S2), with refined occupancies of 76% and 78%, respectively; yet again, no other cysteines reacted. This result indicates that Cys85, Cys160 and Cys300 do not react with *N*-ethylmaleimide because they cannot be reached by this chemical owing to the crystal packing.

### 3.4. The Cys145 thiol remains reduced at pH 6

To understand how the pH influences the reactivity of Cys145 towards oxidation, we grew crystals of SARS-CoV-2 3CL M<sup>Pro</sup> at pH 6 by dialyzing the enzyme in 20 mM bis-Tris pH 6 buffer and using a crystallization well solution with the same buffer and pH (0.1 M bis-Tris pH 6) (Supplementary Fig. S1). The crystals grew in a very similar unit cell to those of structures **I–III**; however, the space group changed from *I2* for the pH 7 crystals to *P2*<sub>1</sub> for the pH 6 crystals (Supplementary Table S1). In the *P2*<sub>1</sub> structure **IV** there are two independent molecules of 3CL M<sup>Pro</sup> in the asymmetric unit forming the dimer. Careful examination of the electron-density maps for structure **IV** revealed that the Cys145 residues in both monomers are not oxidized, remaining in the reduced thiol state even after six weeks of crystal growth (Supplementary Fig. S3). At pH 6 the thiol of Cys145 is expected to be fully protonated and less reactive than the deprotonated thiolate.



**Figure 3** (a) A chemical diagram of an *N*-ethylmaleimide-conjugated cysteine residue. (b) The  $2F_o - F_c$  electron-density map contoured at the  $0.8\sigma$  level (purple mesh) for *N*-ethylmaleimide-conjugated Cys156 in structure **II**. (c) The catalytic site in structure **II**. Possible hydrogen bonds are shown as blue dashed lines; the distance between Cys145 and His41 is shown as a black dotted line. The  $2F_o - F_c$  electron-density map contoured at the  $1.5\sigma$  level is shown as a purple mesh. All distances are given in ångströms.

Protonation of Cys145 can thus explain why we observe it in the reduced state in structure **IV**.

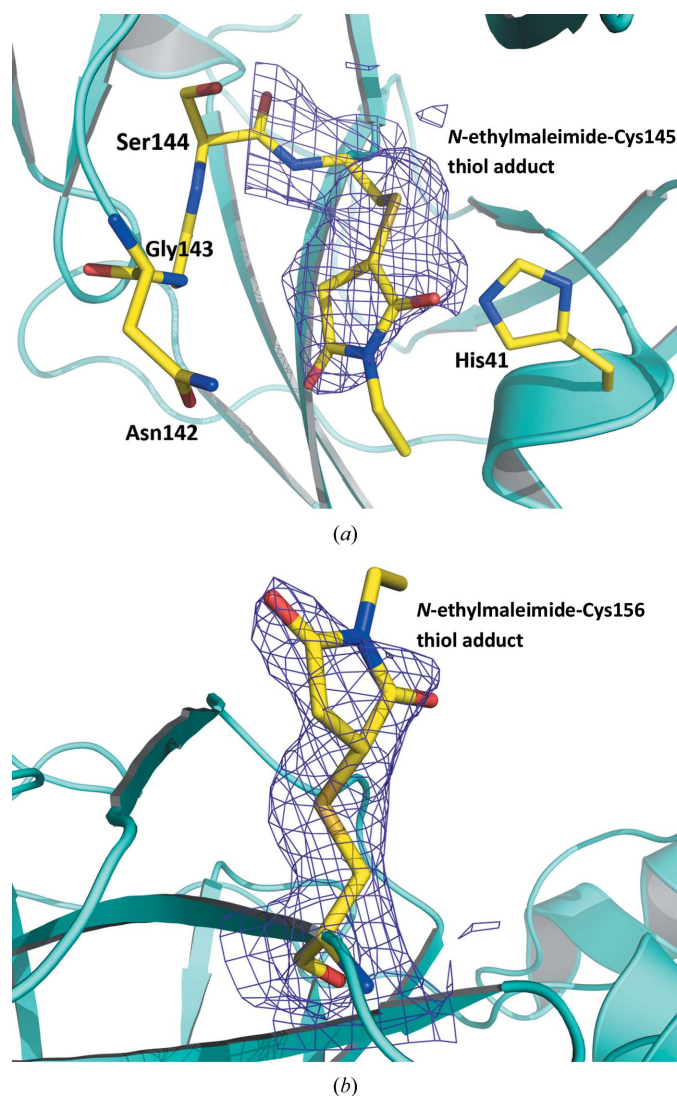
We compared structure **IV** with the previously published pH 6 structure of 3CL M<sup>PRO</sup> from SARS-CoV, which was obtained in the same space group  $P2_1$  (PDB entry 1uj1; Yang *et al.*, 2003; Tan *et al.*, 2005). In PDB entry 1uj1, the S1 substrate-binding subsites have significantly different conformations in the two crystallographically independent monomers. In comparison with structure **IV** (Supplementary Fig. S4), the loop containing residues 138–140 showed dramatic conformational rearrangements owing to the reorientation of Glu166, resulting in the apparent collapse of the S1 subsite and the oxyanion hole. In our structure **IV**, however, we did not detect significant conformational differences within the substrate-binding site between the two monomers or when compared with our room-temperature structure of the native enzyme (Kneller *et al.*, 2020). It is possible that in SARS-CoV-2 3CL M<sup>PRO</sup> such S1 subsite collapse cannot occur, or this might

be a structural artifact owing to the cryocooling of the SARS-CoV 3CL M<sup>PRO</sup> crystals.

#### 4. Discussion

It is important to consider what information can be deduced from our observations of the reactivity of Cys145 in SARS-CoV-2 3CL M<sup>PRO</sup> at different pH values and what insights can be gained regarding the catalytic mechanism of the enzyme. The experimental  $pK_a$  values for Cys145 and His41 were previously determined to be  $8.0 \pm 0.3$  and  $6.3 \pm 0.1$ , respectively, for SARS-CoV 3CL M<sup>PRO</sup>, which is 96% homologous to the SARS-CoV-2 enzyme (Huang *et al.*, 2004; Solowiej *et al.*, 2008). It can be anticipated that the  $pK_a$  values for the Cys–His dyad in the SARS-CoV-2 enzyme are similar. Therefore, at a pH of 7.0 most of the catalytic sites should contain a protonated Cys145 and a neutral His41; thus, Cys145 should not be particularly reactive. This is consistent with a previous suggestion that the ligand-free enzyme adopts the catalytically resting (nonreactive) state and is activated by proton transfer from the Cys145 thiol to the His41 imidazole upon substrate binding, generating a zwitterionic pair with a highly reactive Cys145 thiolate (Paasche *et al.*, 2014). However, theoretical quantum mechanics/molecular mechanics calculations using approximated coupled-cluster approaches have shown that this proton transfer has a high energy barrier of  $\sim 24 \text{ kcal mol}^{-1}$  (Paasche *et al.*, 2013), which is too high considering the measured  $k_{\text{cat}}$  of  $\sim 1\text{--}8 \text{ s}^{-1}$  (Solowiej *et al.*, 2008; Kuo *et al.*, 2004). Conversely, at the pH of 7.0 in our crystallization conditions, around 10–15% of the Cys145–His41 dyad side chains would be in the opposite protonation states, *i.e.* with a deprotonated negatively charged Cys145 and a protonated positively charged His41. It follows that these catalytic sites would be much more reactive and can be prone to oxidation; then, once all of them have reacted with oxygen dissolved in the solution, the rest of the molecules would re-equilibrate, generating more reactive deprotonated Cys145 species. At the physiological pH of 7.4, there would be close to 50% of deprotonated Cys145 residues.

Therefore, we suggest that the reactive species of the 3CL M<sup>PRO</sup> enzyme are those that already contain the Cys–His zwitterionic pair before a substrate (or covalent inhibitor) binds. Notably, our suggestion is in agreement with the normal solvent kinetic isotope effect (SKIE) on the burst rate in pre-steady-state studies and the inverse SKIE in steady-state measurements previously performed for SARS-CoV 3CL M<sup>PRO</sup>. In contrast, these results are not consistent with rate-limiting general acid–base proton transfer from Cys145 to His41 (Solowiej *et al.*, 2008). When the pH of the crystallization solution was lowered to 6.0, we observed no oxidation of Cys145, indicating that the protonated thiol is not reactive towards oxygen, as expected. We also did not detect structural changes in the substrate-binding subsite S1 in our structure at pH 6.0, which was previously suggested to be the pH-triggered activation switch and to which the diminished activity of the SARS-CoV enzyme at acidic pH was attributed (Yang *et al.*, 2003; Tan *et al.*, 2005). Our study instead suggests that if the



**Figure 4**  
The  $2F_o - F_c$  electron-density maps contoured at the  $0.8\sigma$  level (purple mesh) for *N*-ethylmaleimide-conjugated Cys145 (*a*) and Cys156 (*b*) observed in structure **III**.

pH-dependence of SARS-CoV-2 3CL M<sup>Pro</sup> activity follows the same trend as that of the SARS-CoV enzyme, as should be expected, the enzyme activity must be governed by the protonation and deprotonation of the Cys–His dyad rather than by conformational changes of the active-site cavity.

If our hypothesis is correct, structure-based drug-design efforts should consider the reactive species of SARS-CoV-2 3CL M<sup>Pro</sup> in the design of covalent and noncovalent inhibitors. Such inhibitors may need to specifically target the enzyme active site with a negatively charged Cys145 and a positively charged His41 that create a different electrostatic environment compared with the neutral Cys–His dyad. Another important insight provided by our observations relates to whether the reactivity of Cys145 may be directly exploited for the design of novel inhibitors targeting SARS-CoV-2 3CL M<sup>Pro</sup>. One possibility is to consider direct Cys145 oxidation by inhibitors containing an ozonide (or peroxide) reactive group (Dussault *et al.*, 1999), called a warhead, instead of the commonly utilized aldehyde, ketone or Michael acceptor warheads (Dai *et al.*, 2020; Zhang *et al.*, 2020; Jin, Du *et al.*, 2020; Pillaiyar *et al.*, 2016). Such inhibitors have been proposed as potential therapeutics against human cytomegalovirus (Wang *et al.*, 2018). For example, a 1,2,4-trioxolane warhead (Dussault *et al.*, 1999) can directly oxidize the Cys145 thiol to a sulfenic acid, while generating an aldehyde ‘byproduct’, thus ultimately inactivating two 3CL M<sup>Pro</sup> molecules with one inhibitor molecule.

## 5. Conclusions

We succeeded in trapping the SARS-CoV-2 3CL M<sup>Pro</sup> enzyme with Cys145 oxidized to the peroxysulfenic acid in the crystal and performed X-ray crystallographic experiments at a physiologically relevant temperature. We observed that none of the other solvent-exposed surface cysteine residues were oxidized. The catalytic Cys145 and surface Cys156 readily react with the alkylating reagent *N*-ethylmaleimide. Our structures indicate that Cys145 is highly reactive at physiological pH and suggest that the zwitterionic Cys145<sup>−</sup>–His41<sup>+</sup> catalytic dyad is responsible for the initiation of catalysis, rather than proton transfer from Cys145 to His41 via a general acid–base mechanism upon substrate binding. At the acidic pH of 6.0, Cys145, as well as all other cysteines, appears to be nonreactive towards oxidation owing to the full protonation of its thiol. Our structures also point to possible improvements in the design of specific SARS-CoV-2 3CL M<sup>Pro</sup> inhibitors.

## Acknowledgements

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 Oak Ridge National Laboratory (ORNL). 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. The authors declare no conflicts of interest.

## Funding information

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. Funding for this project was provided in part by federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under Contract No. HHSN272201700060C and by the US Department of Energy (DOE) Office of Science, operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## References

- Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. (2003). *Science*, **300**, 1763–1767.
- 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.
- Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020). *Nat. Microbiol.* **5**, 536–544.
- 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. (2020). *Science*, **368**, 1331–1335.
- Driggers, C. M., Cooley, R. B., Sankaran, B., Hirschberger, L. L., Stipanuk, M. H. & Karplus, P. A. (2013). *J. Mol. Biol.* **425**, 3121–3136.
- Dussault, P. H., George, A. D. & Trullinger, T. K. (1999). *Bioorg. Med. Chem. Lett.* **9**, 3255–3258.
- 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.
- Gorbalenya, A. E. & Snijder, E. J. (1996). *Perspect. Drug Discov. Des.* **6**, 64–86.
- Helmy, Y., Fawzy, M., Elasad, A., Sobieh, A., Kenney, S. P. & Shehata, A. A. (2020). *J. Clin. Med.* **9**, 1225.
- Hilgenfeld, R. (2014). *FEBS J.* **281**, 4085–4096.
- Hill, B. G., Reily, C., Oh, J.-Y., Johnson, M. S. & Landar, A. (2009). *Free Radic. Biol. Med.* **47**, 675–683.
- Huang, C., Wei, P., Fan, K., Liu, Y. & Lai, L. (2004). *Biochemistry*, **43**, 4568–4574.
- 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.
- 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. (2020). *Nat. Struct. Mol. Biol.* **27**, 529–532.
- Jo, S., Kim, S., Shin, D. H. & Kim, M.-S. (2020). *J. Enzyme Inhib. Med. Chem.* **35**, 145–151.
- 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.
- Kuo, C.-J., Chi, Y.-H., Hsu, J. T.-A. & Liang, P.-H. (2004). *Biochem. Biophys. Res. Commun.* **318**, 862–867.
- 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, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., Smoot, J., Gregg, A. C., Daniels, A. D., Jervey, S. & Albaiu, D. (2020). *ACS Cent. Sci.* **6**, 315–331.
- 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. (2020). *Cell Res.* **30**, 678–692.
- Ménard, R. & Storer, A. C. (1992). *Biol. Chem. Hoppe Seyler*, **373**, 393–400.
- 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., Schirmeister, T. & Engels, B. (2013). *J. Chem. Theory Comput.* **9**, 1765–1777.
- Paasche, A., Zipper, A., Schäfer, S., Ziebuhr, J., Schirmeister, T. & Engels, B. (2014). *Biochemistry*, **53**, 5930–5946.
- Pillaiyar, T., Manickam, M., Namasivayam, V., Hayashi, Y. & Jung, S.-H. (2016). *J. Med. Chem.* **59**, 6595–6628.
- Poole, L. B., Furdui, C. M. & King, S. B. (2020). *Essays Biochem.* **64**, 1–17.
- Rauwerdink, A. & Kazlauskas, R. J. (2015). *ACS Catal.* **5**, 6153–6176.
- Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Pache, L., Burgstaller-Muehlbacher, S., De Jesus, P. D., Teriete, P., Hull, M. V., Chang, M. W., Chan, J. F.-W., Cao, J., Poon, V. K.-M., Herbert, K. M., Cheng, K., Nguyen, T.-T. H., Rubanov, A., Pu, Y., Nguyen, C., Choi, A., Rathnasinghe, R., Schotsaert, M., Miorin, L., Dejoze, M., Zwaka, T. P., Sit, K.-Y., Martinez-Sobrido, L., Liu, W.-C., White, K. M., Chapman, M. E., Lendy, E. K., Glynne, R. J., Albrecht, R., Rupp, E., Mesecar, A. D., Johnson, J. R., Benner, C., Sun, R., Schultz, P. G., Su, A. I., García-Sastre, A., Chatterjee, A. K., Yuen, K.-Y. & Chanda, S. K. (2020). *Nature*, <https://doi.org/10.1038/s41586-020-2577-1>.
- Rudyk, O. & Eaton, P. (2014). *Redox Biol.* **2**, 803–813.
- Simmons, C. R., Krishnamoorthy, K., Granett, S. L., Schuller, D. J., Dominy, J. E. Jr, Begley, T. P., Stipanuk, M. H. & Karplus, P. A. (2008). *Biochemistry*, **47**, 11390–11392.
- Solowiej, J., Thomson, J. A., Ryan, K., Luo, C., He, M., Lou, J. & Murray, B. W. (2008). *Biochemistry*, **47**, 2617–2630.
- Tan, J., Verschuere, K. H. G., Anand, K., Shen, J., Yang, M., Xu, Y., Rao, Z., Bigalke, J., Heisen, B., Mesters, J. R., Chen, K., Shen, X., Jiang, H. & Hilgenfeld, R. (2005). *J. Mol. Biol.* **354**, 25–40.
- Taranto, A. G., Carvalho, P. & Avery, M. A. (2008). *J. Mol. Graph. Model.* **27**, 275–285.
- Tong, L. (2002). *Chem. Rev.* **102**, 4609–4626.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Wang, L., Bao, B.-B., Song, G.-Q., Chen, C., Zhang, X.-M., Lu, W., Wang, Z., Cai, Y., Li, S., Fu, S., Song, F.-H., Yang, H. & Wang, J.-G. (2017). *Eur. J. Med. Chem.* **137**, 450–461.
- Wang, Y., Mukhopadhyay, R., Roy, S., Kapoor, A., Su, Y., Charman, S. A., Chen, G., Wu, J., Wang, X., Vennerstrom, J. L. & Arav-Boger, R. (2018). *Antimicrob. Agents Chemother.* **63**, e01735-18.
- 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, C.-G., Cheng, S.-C., Chen, S.-C., Li, J.-Y., Fang, Y.-H., Chen, Y.-H. & Chou, C.-Y. (2013). *Acta Cryst. D* **69**, 747–755.
- 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., Yang, H., Shen, W., Zhao, Q., Li, J., Yang, K., Chen, C., Jin, Y., Bartlam, M. & Rao, Z. (2007). *J. Mol. Biol.* **366**, 965–975.
- Yang, H., Xie, W., Xue, X., Yang, K., Ma, J., Liang, W., Zhao, Q., Zhou, Z., Pei, D., Ziebuhr, J., Hilgenfeld, R., Yuen, K. Y., Wong, L., Gao, G., Chen, S., Chen, Z., Ma, D., Bartlam, M. & Rao, Z. (2005). *PLoS Biol.* **3**, e324.
- Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R. & Rao, Z. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 13190–13195.
- Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K. & Hilgenfeld, R. (2020). *Science*, **368**, 409–412.
- Zhu, L., George, S., Schmidt, M. F., Al-Gharabli, S. I., Rademann, J. & Hilgenfeld, R. (2011). *Antiviral Res.* **92**, 204–212.