doi:10.1016/j.jmb.2005.09.004 J. Mol. Biol. (2005) 353, 1137–1151

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Crystal Structures of the Main Peptidase from the SARS Coronavirus Inhibited by a Substrate-like Aza-peptide Epoxide

Ting-Wai Lee[1], Maia M. Cherney[1], Carly Huitema[2], Jie Liu[2] Karen Ellis James[3], James C. Powers[3], Lindsay D. Eltis[2] and Michael N. G. James[1]*

[1]*Canadian Institute of Health Research Group in Protein Structure and Function Department of Biochemistry University of Alberta Edmonton, Alta., Canada T6G 2H7*

[2]*Department of Microbiology and Immunology, University of British Columbia, Vancouver BC, Canada V6T 1Z3*

[3]*School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332-0400, USA*

**Corresponding author*

The main peptidase (M^{pro}) from the coronavirus (CoV) causing severe acute respiratory syndrome (SARS) is one of the most attractive molecular targets for the development of anti-SARS agents. We report the irreversible inhibition of SARS-CoV M^{pro} by an aza-peptide epoxide (APE; $k_{inact}/K_i = 1900(\pm 400)\ M^{-1}\ s^{-1}$). The crystal structures of the M^{pro}:APE complex in the space groups $C2$ and $P2_12_12_1$ revealed the formation of a covalent bond between the catalytic Cys145 S^{γ} atom of the peptidase and the epoxide C3 atom of the inhibitor, substantiating the mode of action of this class of cysteine-peptidase inhibitors. The aza-peptide component of APE binds in the substrate-binding regions of M^{pro} in a substrate-like manner, with excellent structural and chemical complementarity. In addition, the crystal structure of unbound M^{pro} in the space group $C2$ revealed that the "N-fingers" (N-terminal residues 1 to 7) of both protomers of M^{pro} are well defined and the substrate-binding regions of both protomers are in the catalytically competent conformation at the crystallization pH of 6.5, contrary to the previously determined crystal structures of unbound M^{pro} in the space group $P2_1$.

Keywords: X-ray crystallography; SARS coronavirus; main peptidase; aza-peptide epoxide; structure-based drug design

Introduction

Severe acute respiratory syndrome (SARS) first emerged in China in November 2002. This highly transmissible, infectious and often fatal disease spread to 32 countries across five continents, causing close to 8500 infections and over 900 deaths, until being contained by the summer of 2003. Several infections in Asia were reported subsequently, alerting the world that it remains at risk of another outbreak of SARS (World Health Organization: Severe acute respiratory syndrome†). Although the development of anti-SARS vaccines and drugs are in progress, these agents are still far from clinical use.[1] Additional efforts in these areas of study therefore remain paramount.

SARS is caused by a novel coronavirus (CoV);[2–4] it is an enveloped positive-sense single-stranded RNA virus infecting respiratory and gastrointestinal epithelial cells, macrophages, and other cell types, thereby causing systemic changes and damaging many vital organs such as lung, heart, liver, kidney and adrenal gland.[4,5] Anti-SARS therapeutics could target several major steps in the viral life-cycle, such as virus–cell interactions, virus entry, intracellular viral replication, and virus assembly and exit.[1] The intracellular replication of SARS-CoV is mediated by a "replicase" complex derived from two virally coded polyprotein precursors, pp1a (486 kDa) and pp1ab (790 kDa).[6,7] The formation of this replicase complex requires the extensive processing of the two polyproteins by two cysteine peptidases within them, namely the main peptidase (M^{pro}), also known as the 3C-like

Abbreviations used: SARS, severe acute respiratory syndrome; CoV, coronavirus; APE, aza-peptide epoxide; M^{pro}, the main peptidase from CoV; CMK, chloromethyl ketone.

E-mail address of the corresponding author: michael.james@ualberta.ca

† http://www.who.int/csr/sars/en/

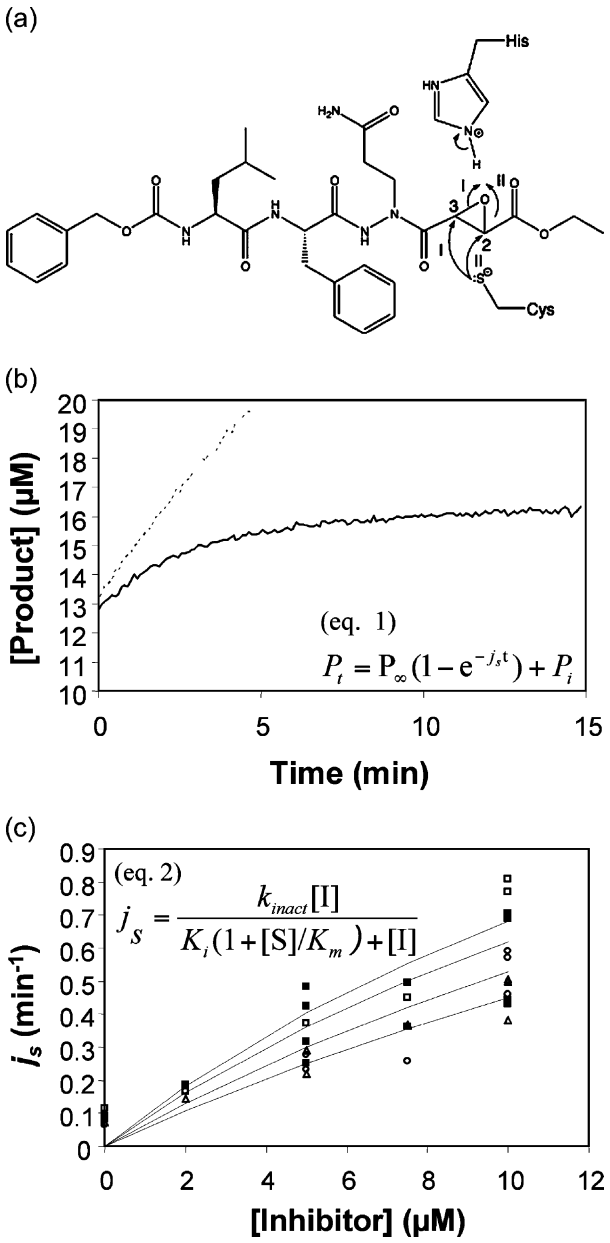

Figure 1. Inhibition of SARS-CoV M^{pro} by aza-peptide epoxides (APEs). (a) APEs synthesized for our study, Cbz-Leu-Phe-AGln-EP-COOEt. The epoxide carbon atoms are numbered and their stereochemistries are omitted for simplicity. The proposed mechanism for the irreversible inhibition of clan CD cysteine peptidases by APEs is indicated by arrows. Cbz, the benzyloxycarbonyl group; AGln, aza-glutamine; EP, epoxide; COOEt, ethyl ester. (b) Progress curves for the steady-state cleavage of a fluorogenic peptide substrate observed using 64 μM peptide and either no APE (dotted line) or 5 μM Cbz-Leu-Phe-AGln-(*S*,*S*)EP-COOEt (continuous line). The appearance of product was followed using excitation and emission wavelengths of 320 and 420 nm, respectively. Analysis of these data using equation (1), in which P_i and P_∞ represent the initial and final product concentrations, respectively, yielded inactivation rates, j_s, of 0.07 min^{-1} and 0.29 min^{-1}, respectively. (c) The rate of M^{pro} inhibition was determined using 16 μM (□), 32 μM (■), 64 μM (Δ) and 100 μM (○) peptide. Equation (2) was fit to the data using the least-squares, dynamic

peptidase (3CL^{pro}) because of its similarity to the 3C peptidases of *Picornaviridae*,[8] and the accessory papain-like peptidase 2 (PL2^{pro}). PL2^{pro} cleaves at three sites in the N-proximal regions of the two polyproteins, whereas M^{pro} cleaves at 11 sites in the central and C-proximal regions of the two polyproteins. M^{pro} releases the key proteins in viral replication, such as the RNA-dependent RNA polymerase and the helicase.[7] Playing such an essential role, SARS-CoV M^{pro} is an attractive molecular target for the development of anti-SARS drugs acting as the inhibitors of the peptidase.

SARS-CoV M^{pro} has a molecular mass of 33.8 kDa per protomer; it exists as a homodimer over a wide range of concentrations in solution.[9–12] The crystal structures of M^{pro} in the space group $P2_1$ showed that the two protomers of the dimeric peptidase are oriented almost perpendicular to each other and that each protomer consists of three domains. Domain I (residues 8 to 101) and domain II (residues 102 to 184) comprise a two-β-barrel fold similar to that of the chymotrypsin-type serine peptidases. Domain III (residues 201 to 300) has five α-helices and is connected to domain II by a long loop (residues 185 to 200). Each protomer has its own substrate-binding region situated in the cleft between domains I and II.[13] A recent mutagenesis study has confirmed that M^{pro} is a cysteine peptidase with a Cys-His catalytic dyad at the active site.[14] As suggested by the structure-based sequence alignment of the main peptidases (including their flanking residues in the polyproteins) from SARS-CoV and other coronaviruses,[15] and confirmed by *in vitro* studies,[7,9] these peptidases preferentially cleave at a consensus sequence for the P4 to P1′ residues of substrates (nomenclature based on that of Schechter and Berger[16]): (amino acid with a small side-chain)-(any amino acid)-Leu-Gln↓(Ala, Ser, Gly).

A number of small-molecule inhibitors of SARS-CoV M^{pro} have been proposed using various methodologies, such as knowledge-based discovery[17,18] and high-throughput screening (experimental[19–21] or virtual[22,23]). Although the efficacies of many of these inhibitors were supported by assay results, the modes of action are unknown for most of them. Also, the lack of structural information for inhibitors binding to M^{pro} impedes the structure-based optimization of these inhibitors. To our knowledge, the crystal structure of M^{pro} bound by chloromethyl ketone (CMK) has so far been the only published structure for an inhibitor-bound M^{pro}.[13]

Aza-peptide epoxides (APEs) were synthesized as a new class of inhibitors apparently specific for clan CD cysteine peptidases[24] (based on the

weighting options of LEONORA,[37] yielding the following parameters: $k_{inact}=35(\pm17)\times10^{-3}$ s^{-1}, $K_i=18(\pm9)$ μM and $K_m=96(\pm31)$ μM. Additional experimental details are provided in Materials and Methods.

classification by Barrett and Rawlings[25]), including the legumains[26] and the caspases.[27] Each APE has an aza-peptide component, with an epoxide moiety attached to the carbonyl group of the P1 residue. The side-chain of the P1 residue predominantly determines the target-peptidase specificity of an APE. The substituent on the epoxide C2 atom also allows some tuning of both the inhibitory activity and specificity of APE towards a particular target peptidase. The aza-peptide component resembles a peptide, except that the C^{α} atom of the P1 residue in the former is replaced by a nitrogen atom to form an aza-amino acid residue. This introduces trigonal planar geometry to the α-atom of the P1 residue and reduces the electrophilicity of the carbonyl C atom of the P1 residue, thereby making the carbonyl group of the P1 residue resistant to nucleophilic attack.[28] It has been proposed that APEs inhibit their target peptidases irreversibly by a mechanism in which the catalytic Cys S^{γ} atom nucleophilically attacks one of the two epoxide carbon atoms (C2 or C3) of APE (Figure 1(a)).[24,26,27] This results in the opening of the conformationally strained epoxide ring, and the formation of a covalent bond between the Cys S^{γ} atom and the attacked APE atom.

Results

Inhibition of SARS-CoV M^{pro} by an APE

We hypothesized that an APE possessing an aza-glutamine (AGln) as the P1 residue to mimic the S1 specificity of SARS-CoV M^{pro} for Gln (Figure 1(a)) would irreversibly inhibit the peptidase. Accordingly, we synthesized Cbz-Leu-Phe-AGln-(*S*,*S*)EP-COOEt and Cbz-Leu-Phe-AGln-(*R*,*R*)EP-COOEt. Micromolar quantities of the *S*,*S* diastereomer strongly inhibited the cleavage of a peptidic substrate, manifesting itself as a pronounced slowing of the reaction velocity as the reaction progressed (Figure 1(b)). Under these conditions, the *R*,*R* diastereomer did not detectably inhibit M^{pro}. Analysis of the rates of inactivation (j_s) at different concentrations of substrate and inhibitor

Table 1. Parameters and statistics derived from X-ray diffraction data processing and structure refinement

	Unbound M^{pro}	M^{pro}:APE	$M^{pro}_{+A(-1)}$: APE
Data processing			
Wavelength (Å)	1.116	1.116	1.116
Resolution limit[a] (Å)	50.00–2.08 (2.15–2.08)	50.00–1.88 (1.95–1.88)	50.00–2.30 (2.38–2.30)
Space group	*C*2	*C*2	$P2_12_12_1$
Unit-cell constants			
a (Å)	107.76	106.71	65.35
b (Å)	82.14	83.68	67.48
c (Å)	53.13	52.87	167.32
α (°)	90.00	90.00	90.00
β (°)	104.87	105.66	90.00
γ (°)	90.00	90.00	90.00
Mosaicity (°)	1.00	0.61	0.23
Total number of reflections	50,960 (3872)	74,092 (6410)	134,161 (10,755)
Number of unique reflections	26,196 (2482)	35,576 (3301)	33,305 (3147)
Redundancy	1.9 (1.6)	2.0 (1.7)	4.0 (3.4)
Completeness (%)	95.8 (91.0)	96.3 (90.0)	99.2 (95.5)
R_{sym}[b] (%)	3.7 (22.5)	3.3 (31.5)	5.2 (51.8)
$I/\sigma(I)$	14.9 (3.4)	21.5 (2.1)	24.8 (2.0)
Structure refinement			
Resolution range (Å)	40.00–2.08	40.00–1.88	40.00–2.30
R_{work}[c] (%)	19.2	19.8	17.5
R_{free}[c] (%)	24.6	24.2	24.9
Number of non-hydrogen atoms per asymmetric unit (average *B* factor, Å^2)			
Protein	2371 (56.03)	2371 (42.72)	4690 (46.61)
APE	Not applicable	46 (45.28)	92 (65.29)
Solvent	150.5 (66.13)	264.5 (55.29)	279.5 (56.28)
rms deviation from ideal geometry			
Bond lengths (Å)	0.029	0.022	0.021
Bond angles (°)	2.801	2.342	1.944
Ramachandran plot			
Favored (%)	87.2	91.3	88.5
Allowed (%)	10.6	6.8	9.9
Generously allowed (%)	1.1	1.5	0.8
Disallowed (%)	1.1	0.4	0.8

[a] Numbers in the parentheses refer to the highest resolution bins.

[b] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_{hkl,i} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_i I_{hkl,i}$, where $I_{hkl,i}$ and $\langle I_{hkl}\rangle$ are the *i*th observed intensity and average intensity of the reflection *hkl*, respectively.

[c] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes of a particular reflection, respectively, and the summation is over 95% of the reflections in the specified resolution range. The remaining 5% of the reflections were randomly selected before the structure refinement and not included in the structure refinement. R_{free} was calculated over these reflections using the same equation as for R_{work}.[49]

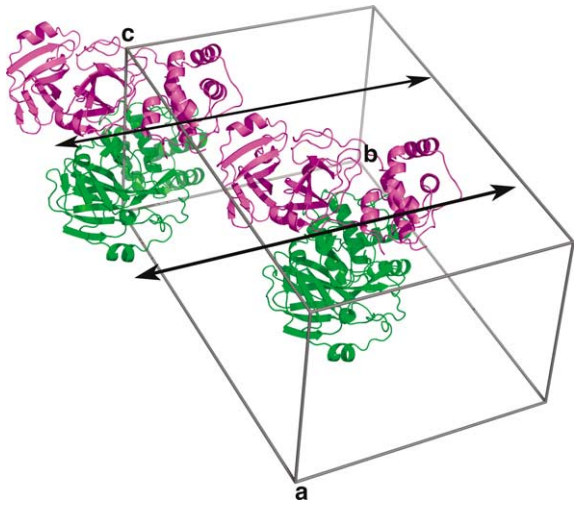

Figure 2. SARS-CoV M^{pro} in a *C*2 unit cell. The two protomers making up the dimeric peptidase are from two different asymmetric units (green and magenta) related by the crystallographic 2-fold symmetry along the *b*-axis (shown as arrows).

indicated that APE inhibited M^{pro} with a $k_{inact}/K_i = 1900(\pm 400)\ M^{-1}\ s^{-1}$.

Structure determination

We have determined the crystal structures of SARS-CoV M^{pro} in three forms: wild-type peptidase in the absence and presence of APE in the space group *C*2, and a variant of the peptidase with an Ala added to the N terminus of the wild-type sequence, $M^{pro}_{+A(-1)}$, bound by APE in the space group $P2_12_12_1$. The parameters and statistics derived from X-ray diffraction data processing and structure refinement are summarized in Table 1. For unbound M^{pro} and the M^{pro}:APE complex in the space group *C*2, each asymmetric unit has only one protomer of the dimer. The two protomers of each dimer are related by the crystallographic 2-fold symmetry (Figure 2). All residues of the protomer (residues 1 to 306) were identified in the electron density maps. In the Ramachandran plot for the structure of unbound M^{pro}, Asp33, Ala46 and Glu47 are in the generously allowed regions, whereas Asn84, Tyr154 and Ile286 are in the disallowed regions. The Asp33 $O^{\delta 2}$ atom forms a hydrogen bond with the Tyr101 phenolic OH group (2.9 Å). The Asn84 $N^{\delta 2}$ atom forms a hydrogen bond with the Glu178 carbonyl O atom (3.2 Å), and possibly there are hydrogen bonds and van der Waals forces between the side-chains of Asn84 and Lys180 as well. Hydrophobic interactions occur between the side-chains of Thr285 and Ile286 of opposite protomers at the dimer interface. The electron densities for the side-chains of Ala46, Glu47 and Tyr154 are not well defined. Similarly, in the Ramachandran plot for the structure of the M^{pro}:APE complex, Asp33, Asn84, Tyr154 and Asn277 are in the generously allowed

Table 2. Root-mean-square differences (rmsd; in Å) for superimpositions of the structures of SARS-CoV M^{pro} determined in different conditions

	Unbound M^{pro} (*C*2, pH 6.5)	M^{pro}:APE (*C*2, pH 6.5)	$M^{pro}_{+A(-1)}$: *APE* ($P2_12_12_1$, pH 6.5)
Unbound M^{pro} (*C*2, pH 6.5)		PP: 0.24	PA: 0.35
	–	DD: 0.37	PB: 0.35
			DD: 0.65
M^{pro}:APE (*C*2, pH 6.5)			PA: 0.39
	–	–	PB: 0.38
			DD: 0.74
Unbound M^{pro} ($P2_1$, pH 6.0)	AP: 0.63	AP: 0.75	AA: 0.62
	BP: 0.32	BP: 0.37	BB: 0.39
	DD: 0.48	DD: 0.48	AB: 0.62
			BA: 0.34
			DD: 0.79
Unbound M^{pro} ($P2_1$, pH 7.6)	AP: 0.92	AP: 0.99	AA: 0.76
	BP: 0.99	BP: 1.02	BB: 0.89
	DD: 1.03	DD: 1.04	AB: 0.82
			BA: 0.90
			DD: 0.83
Unbound M^{pro} ($P2_1$, pH 8.0)	AP: 0.92	AP: 0.98	AA: 0.75
	BP: 0.96	BP: 1.01	BB: 0.85
	DD: 1.08	DD: 1.06	AB: 0.78
			BA: 0.84
			DD: 0.85
M^{pro}:CMK ($P2_1$, pH 6.0)	AP: 0.58	AP: 0.68	AA: 0.58
	BP: 0.33	BP: 0.37	BB: 0.41
	DD: 0.46	DD: 0.51	AB: 0.57
			BA: 0.40
			DD: 0.70

Superimpositions of structures were done for pairs of protomers (PA, AP, PB, BP, AA, BB and AB) and pairs of dimers (DD). The first letter corresponds to the part from the structure listed at the left of the Table superimposed and the second letter corresponds to the part from the structure listed along the top of the Table superimposed. P, protomer (applicable to *C*2 structures only); A, protomer A; B, protomer B; D, dimer.

region, whereas Ile286 is in the disallowed region. The electron densities for Tyr154 and Asn277 are not well defined. For the $M^{pro}_{+A(-1)}$:APE complex in the space group $P2_12_12_1$, there is a dimer in each asymmetric unit. Only the residues 1A to 304A and residues 1B to 300B were identified in the electron density maps. In the Ramachandran plot for this structure, Asp33 and Asn84 of both protomers are in the generously allowed regions, whereas Tyr154 and Ile286 of both protomers are in the disallowed regions. The electron density for Tyr154 is not well defined. Superimposition (see Materials and Methods) of protomers A and B yielded a root-mean-square difference (rmsd) of 0.29 Å. Positional differences (up to 4.18 Å) occur mainly among the N and C-terminal residues as well as those poorly defined residues on the flexible loops. The three structures are in close agreement (Table 2). With regard to the protomer orientation and protein fold, these structures, in general, are identical with the crystal structures of M^{pro} previously determined in the space group $P2_1$.[13]

Binding of APE to SARS-CoV M^{pro}

Crystals of SARS-CoV M^{pro} and $M^{pro}_{+A(-1)}$ were soaked in the solutions of Cbz-Leu-Phe-AGln-(*S*,*S*) EP-COOEt, Cbz-Leu-Phe-AGln-(*R*,*R*)EP-COOEt, and a racemic mixture of the *S*,*S* and *R*,*R* diastereomers (*trans*). Only the *S*,*S* diastereomer showed up in the electron density maps. APE binds in the substrate-binding regions of M^{pro} (Figure 3(a) and (b)). As visualized in all three structures, the residues forming the substrate-binding regions of both protomers of the peptidase are in the catalytically competent conformation, similar to their counterparts in the structures of the main peptidases from other coronaviruses,[15,29] and to those in protomer A of the $P2_1$ structures of SARS-CoV M^{pro}.[13] In the structure of unbound M^{pro}, the catalytic dyad has a distance of 3.7 Å between the His41 $N^{\varepsilon 2}$ atom and the Cys145 S^{γ} atom, and the Cys145 S^{γ} atom is coplanar with the atoms of the His41 imidazole ring. Superimposition of the structures of unbound M^{pro} and the M^{pro}:APE complex shows that the binding of APE does not cause any major changes in the structure of the peptidase (Table 2). The Cys145 C^{α}–C^{β} bond undergoes a 95° rotation (χ_1: from −64.0° to +30.7°) accompanying the formation of a covalent bond with a distance of 2.01 Å between the Cys145 S^{γ} atom of the peptidase and the epoxide C3 atom of APE (Figure 4(a) and (b)). The length of a C–S single bond is normally about 1.8 Å. However, with an estimated overall coordinate error (based on maximum likelihood) of 0.12 Å for the structure of the M^{pro}:APE complex, the difference between the refined and expected distances (0.2 Å) is not considered significant. This new covalent bond makes a torsion angle,

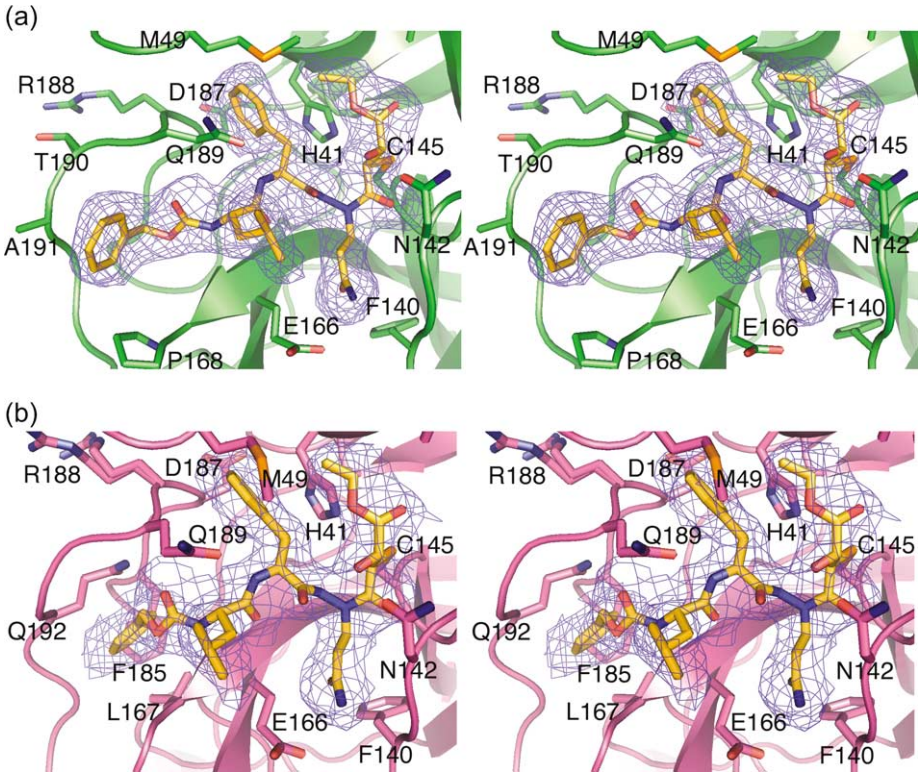

Figure 3. Binding of APE (orange) in the substrate-binding regions of SARS-CoV M^{pro}. (a) Stereo view of the outstanding density in the F_o-F_c map for the structures of the M^{pro}:APE complex and protomer B of the $M^{pro}_{+A(-1)}$:APE complex. (b) The corresponding stereo view for protomer A of the $M^{pro}_{+A(-1)}$:APE complex.

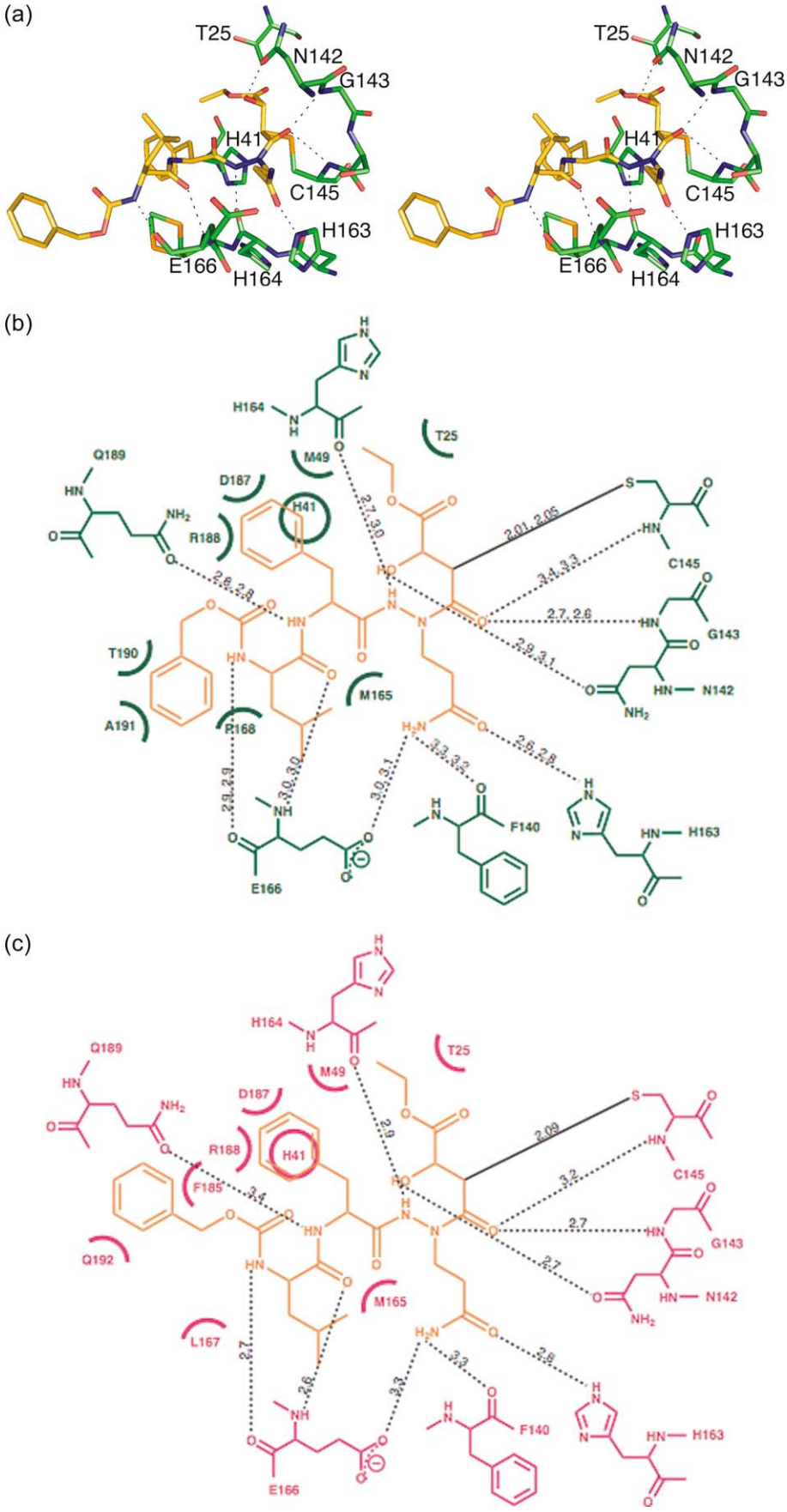

Figure 4 (*legend on next page*)

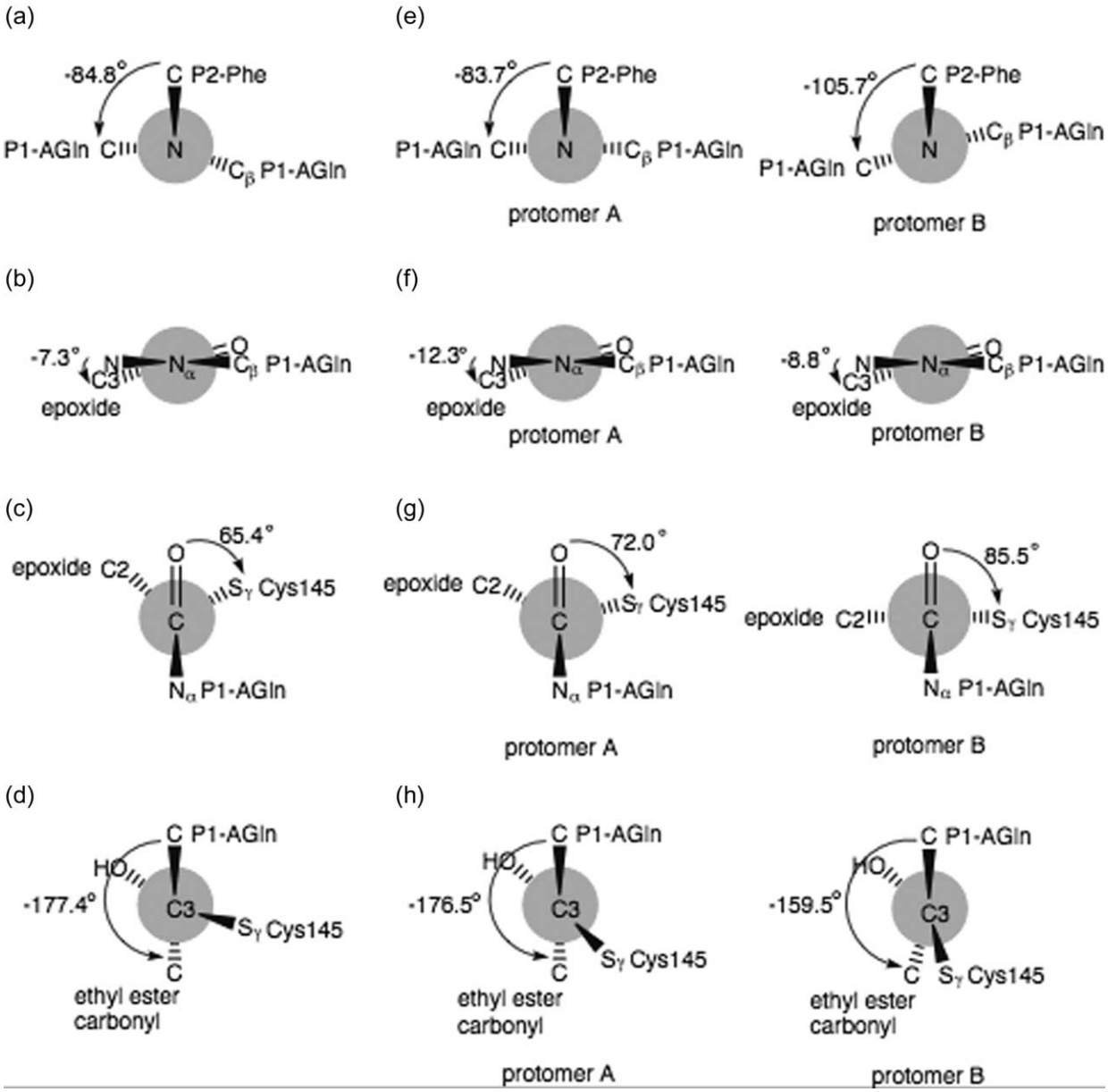

Figure 5. Newman projections. (a) The equivalent to ϕ of P1-AGln, C(P2-Phe)-N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln), (b) the equivalent to ψ of P1-AGln, N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln)-C3(epoxide), (c) the torsion angle O═C(P1-AGln)-C3(epoxide)-S^{γ}(Cys145), and (d) the torsion angle C(P1-AGln)-C3(epoxide)-C2(epoxide)-C(ethyl ester carbonyl) of APE in the structure of the SARS-CoV M^{pro}:APE complex. (e) The equivalent to ϕ of P1-AGln, C(P2-Phe)-N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln), (f) the equivalent to ψ of P1-AGln, N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln)-C3(epoxide), (g) the torsion angles O═C(P1-AGln)-C3(epoxide)-S^{γ}(Cys145) and (h) the torsion angle C(P1-AGln)-C3(epoxide)-C2(epoxide)-C(ethyl ester carbonyl) of APE in protomers A and B of the $M^{pro}_{+A(-1)}$:APE complex.

O═C(P1-AGln)-C3(epoxide)-S^{γ}(Cys145), of 65.4° (Figure 5(c)). The epoxide ring of APE opens, leaving a hydroxyl group on the C2 atom to form hydrogen bonds with the Asn142 $O^{\delta 1}$ atom of the peptidase (2.9 Å) and the P2-Phe carbonyl O atom of APE (3.2 Å) (Figure 4(a) and (b)). The configurations of the C2 and C3 atoms are inverted from *S*,*S* to *R*,*R*. The conversion at the C3 atom is due to the

Figure 4. (a) View from the floor of the S1 specificity pocket in all structures. (b) Schematic diagram for the interactions of APE with SARS-CoV M^{pro} in the structures of the M^{pro}:APE complex and protomer B of the $M^{pro}_{+A(-1)}$:APE complex. (c) The corresponding schematic diagram for protomer A of the $M^{pro}_{+A(-1)}$:APE complex. Hydrogen bonds are shown as broken lines, with their distance(s) (in Å) given in (b) and (c). In (b), the first value corresponds to the M^{pro}:APE complex and the second value corresponds to protomer B of the $M^{pro}_{+A(-1)}$:APE complex. The residues of the peptidase in contact with APE are shown as arcs.

formation of this new covalent bond, whereas that at the C2 atom is a result of the change in the ranking order of substituents (Figure 5(d)). Asn142 and Cys145 of the peptidase flank a short loop constituting the oxyanion hole, with the amide hydrogen atoms of Gly143 and Cys145 as the hydrogen-bond donors to stabilize the carbonyl O atom in the scissile peptide group of the substrate. The conformation of this loop remains essentially unchanged upon the binding of APE. The P1-AGln carbonyl O atom of APE is accommodated in the oxyanion hole of the peptidase, accepting hydrogen bonds from the amide hydrogen atoms of Gly143 and Cys145 at distances of 2.7 Å and 3.4 Å, respectively (Figure 4(a) and (b)). The P1-AGln N^{α} atom of APE remains sp^2-hybridized and adopts trigonal planar geometry. For P1-AGln, the equivalent to ϕ, C(P2-Phe)-N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln), is −84.8° (Figure 5(a)), and the equivalent to ψ, N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln)-C3(epoxide), is −7.3° (Figure 5(b)). With the error in atomic coordinates considered, the trigonal plane centered at the P1-AGln N^{α} atom is coplanar with that centered at the P1-AGln carbonyl C atom, allowing the N^{α} atom to reduce the electrophilicity of the carbonyl C atom by π-electron delocalization. The equivalent to χ_1 of P1-AGln, N-N^{α}-C^{β}-C^{γ}, is −77.4°.

Compared to the structure of the M^{pro}:APE complex, the structure of the $M^{pro}_{+A(-1)}$:APE complex shows some differences in the geometry of binding. In the latter, the rotation of the Cys145 C^{α}–C^{β} bond of the peptidase reaches a more positive value of χ_1 (protomer A, 47.5°; protomer B, 46.5°). The length of the covalent bond between the Cys145 S^{γ} atom of the peptidase and the epoxide C3 atom of APE is 2.09 Å in protomer A and 2.05 Å in protomer B (Figure 4(a) to (c)). Note that the estimated overall coordinate error (based on maximum likelihood) for this structure is 0.17 Å. O=C(P1-AGln)-C3(epoxide)-S^{γ}(Cys145) makes a torsion angle of 72.0° in protomer A and 85.5° in protomer B (Figure 5(g)). In both protomers, the configurations of the C2 and C3 atoms of APE are inverted from *S*,*S* to *R*,*R* (Figure 5(h)), and the P1-AGln N^{α} atom of APE remains sp^2-hybridized and adopts trigonal planar geometry. For the P1-AGln of APE, the equivalent to ϕ, C(P2-Phe)-N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln), is −83.7° in protomer A and −105.7° in protomer B (Figure 5(e)), and the equivalent to ψ, N(P1-AGln)-N^{α}(P1-AGln)-C(P1-AGln)-C3(epoxide), is −12.3° in protomer A and −8.8° in protomer B (Figure 5(f)). With the errors in atomic coordinates considered, the trigonal plane centered at the P1-AGln N^{α} atom is roughly coplanar with that centered at the P1-AGln carbonyl C atom in both protomers. The equivalent to χ_1 of P1-AGln, N-N^{α}-C^{β}-C^{γ}, is −82.7° in protomer A and −119.5° in protomer B.

The aza-peptide component of APE binds in the substrate-binding regions of both protomers of M^{pro} in a substrate-like manner. Consistent with the results of the secondary-structure studies for the substrates of this peptidase,[9] the structures of the M^{pro}:APE and $M^{pro}_{+A(-1)}$:APE complexes show that the main chain of the aza-peptide component of APE interacts with the main chain of the residues 164 to 166 of the peptidase through amide hydrogen–carbonyl oxygen hydrogen bonding in the manner of an anti-parallel β sheet in both protomers (Figure 4(a) to (c)).

Specificity pockets

The predominant S1 specificity of SARS-CoV M^{pro} for Gln is determined primarily by His163. In the structure of unbound M^{pro}, the His163 $N^{\varepsilon 2}$ atom interacts with a chloride ion at a distance of 3.3 Å and in the plane of the His163 imidazole ring, whereas in the structures of the M^{pro}:APE and $M^{pro}_{+A(-1)}$:APE complexes, the chloride ion is displaced by the P1-AGln side-chain amide group of APE with its $O^{\varepsilon 1}$ atom accepting a hydrogen bond from the His163 $N^{\varepsilon 2}$ atom of the peptidase (2.6 Å to 2.8 Å). Additional hydrogen bonds may be donated, though not of ideal geometry, by the P1-AGln $N^{\varepsilon 2}$ atom of APE to the Phe140 carbonyl O atom and the Glu166 $O^{\varepsilon 2}$ atom of the peptidase (Figure 4(b) and (c)). In both protomers of all three structures, the Phe140 phenyl ring interacts with the His163 imidazole ring through π-stacking (distance between the geometric centers of the rings: 3.7 to 3.8 Å), and the latter is properly oriented for its $N^{\delta 1}$ atom to accept a hydrogen bond from the Tyr161 phenolic OH group (2.9 to 3.1 Å). This hydrogen bond maintains the neutral tautomeric state of the His163 imidazole ring with its $N^{\varepsilon 2}$ atom protonated over a broad range of pH. This is crucial for the interaction of the His163 of the peptidase with the P1-Gln of the substrate.[13,15,29]

In each peptidase dimer, the integrity of the S1 specificity pocket in one protomer requires the protonated amino group of Ser1 from the other protomer. This residue is at the tip of the N-finger (N-terminal residues 1 to 7) propagating between the domain III of its parent protomer and the domain II of the partner protomer. In the structures of unbound M^{pro} and the M^{pro}:APE complex, the Ser1 amino group of each protomer forms an ionic interaction with the Glu166 side-chain carboxyl group of the other protomer (2.7 Å). The Ser1 of each protomer also interacts with the Phe140 of the other protomer through amide hydrogen–carbonyl oxygen hydrogen bonding (3.2 Å). These three residues are thus held together to form the "floor" of the S1 specificity pocket (Figure 6(a)). In both protomers of the $M^{pro}_{+A(-1)}$:APE complex, however, Ser1 is N-terminally blocked by the additional Ala, and the ionic interaction between the Ser1 amino group and Glu166 side-chain carboxyl group is lost. This is also followed by the loss of the amide hydrogen–carbonyl oxygen hydrogen bond between Ser1 and Phe140 because Ser1 is no longer properly oriented. The additional Ala is disordered, leaving Ser1 unanchored and the floor of the S1 specificity pocket partly disrupted (Figure 6(b)).

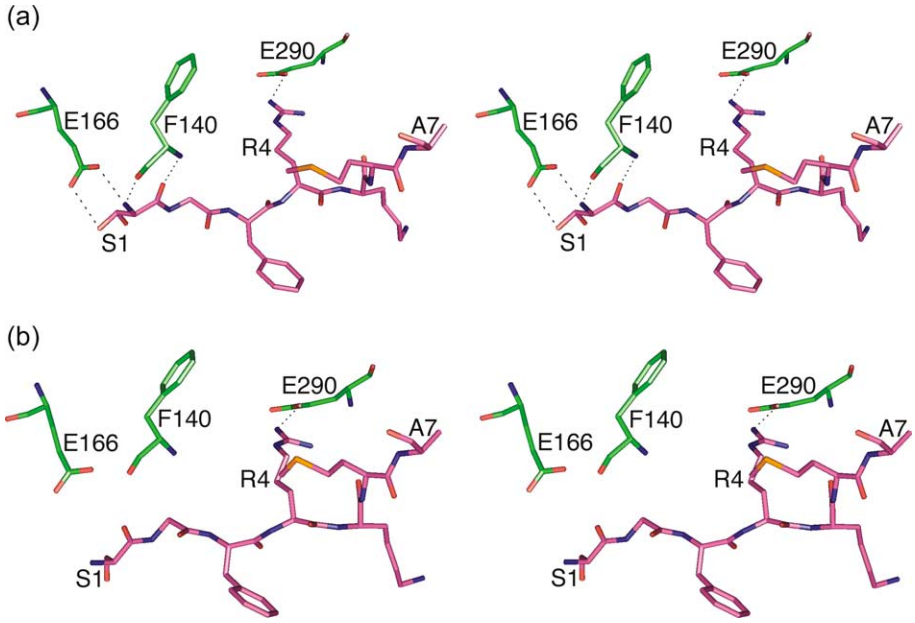

Figure 6. Floor of the S1 specificity pocket of SARS-CoV M^{pro}. Phe140, Glu166 and Glu290 of the first protomer (green) and the N-finger (N-terminal residues 1 to 7) of the second protomer (magenta) are shown. (a) In the structures of unbound M^{pro} and the M^{pro}:APE complex, Phe140 and Glu166 of the first protomer interact with Ser1 of the second protomer to form the floor of the S1 specificity pocket. (b) In the structure of the $M^{pro}_{+A(-1)}$:APE complex, the disordered N-terminal Ala of the second protomer (not shown) orients Ser1 poorly for any interaction with Phe140 and Glu166 of the first protomer. The interaction between Glu290 of the first protomer and Arg4 of the second protomer, however, remains in both (a) and (b). Hydrogen bonds are shown as broken lines.

Despite the disruption of the S1 pocket by the presence of a single additional N-terminal residue, the presence of a ten-residue affinity tag at the N terminus of M^{pro} reduced the specific activity of the peptidase by less than an order of magnitude (results not shown).

M^{pro} has greatest preference for Leu and Ile at P2, followed by Phe, Val and Met in that order.[7,9] In the structures of the M^{pro}:APE and $M^{pro}_{+A(-1)}$:APE complexes, the P2-Phe side-chain of APE fits snugly in the S2 specificity pocket of the peptidase, where the interactions are mainly hydrophobic. The P2-Phe phenyl ring of APE interacts with the His41 imidazole ring of the peptidase through π-stacking (distance between the geometric centers of the two rings: 4.3 to 4.6 Å). Superimposition of the structures of unbound M^{pro} and the M^{pro}:APE complex shows that, upon the binding of APE, the side-chain of Met49 undergoes a large conformational change, thereby opening the S2 specificity pocket for the P2-Phe side-chain of APE. Also, the side-chain of Gln189 is no longer disordered after its reorientation and formation of a hydrogen bond through its $O^{\varepsilon 1}$ atom with the P2-Phe amide hydrogen atom of APE (2.8 Å to 3.4 Å). This appears to secure the P2-Phe of APE in the S2 specificity pocket of the peptidase (Figures 3(a) and (b) and 4(a)).

In the structures of the M^{pro}:APE and $M^{pro}_{+A(-1)}$:APE complexes, the P3-Leu side-chain of APE extends into the solvent and does not have any significant interactions with the peptidase (Figures 3(a) and (b) and 4(a)). This lack of interactions is consistent with the fact that no S3 specificity of M^{pro} could be established.[7]

The aza-peptide component of APE consists of only three residues, so its benzyloxycarbonyl (Cbz) group partly takes up the space for the P4 residue of a substrate. M^{pro} has a shallow S4 specificity pocket that accommodates small side-chains (Ser, Thr, Val, Pro and Ala).[7] In the M^{pro}:APE complex and protomer B of the $M^{pro}_{+A(-1)}$:APE complex, the binding of the Cbz group of APE does not use the S4 specificity pocket of the peptidase. The position for the C^{α} atom of a P4 residue of a substrate is occupied by the O2 atom in the Cbz group of APE. In this conformation, the benzyl group of APE makes contacts with Pro168 and with residues 190 to 192 of the peptidase; the Cbz group is exposed to the solvent (Figures 3(a) and 4(b)). In contrast, in protomer A of the $M^{pro}_{+A(-1)}$:APE complex, the benzyl group of APE squeezes into and thereby widens the S4 specificity pocket of the peptidase, so that it is snugly accommodated in this enlarged pocket now formed by the residues 165 to 168, Phe185, Gln192 and the main-chain atoms of Val186 (Figures 3(b) and 4(c)).

The S1′ specificity pocket of M^{pro} is also shallow, accommodating only small side-chains (Ser, Ala, Gly, Asn and Cys).[7,9] In the design of APEs, several different epoxide derivatives were attached to the aza-peptide component to modulate the interactions of APE with the S1′ specificity pocket of a clan CD cysteine peptidase.[24,26,27] In the structures of the M^{pro}:APE and $M^{pro}_{+A(-1)}$:APE complexes, however, the S1′ specificity pocket of the peptidase

is apparently not used in the binding of APE. The epoxide C2 atom of APE sits close to the position for the C^{α} atom of the P1′ residue of a substrate and the hydroxyl group on the C2 atom is exposed to the solvent. The ethyl ester group of APE lies against the "ceiling" of the active site lined by Leu27, Pro39, His41, and the peptide group between His41 and Val42.

Dimer interface

The structures reported here show similar features at their dimer interfaces. The solvent-accessible surface area (per protomer) buried upon dimerization of SARS-CoV M^{pro} is from 1250 Å^2 to 1260 Å^2. In the structures of unbound M^{pro} and the M^{pro}:APE complex, the crystallographic 2-fold axis passes through the dimer interface and brings the opposite interacting residue-pairs into exact 2-fold symmetry (Figure 2); whereas in the structure of the $M^{pro}_{+A(-1)}$:APE complex, even in the absence of a crystallographic 2-fold axis, the dimer interface still exhibits approximate 2-fold symmetry. In all three structures, the dimer interface concentrates on one face of each protomer: that containing the residues on the N-finger, and domains II and III. The majority of the interactions occur between the residues on the N-finger and domain II. Two ionic interactions, one between the Ser1 amino group and the Glu166 side-chain carboxyl group of opposite protomers, and one between the side-chains of the Arg4 and Glu290 of opposite protomers, are observed in the structures of unbound M^{pro} and the $M^{pro}_{+A(-1)}$:APE complex (Figure 6(a)). However, in the structure of the $M^{pro}_{+A(-1)}$:APE complex, only the latter ion pair is observed (Figure 6(b)). In contrast to the N-finger immobilized at the dimer interface, the C-terminal loop (residues 301 to 306) is highly mobile as it is exposed to the solvent and anchored by only three or four solvent-exposed hydrogen bonds to some residues along the rim of the dimer interface. Interestingly, in the structures of unbound M^{pro} and the M^{pro}:APE complex, the C-terminal loop of one protomer extends towards the S1 specificity pocket of the other protomer, whereas in the structure of the $M^{pro}_{+A(-1)}$:APE complex, the C-terminal loop of protomer A (the only one visualized in this structure) propagates away from the S1 specificity pocket of protomer B.

Discussion

The kinetic data and crystal structures reported herein indicate that APEs have excellent potential as inhibitors of SARS-CoV M^{pro} and are worthy of further evaluation in the development of lead compounds for anti-SARS agents. Thus, the k_{inact}/K_i of M^{pro} for Cbz-Leu-Phe-AGln-(*S*,*S*)EP-COOEt is similar in magnitude to that of the first generation APEs produced to inhibit other cysteine peptidases.[24,26,27] Optimization of the latter has yielded inhibitors of caspases with k_{inact}/K_i well over $10^6\ M^{-1}\ s^{-1}$.[27]

The excellent specificities of APEs for clan CD cysteine peptidases[24,26,27] suggest that APEs have better potential as inhibitors of M^{pro} than do chloromethyl ketones (CMKs), the first class of potential inhibitors proposed on a structural basis. The structure of the M^{pro}:CMK complex previously determined in the space group $P2_1$ shows different and unexpected modes of binding for CMK to the two protomers of the peptidase.[13] CMKs are highly active alkylating agents and react with good nucleophiles, such as hydroxyl and thiol groups. They therefore inhibit serine peptidases as well as cysteine peptidases.[28] A recent study showed that CMKs efficiently inhibit some clan CA cysteine peptidases, such as papain and the cathepsins.[30] This casts doubt on the utility of CMKs as specific inhibitors for M^{pro}. In contrast, the structures of unbound M^{pro} and the M^{pro}:APE complex show that the aza-peptide component of APE binds to the peptidase in a substrate-like manner. The main-chain of the aza-peptide component of APE forms amide hydrogen–carbonyl oxygen hydrogen bonds with the main chain of the residues 164 to 166 of the peptidase in the manner of an anti-parallel β-sheet. The P1 and P2 side-chains of APE occupy the S1 and S2 specificity pockets of the peptidase, respectively. Based on the definitions for the binding of epoxysuccinyl peptides to clan CA cysteine peptidases,[28] this corresponds to the S and S′ binding mode, with inclination to the S binding mode because the pre-cleavage portion of the substrate-binding region of the peptidase makes the major contribution to M^{pro}:APE interactions.

The structures of the M^{pro}:APE and $M^{pro}_{+A(-1)}$:APE complexes substantiate the mechanism by which APEs have been proposed to irreversibly inhibit their target peptidases (Figure 1(a)). Nucleophilic attack at the epoxide C3, rather than the C2, atom of APE, by the Cys145 S^{γ} atom is consistent with the expected transition-state geometry for proteolysis catalyzed by M^{pro}. In caspase-3, the epoxide C3 atom is attacked (M. Grutter, unpublished results), whereas in caspase-1, the C2 atom is attacked (R. Rubin, unpublished results). In the case of epoxysuccinyl peptides binding to clan CA cysteine peptidases, the position of attack depends on the orientation of the epoxysuccinyl peptide in the substrate-binding region.[28] E-64 binds to papain in the S binding mode, similar to the mode of APE binding to M^{pro}. In the papain:E-64 complex, however, the epoxide C2 atom is the one attacked.[31] Nucleophilic attack at the epoxide C3 atom is observed in the S′ binding mode, as exemplified by CA-074 binding to cathepsin B.[32]

The kinetic data and crystal structures indicate that M^{pro} reacts only with the *S*,*S* diastereomer of the APE and not its *R*,*R* diastereomer. Interestingly, the order of inhibitory activities for APEs towards most clan CD cysteine peptidases is *S*,*S* > *R*,*R* > *trans* > *cis* (the racemic mixture of the *S*,*R* and *R*,*S* diastereomers).[24,26,27] Based on the M^{pro}:APE

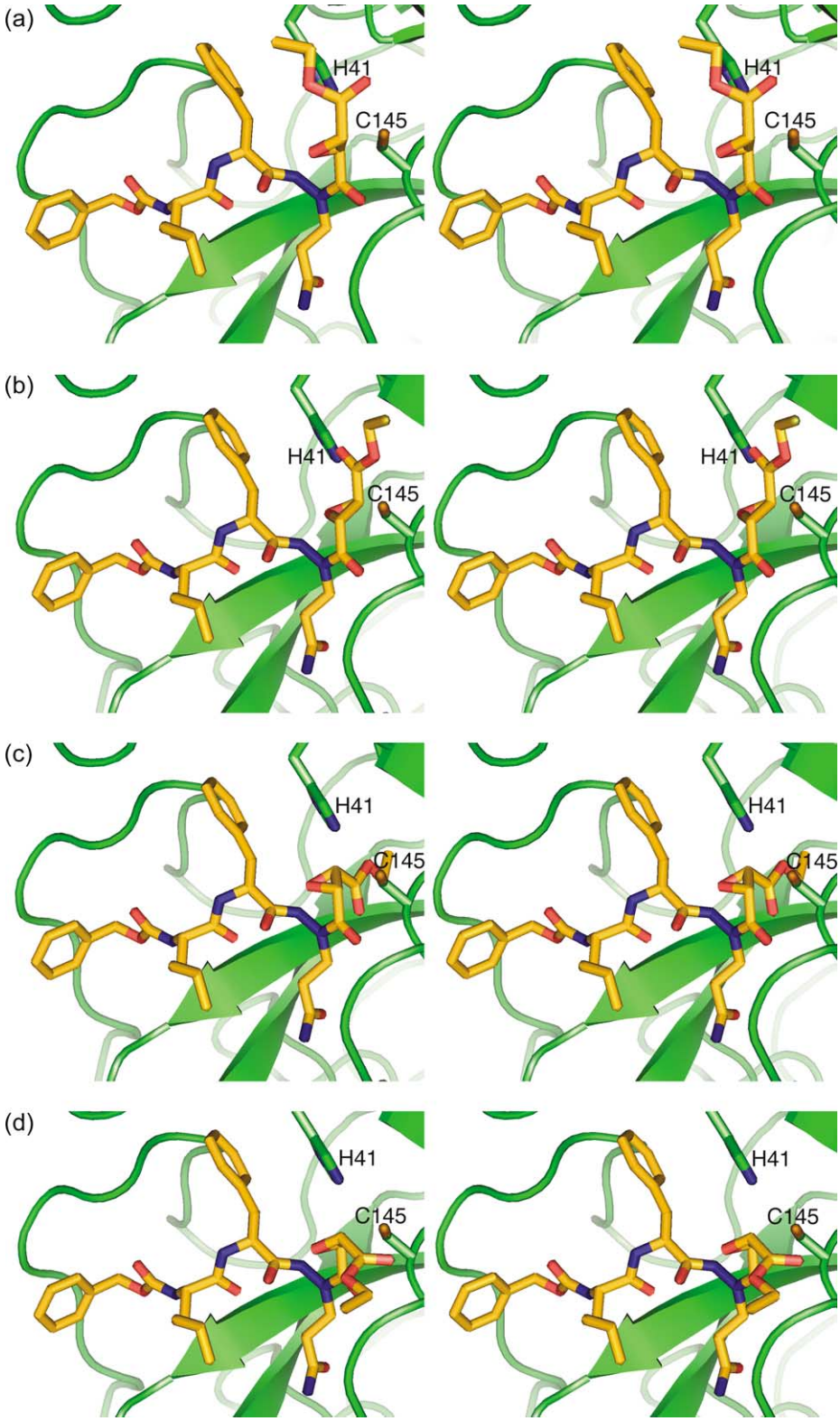

Figure 7. Models for each of the four diastereomers of Cbz-Leu-Phe-AGln-EP-COOEt binding to SARS-CoV M^{pro} before the nucleophilic attack by the Cys145 S^{γ} atom of the peptidase. (a) 2*S*, 3*S*, (b) 2*R*, 3*R*, (c) 2*S*, 3*R*, (d) 2*R*, 3*S*.

structures, we built models of each of the four diastereomers of APE at the active site of M^{pro} to explain the results of our trials (Figure 7(a) to (d)). These models show that, for APE to be accommodated in the substrate-binding regions of the peptidase, the epoxide C3 atom of APE must be in the *S* configuration, otherwise the epoxide moiety sterically clashes with the "back-wall" of the active site of the peptidase and with the aza-peptide component of APE itself (Figure 7(b) and (c)). In the

S configuration, the epoxide C3 atom of APE is also in better geometry with respect to the Cys145 S^{γ} atom for the nucleophilic attack. The epoxide C2 atom of APE must be in the *S* configuration as well to allow the interactions between the epoxide moiety and the active site of the peptidase, otherwise the epoxide moiety sterically clashes with the loop constituting the oxyanion hole of the peptidase (Figure 7(d)). The model with the *S*,*S* diastereomer of APE also indicates that the distance between the His41 $N^{\varepsilon 2}$ atom of the peptidase and the epoxide O atom is 4 Å to 5 Å, and that these two atoms are not well aligned for proton transfer (Figure 7(a)). This suggests that the opening of the epoxide ring likely involves two steps separated by a conformational rearrangement of APE: (1) the protonation of the epoxide O atom and (2) the nucleophilic attack at the epoxide C3 atom. From the model, it is not possible to determine the order in which these steps occur. However, it would be energetically more favorable for protonation to be the first step. On the other hand, M^{pro} may be sufficiently flexible in solution to allow the alignment of the His41 $N^{\varepsilon 2}$ atom and the epoxide O atom for proton transfer. In such a scenario, protonation and nucleophilic attack could occur in a concerted manner, enabling the epoxide ring to open in a single step. Much of the mechanism for the inhibition of SARS-CoV M^{pro} by APE remains to be elucidated. Rigorous treatment of this issue using methodologies in organic chemistry will be required.

All three structures successfully visualize the N-fingers of both protomers of M^{pro}. Despite the appreciable participation by the N-fingers in dimer interactions, it was shown experimentally that the deletion of the N-fingers inactivates the peptidase but has little effect on its dimerization properties. Molecular dynamic simulations suggested that the main role of the N-finger is one directing the peptidase to dimerize at an orientation facilitating the formation of substrate-binding regions in the catalytically competent conformation.[10] This very likely relies on the two ionic interactions observed in the structures of unbound M^{pro} and the M^{pro}:APE complex, one between the protonated amino group of Ser1 and the Glu166 side-chain carboxyl group of opposite protomers, and one between the side-chains of the Arg4 and Glu290 of opposite protomers. In the structure of the $M^{pro}_{+A(-1)}$:APE complex, the former ion-pair does not exist but both substrate-binding regions are still capable of accommodating APE in a manner similar to that exhibited by the structure of the M^{pro}:APE complex. This suggests that the former ionic interaction, possibly because it is relatively accessible to the solvent and therefore weaker, is of less importance.

The crystal structure of M^{pro} previously determined in the space group $P2_1$ at pH 6.0 showed the collapse of the active site and S1 specificity pocket of one of the protomers, whereas the $P2_1$ structures at pH 7.6 and 8.0 showed the recovery of the collapsed parts. Based on this trend, a pH-triggered switch for the catalytic activity of the peptidase was proposed.[13] In our study, all crystals were grown at pH 6.5 and the resulting structures show that the substrate-binding regions of both protomers are in the catalytically competent conformation. This suggests the possibility of an alternative or additional mechanism underlying the pH-dependence of the activity of the peptidase, especially at pH 6.5 or above. The change in protonation/deprotonation state of the catalytic dyad with pH is one of the possible second mechanisms. Insights into this possibility could be provided by the direct determinations of the pK_a values of the catalytic dyad by nuclear magnetic resonance (NMR) spectroscopy, as exemplified by the similar studies of the catalytic triad of α-lytic peptidase.[33]

Materials and Methods

Preparation of the protein and APEs

SARS-CoV $M^{pro}_{+A(-1)}$ was cloned, overexpressed and purified as described.[21] A clone expressing M^{pro} was generated using oligonucleotide-directed evolution to delete the codon corresponding to the N-terminal Ala of $M^{pro}_{+A(-1)}$. Using this clone, M^{pro} was overexpressed and purified essentially as described for $M^{pro}_{+A(-1)}$. Cbz-Leu-Phe-AGln-(*S*,*S*)EP-COOEt and Cbz-Leu-Phe-AGln-(*R*,*R*)EP-COOEt were synthesized using the methods established to synthesize other APEs[24,26,27] with minor modifications.

Inhibition studies

Enzymatic activity was measured by following the increase in fluorescence due to the cleavage of a fluorogenic peptide: Abz-Ser-Val-Thr-Leu-Gln-Ser-Gly-(NO_2)Tyr-Arg, where Abz is aminobenzoate and (NO_2)Tyr is nitrotyrosine. Fluorescence was measured using a Cary Eclipse Fluorescence spectrophotometer (Varian Canada, Mississauga, Ontario, Canada) equipped with a circulating water-bath. Experiments were performed using a 100 μl quartz cuvette. The standard assay contained 25 nM M^{pro}, 20 mM Bis–Tris (pH 7.0), 2 mM DTT, and was performed at 37.0($\pm$0.1) °C. The reaction was monitored using an excitation wavelength of 320 nm (5 nm bandpass) and an emission wavelength of 420 nm (10 nm bandpass). Initial velocities were determined from a least-squares analysis of the linear portion of the progress curves (at least 1 min) using Excel 2003 (Microsoft, Redmond, WA). All rates were corrected for the inner filter effect using an empirical correction.[34]

In inhibition studies, the concentration of APE was varied from 0 μM to 10 μM and the concentration of peptidic substrate was varied from 16 μM to 100 μM. These substrate and inhibitor concentrations were dictated by solubility limitations and the observed rates of inhibition. The rate of inactivation at each concentration of substrate and inhibitor, j_s, was determined by fitting equation (1)[35] to the corresponding progress curve using SCIENTIST version 2.01 (Micromath Scientific Software, Salt Lake City, UT). The parameters of inactivation, k_{inact} and K_i, were evaluated by fitting equation (2) to the j_s obtained at each concentration of S

and I,[36] using the least-squares and dynamic weighting options of LEONORA.[37]

Crystallization, crystal soaking and cryo-protection

Before crystallization, both SARS-CoV M^{pro} and $M^{pro}_{+A(-1)}$ were dialyzed against 20 mM NaCl, 20 mM Tris–HCl (pH 7.5), and concentrated to 10 mg/ml. All crystals were grown at ambient temperature by the hanging-drop, vapor-diffusion method. For the *C*2 crystals, the reservoir solution contained 50 mM ammonium acetate, 5% (w/v) polyethylene glycol (M_r 10,000), 3% ethylene glycol, 3% dimethyl sulfoxide, 1 mM dithiothreitol and 0.1 mM Mes (pH 6.5). The drop contained equal amounts of the M^{pro} solution and the reservoir solution. Block-shaped crystals grew in two to three days to a size of about 0.1 mm×0.1 mm×0.1 mm. For the $P2_12_12_1$ crystals, the reservoir solution had essentially the same composition as that for the *C*2 crystals, except the replacement of 5% polyethylene glycol (M_r 10,000) and 3% dimethyl sulfoxide by 6% polyethylene glycol (M_r 8000). The drop contained equal amounts of the $M^{pro}_{+A(-1)}$ solution and the reservoir solution. Needle-shaped crystals grew in three to five days to a size of about 0.05 mm×0.05 mm×0.5 mm. Crystals of good quality were selected and soaked overnight in drops with the same compositions as their reservoir solutions plus the APE chosen for this study at 3 mM. Cryo-protectants had essentially the same compositions as reservoir solutions, except for the inclusion of 25% (v/v) glycerol in the case of the *C*2 crystals and the increase of ethylene glycol to 25% in the case of the $P2_12_12_1$ crystals. Crystals were briefly soaked and then immediately frozen in liquid nitrogen for storage and shipment to the synchrotron beamline.

Data collection and processing, structure solution, refinement and analysis

The X-ray diffraction data from all crystals were collected at the synchrotron Beamline 8.3.1 (equipped with an ADSC-Q210 CCD detector) at the Advanced Light Source in the Lawrence Berkeley National Laboratory. All data sets were indexed, scaled and merged using DENZO and SCALEPACK.[38] Structure solution and refinement were carried out in CCP4.[39,40] All structures were solved by the molecular replacement method,[41] using the structure of unbound SARS-CoV M^{pro} at pH 8.0 in the space group $P2_1$ (PDB accession code 1UK2)[13] as the search model for the structure of unbound M^{pro} in the space group *C*2, and the structure of unbound M^{pro} in the space group *C*2 as the search model for the structure of the M^{pro}:APE complex in the space group *C*2 and the structure of the $M^{pro}_{+A(-1)}$:APE complex in the space group $P2_12_12_1$. The structures of unbound M^{pro} and the $M^{pro}_{+A(-1)}$:APE complex were solved using AMoRe,[42] and the structure of the $M^{pro}_{+A(-1)}$:APE complex was solved using MOLREP.[43] APE was located as outstanding electron densities in the substrate-binding region of the peptidase in both the F_o-F_c (contoured at 3σ and 4σ) and $2F_o-F_c$ (contoured at 1σ) maps for each structure. All structures were iteratively refined using REFMAC[44] and adjusted using XtalView/Xfit.[45] The stereochemical qualities of the final structures were assessed using PROCHECK.[46] Graphical representations of the structures were prepared using PyMOL†. Superimpositions of structures were carried out using ALIGN,[47,48] based on the main-chain atoms (amide N, C^{α}, and carbonyl C and O). The surface areas of structures were calculated using NACCESS.‡ APE–peptidase interactions and dimer interactions were analyzed using LIGPLOT and DIMPLOT,[49] respectively.

Protein Data Bank accession codes

The atomic coordinates and structure factors of all structures have been deposited in the RCSB Protein Data Bank. The accession code is 2A5A for the structure of unbound SARS-CoV M^{pro}, 2A5I for the structure of the M^{pro}:APE complex and 2A5K for the structure of the $M^{pro}_{+A(-1)}$:APE complex.

Acknowledgements

This study was supported by the Protein Engineering Network of Centres of Excellence (PENCE). M.N.G.J. holds a Canada Research Chair in Protein Structure and Function and received research grants from the Canadian Institute of Health Research (CIHR). T.-W.L. received the 75th Anniversary Award from the Faculty of Medicine and Dentistry, University of Alberta in 2004. We are grateful to the Alberta Synchrotron Institute and the Lawrence Berkeley National Laboratory for their arrangements of our access to the Beamline 8.3.1 of the Advanced Light Source. We also thank J. B. Bonanno, R. Fowler, D. Lorimer and R. Romero in Structural GenomiX, Inc. for their sharing of experience, and Y. Batsiolas, E. M. Bergmann, B. K. Biswal, C. Garen, S. Khan, and M. J. Lemieux in the laboratory of M.N.G.J. for their advice and help in the course of this study and the writing of this article.

References

1. Groneberg, D. A., Hilgenfeld, R. & Zabel, P. (2005). Molecular mechanisms of severe acute respiratory syndrome (SARS). *Respir. Res.* **6**, 8.
2. Peiris, J. S., Lai, S. T., Poon, L. L., Guan, Y., Yam, L. Y., Lim, W. *et al*. (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*, **361**, 1319–1325.
3. Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S. *et al*. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1967–1976.
4. Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S. *et al*. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1953–1966.
5. Ding, Y., Wang, H., Shen, H., Li, Z., Geng, J., Han, H. *et al*. (2003). The clinical pathology of severe acute respiratory syndrome (SARS): a report from China. *J. Pathol.* **200**, 282–289.

† http://www.pymol.org/

‡ http://wolf.bms.umist.ac.uk/naccess/

6. Thiel, V., Herold, J., Schelle, B. & Siddell, S. G. (2001). Viral replicase gene products suffice for coronavirus discontinuous transcription. *J. Virol.* **75**, 6676–6681.
7. Thiel, V., Ivanov, K. A., Putics, A., Hertzig, T., Schelle, B., Bayer, S. *et al*. (2003). Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol.* **84**, 2305–2315.
8. Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. (2000). Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **81**, 853–879.
9. Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L. *et al*. (2004). Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. *J. Biol. Chem.* **279**, 1637–1642.
10. Chen, S., Chen, L., Tan, J., Chen, J., Du, L., Sun, T. *et al*. (2005). Severe acute respiratory syndrome coronavirus 3C-like proteinase N terminus is indispensable for proteolytic activity but not for enzyme dimerization: biochemical and thermodynamic investigation in conjunction with molecular dynamic simulations. *J. Biol. Chem.* **280**, 164–173.
11. Chou, C. Y., Chang, H. C., Hsu, W. C., Lin, T. Z., Lin, C. H. & Chang, G. G. (2004). Quaternary structure of the severe acute respiratory syndrome (SARS) coronavirus main protease. *Biochemistry*, **43**, 14958–14970.
12. Shi, J., Wei, Z. & Song, J. (2004). Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors. *J. Biol. Chem.* **279**, 24765–24773.
13. Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z. *et al*. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl Acad. Sci. USA*, **100**, 13190–13195.
14. Huang, C., Wei, P., Fan, K., Liu, Y. & Lai, L. (2004). 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry*, **43**, 4568–4574.
15. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science*, **300**, 1763–1767.
16. Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162.
17. Bacha, U., Barrila, J., Velazquez-Campoy, A., Leavitt, S. A. & Freire, E. (2004). Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro. *Biochemistry*, **43**, 4906–4912.
18. Jain, R. P., Pettersson, H. I., Zhang, J., Aull, K. D., Fortin, P. D., Huitema, C. *et al*. (2004). Synthesis and evaluation of keto-glutamine analogues as potent inhibitors of severe acute respiratory syndrome 3CLpro. *J. Med. Chem.* **47**, 6113–6116.
19. Wu, C. Y., Jan, J. T., Ma, S. H., Kuo, C. J., Juan, H. F., Cheng, Y. S. *et al*. (2004). Small molecules targeting severe acute respiratory syndrome human coronavirus. *Proc. Natl Acad. Sci. USA*, **101**, 10012–10017.
20. Kao, R. Y., Tsui, W. H., Lee, T. S., Tanner, J. A., Watt, R. M., Huang, J. D. *et al*. (2004). Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics. *Chem. Biol.* **11**, 1293–1299.
21. Blanchard, J. E., Elowe, N. H., Huitema, C., Fortin, P. D., Cechetto, J. D., Eltis, L. D. & Brown, E. D. (2004). High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. *Chem. Biol.* **11**, 1445–1453.
22. Liu, Z., Huang, C., Fan, K., Wei, P., Chen, H., Liu, S. *et al*. (2005). Virtual screening of novel noncovalent inhibitors for SARS-CoV 3C-like proteinase. *J. Chem. Inf. Model.* **45**, 10–17.
23. Zhang, X. W., Yap, Y. L. & Altmeyer, R. M. (2005). Generation of predictive pharmacophore model for SARS-coronavirus main proteinase. *Eur. J. Med. Chem.* **40**, 57–62.
24. Asgian, J. L., James, K. E., Li, Z. Z., Carter, W., Barrett, A. J., Mikolajczyk, J. *et al*. (2002). Aza-peptide epoxides: a new class of inhibitors selective for clan CD cysteine proteases. *J. Med. Chem.* **45**, 4958–4960.
25. Barrett, A. J. & Rawlings, N. D. (2001). Evolutionary lines of cysteine peptidases. *Biol. Chem.* **382**, 727–733.
26. James, K. E., Gotz, M. G., Caffrey, C. R., Hansell, E., Carter, W., Barrett, A. J. *et al*. (2003). Aza-peptide epoxides: potent and selective inhibitors of Schistosoma mansoni and pig kidney legumains (asparaginyl endopeptidases). *Biol. Chem.* **384**, 1613–1618.
27. James, K. E., Asgian, J. L., Li, Z. Z., Ekici, O. D., Rubin, J. R., Mikolajczyk, J. *et al*. (2004). Design, synthesis, and evaluation of aza-peptide epoxides as selective and potent inhibitors of caspases-1, -3, -6, and -8. *J. Med. Chem.* **47**, 1553–1574.
28. Powers, J. C., Asgian, J. L., Ekici, O. D. & James, K. E. (2002). Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* **102**, 4639–4750.
29. Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J. & Hilgenfeld, R. (2002). Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **21**, 3213–3224.
30. Rozman-Pungercar, J., Kopitar-Jerala, N., Bogyo, M., Turk, D., Vasiljeva, O., Stefe, I. *et al*. (2003). Inhibition of papain-like cysteine proteases and legumain by caspase-specific inhibitors: when reaction mechanism is more important than specificity. *Cell Death Differ.* **10**, 881–888.
31. Varughese, K. I., Ahmed, F. R., Carey, P. R., Hasnain, S., Huber, C. P. & Storer, A. C. (1989). Crystal structure of a papain-E-64 complex. *Biochemistry*, **28**, 1330–1332.
32. Yamamoto, A., Tomoo, K., Hara, T., Murata, M., Kitamura, K. & Ishida, T. (2000). Substrate specificity of bovine cathepsin B and its inhibition by CA074, based on crystal structure refinement of the complex. *J. Biochem. (Tokyo)*, **127**, 635–643.
33. Bachovchin, W. W. & Roberts, J. D. (1978). Nitrogen-15 nuclear magnetic resonance spectroscopy of the state of histidine in the catalytic triad of α-lytic protease. Implications for the charge-relay mechanism of peptide-bond cleavage by serine proteases. *J. Am. Chem. Soc.* **100**, 8041–8047.
34. Liu, Y., Kati, W., Chen, C. M., Tripathi, R., Molla, A. & Kohlbrenner, W. (1999). Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem.* **267**, 331–335.
35. Tudela, J., Garcia Canovas, F., Varon, R., Garcia Carmona, F., Galvez, J. & Lozano, J. A. (1987). Transient-phase kinetics of enzyme inactivation induced by suicide substrates. *Biochim. Biophys. Acta*, **912**, 408–416.
36. Cornish-Bowden, A. (2004). *Fundamentals of Enzyme Kinetics*, Portland Press, London.
37. Cornish-Bowden, A. (1995). *Analysis of Enzyme Kinetic Data*, Oxford University Press, New York.

38. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
39. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
40. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). A graphical user interface to the CCP4 program suite. *Acta Crystallog. sect. D*, **59**, 1131–1137.
41. Rossmann, M. G. & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallog.* **15**, 24–31.
42. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.
43. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022–1025.
44. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
45. McRee, D. E. (1999). XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
46. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
47. Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. (1986). Phosphocholine binding immunoglobulin Fab McPC603: an X-ray diffraction study at 2.7 Å. *J. Mol. Biol.* **190**, 593–604.
48. Cohen, G. H. (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallog.* **30**, 1160–1161.
49. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134.

Edited by M. Guss

(Received 14 July 2005; received in revised form 2 September 2005; accepted 5 September 2005)
Available online 27 September 2005