

Structural Basis for the Development of SARS 3CL Protease Inhibitors From a Peptide Mimic to an *aza*-Decaline Scaffold

Kenta Teruya,¹ Yasunao Hattori,² Yasuhiro Shimamoto,² Kazuya Kobayashi,² Akira Sanjoh,³ Atsushi Nakagawa,⁴ Eiki Yamashita,⁴ Kenichi Akaji²

¹Department of Neurochemistry, Tohoku University Graduate School of Medicine, Aoba-Ku Sendai 980-8575, Japan

²Department of Medicinal Chemistry, Kyoto Pharmaceutical University, Yamashina-Ku, Kyoto 607-8412, Japan

³R&D Center, Protein Wave Co, Nara 631-0006, Japan

⁴Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan

Received 31 August 2015; revised 22 October 2015; accepted 2 November 2015

Published online 17 November 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22773

ABSTRACT:

Design of inhibitors against severe acute respiratory syndrome (SARS) chymotrypsin-like protease (3CL^{Pro}) is a potentially important approach to fight against SARS. We have developed several synthetic inhibitors by structure-based drug design. In this report, we reveal two crystal structures of SARS 3CL^{Pro} complexed with two new inhibitors based on our previous work. These structures combined with six crystal structures complexed with a series of related ligands reported by us are collectively analyzed. To these eight complexes, the structural basis for inhibitor binding was analyzed by the COMBINE method, which is a chemometrical analysis optimized for the protein–ligand complex. The analysis revealed that the first two latent variables gave a cumulative contribution ratio of $r^2 = 0.971$. Interestingly, scores using the second latent variables for each complex were strongly correlated with root

mean square deviations (RMSDs) of side-chain heavy atoms of Met⁴⁹ from those of the intact crystal structure of SARS-3CL^{Pro} ($r = 0.77$) enlarging the S₂ pocket. The substantial contribution of this side chain (~10%) for the explanation of pIC₅₀s was dependent on stereochemistry and the chemical structure of the ligand adapted to the S₂ pocket of the protease. Thus, starting from a substrate mimic inhibitor, a design for a central scaffold for a low molecular weight inhibitor was evaluated to develop a further potent inhibitor. © 2015 Wiley Periodicals, Inc. *Biopolymers (Pept Sci)* 106: 391–403, 2016.

Keywords: SARS 3CL protease; inhibitor; *aza*-decaline; SARS 3CL protease; crystal structure complexed with inhibitor; *aza*-decaline scaffold; COMBINE analysis for inhibitor design, hydrophobic interaction

This paper is dedicated to Emeritus Professor Haruaki Yajima on the occasion of his 90th birthday.

Correspondence to: Kenta Teruya, Division of Neurochemistry, Tohoku University Graduate School of Medicine, 2-1 Seiryō, Aoba, Sendai 980-8575, Japan; e-mail: kenta.teruya.d4@tohoku.ac.jp and Kenichi Akaji, Department of Medicinal Chemistry, Kyoto Pharmaceutical University, Yamashina-ku, Kyoto 607-8412, Japan; e-mail: akaji@mb.kyoto-phu.ac.jp

Contract grant sponsor: Japan Society for the Promotion of Science

Contract grant number: 23550195 (to K.T.) and 25460160 (to K.A.)

Contract grant sponsor: Adaptable and Seamless Technology Transfer Program through Target-driven R&D from Japan Science and Technology Agency

Contract grant number: AS251Z01976Q (to Y.H.)

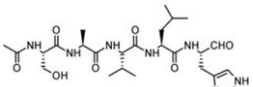
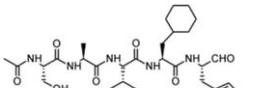
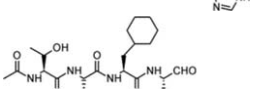
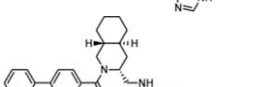
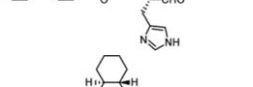
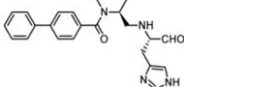
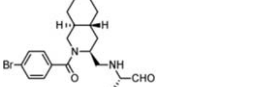
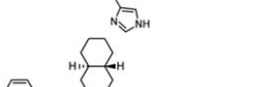
© 2015 Wiley Periodicals, Inc.

This article was originally published online as an accepted preprint. The “Published Online” date corresponds to the preprint version. You can request a copy of any preprints from the past two calendar years by emailing the *Biopolymers* editorial office at biopolymers@wiley.com.

INTRODUCTION

Although the primary epidemic of severe acute respiratory syndrome (SARS)^{1–3} was eventually brought under control, the recent identification of a SARS CoV (coronavirus)-like virus in Chinese bats^{4,5} indicates that there are natural

Table I Data Set of the Properties of an Inhibitor Co-crystallized With 3CL^{PTO}

Compound entry No.	Chemical Structure of Ligand	Complex PDB Code	pIC ₅₀	Reference
1		3AW0	5.244	8
2		3AVZ	7.187	8
3		3ATW	7.009	8
4		4WY3	3.620	24
5		4TWY	3.967	24
6		4TWW	4.201	24
7		5C5O	3.668	This work
8		5C5N	3.561	This work
9	—	3AW1	—	8

reservoirs for this virus. Since no effective therapy exists for these viral infections, developing anti-SARS agents against future outbreaks remains a formidable challenge. In the early stage of the infection, SARS 3CL protease (3CL^{PTO}) is a key enzyme to cleave polyproteins to yield functional polypeptides for replication.^{6,7} The 3CL^{PTO} is a cysteine protease containing a Cys-His catalytic dyad. Because of its functional importance in the viral lifecycle, 3CL^{PTO} is considered an attractive target for the structure-based design of drugs against SARS. Thus, numerous inhibitors of 3CL^{PTO} have been reported including peptide-mimics^{8–12} and small molecules derived from natural products,^{13–15} anti-viral agents,^{16,17} anti-malaria agents,¹⁸ or high throughput screening.^{19–22}

In the course of our own studies on SARS 3CL^{PTO} and its inhibitors,²³ we found that mature SARS 3CL^{PTO} is sensitive to degradation at the 188Arg/189Gln site, and the Arg188Ile mutant is resistant to this degradation. As a result, this stable mutant protease makes it feasible for evaluating designed and synthesized inhibitors. Exploring useful inhibitors for SARS 3CL^{PTO} based on the substrate revealed that peptide aldehyde is a facile for designing an inhibitor for SARS 3CL^{PTO}. Through combinations of the systematic substitution of peptide aldehyde amino acid and the crystallographic analysis of the complexes, a leading inhibitor (Table I, entry 1) emerged. This peptide-based inhibitor has all of its side-chain structures differing from the substrate non-prime site sequence, Thr-Ser-Ala-Val-Leu-Gln, except at the

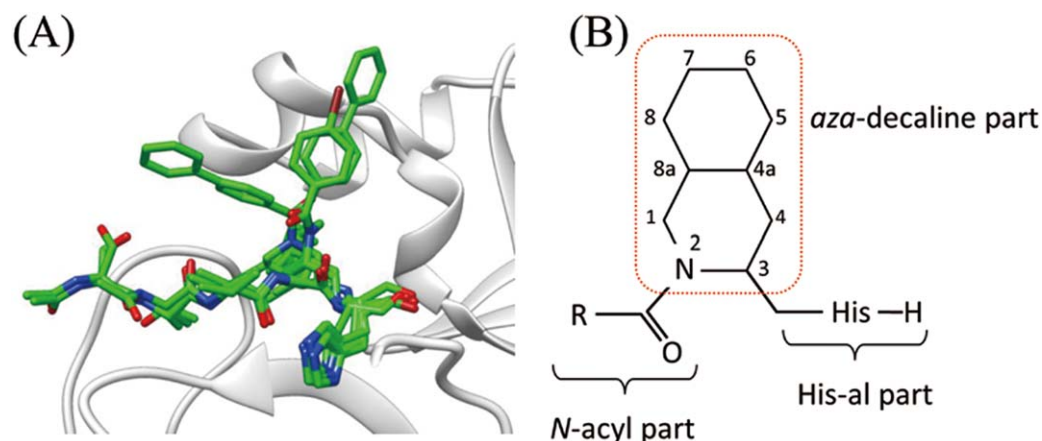


FIGURE 1 (A) The catalytic site of 3CL^{PRO}. Six structures containing inhibitors 1-6 were superimposed. The inhibitors are shown in the stick model. The S₂ pocket is located at approximately the middle position along the direction of depth. The P₃ and P₄ of peptide-based inhibitors 1-3 are extended to the left along with the groove, while N-acyl-parts of an aza-decaline based inhibitor are directed outward. (B) Schematic drawing of aza-decaline-based inhibitors 4-8. The basic structure was divided into three parts. Stereochemistry of positions 4a, 8a, and 3 and the structures of the N-acyl group are different among 4-8.

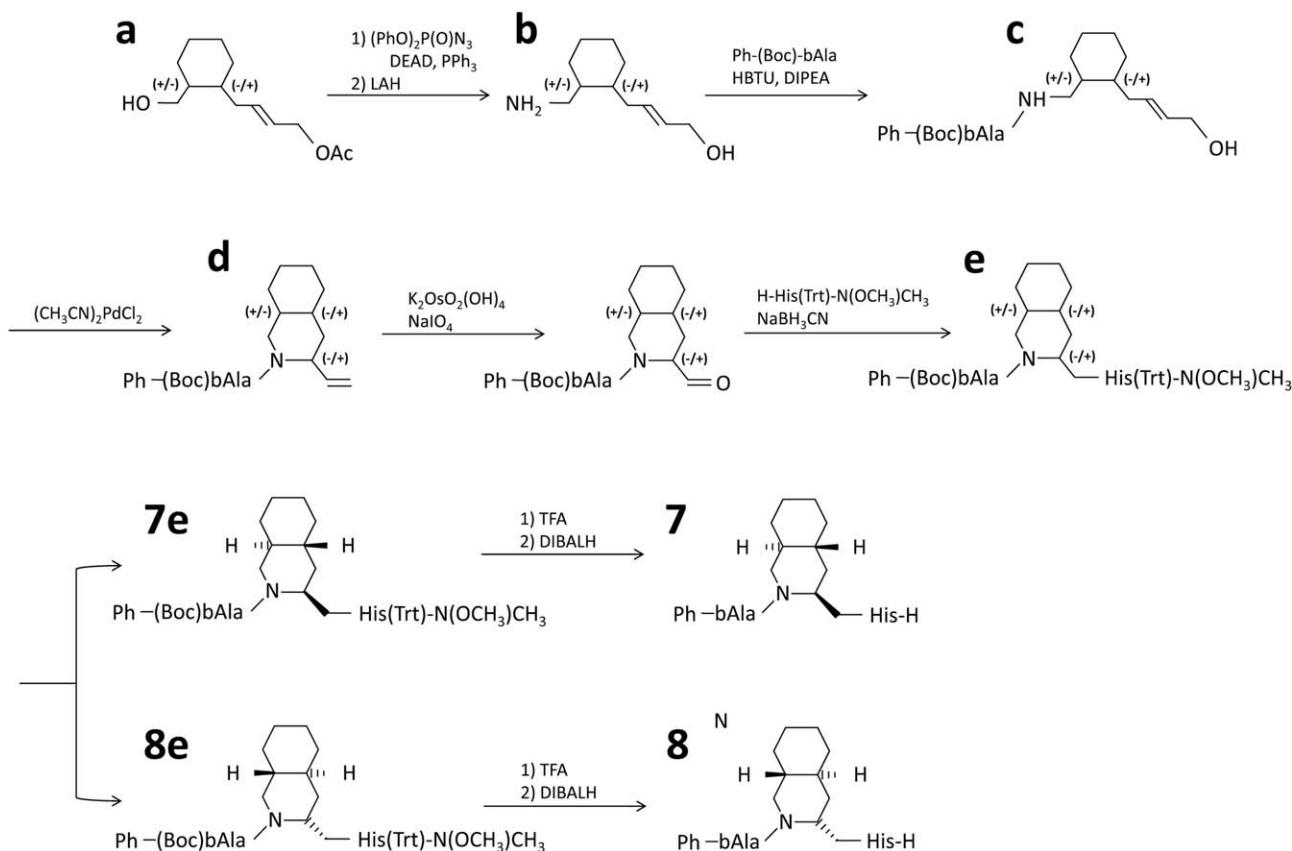
P₃ site where the side-chain was directed outward. Especially, substitution of Leu to cyclohexyl alanine (Cha) P₂ of the peptide inhibitor gave a higher increase in inhibitory potency to SARS 3CL^{PRO} (Table I, entries 1 and 2). Moreover, crystallographic analysis of the inhibitor complex provided an extended design of both deletion at the P₅ residue and mutation providing a hetero atom at the P₄ position of the inhibitor. Evaluation of IC₅₀ in the protease and crystallographic analysis revealed that this down size design was clearly realized without any substantial decrease in potency (Table I, entry 3).

Based on this wide range surveillance of the surroundings of the active sites by the peptide aldehyde inhibitor, a novel non-peptide inhibitor based on interactions at the S₁ and S₂ sites of SARS 3CL^{PRO} was designed and synthesized.²⁴ In the study, focusing on the fact that P₁ and P₂ sites interact with the catalytic dyad and hydrophobic pocket, respectively, compounds were designed by connecting the P₂ site cyclohexyl group of the substrate-based inhibitor to the main-chain at α -nitrogen atom of the P₂ position Cha via a methylene linker. This strategy gave aza-decaline as a central scaffold of the inhibitor with expected restricted conformational diversity of the P₂ site. The designed compounds showed substantial inhibitory activity. By X-ray crystallographic studies (Figure 1), it was confirmed that the aza-decaline inhibitors were at the active site cleft of 3CL^{PRO} and the aza-decaline scaffold was inserted into a large S₂ pocket, while most of the S₂ pocket was occupied. Interestingly, the stereochemistry on the aza-decaline scaffold, (4aS,8aR) and (4aR,8aS), showed clear dif-

ferences in inhibitory activities for SARS 3CL^{PRO} (Table I entries 4-6). Thus, starting from the substrate mimic peptide aldehyde as a leading compound, we have been developing nonpeptidic inhibitors by extracting fine structure-activity relationships from these crystallographic analyses. However, this stereostructure-activity relationship was not fully evaluated because of a shortage of samples for structural comparison.

The COMparative BINDing Energy (COMBINE) method developed by Ortiz²⁵ is a powerful pipeline for the decomposition of the ligand-receptor interaction energies into a series of terms, extracting latent variables (LV) for the explanation of the interactions and then, assigning weights to selected ligand-residue interactions using multivariate statistics and partial least square (PLS) analysis to correlate with a set of ligand-receptor complexes data on the experimental activities or binding affinities. Gil-Redondo and colleagues²⁶ developed gCOMBINE, which is a user-friendly graphical interface, to perform the analysis with defined input files and sophisticated result presentations. Since the ligand-induced structural alternation is taken into account, ligand-receptor interactions can be better understood by chemometrical analysis. This method has been applied to various analyses of quantitative structure-activity relationships.²⁷⁻⁴¹

To get further insight into the relations between the configuration of the side chain of the aza-decaline scaffold and the structure of the complexes, we analyzed two new complexes in which the ligands were (4aS,8aR)- and (4aR,8aS)- aza-decaline scaffold. Then, these structures and our previously reported six



SCHEME 1 Synthesis of inhibitors **7** and **8**. Abbreviations: DEAD: diethyl azodicarboxylate, DIBALH: diisobutylaluminum hydride, DIPEA: diisopropylethylamine, HBTU: *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate, LAH: lithium aluminum hydride, PPh₃: triphenylphosphine, TFA: trifluoroacetic acid.

crystal structures of a complex with a series of ligands and one crystal structure without ligands were collectively referred to derive a quantitative model.

MATERIALS AND METHODS

In the present study, two inhibitors were newly synthesized and examined for their inhibitory activity. These inhibitors were complexed with SARS 3CL^{pro} and the complexes were subjected to X-ray structural analysis. After these structural analyses, another six complex structures containing previously developed inhibitors were subjected to the gCOMBINE, a chemometric analysis. Table I shows the chemical structures, PDB IDs, and the properties of the inhibitors.

Synthesis and Determination of Inhibitory Activity of Compounds **7** and **8**

Inhibitors **7** and **8** were synthesized based on the Scheme 1 as previously described with modifications. Briefly, starting from an enantiomeric mixture of alcohol **a**, previously reported product²⁴ by Diels-Alder reaction, the primary alcohol was converted to azide and then the

amino group (**b**). Amine **b** was condensed with *N*-Boc-*N*-phenyl β -alanine (Ph-(Boc)bAla) to afford an enantio-mixture of compound **c**, a precursor of the cyclization reaction. Amide **c**, in which the nitrogen of the amide group located at 5-carbon apart from olefin, was cyclized with a Pd-based catalysis.⁴² The obtained enantio-mixture of an *aza*-decaline derivative, olefin **d** was converted to diol and then to aldehyde by Lemieux-Johnson oxidation. The resulting *aza*-decaline-derived aldehyde and H-His(Trt)-N(CH₃)OCH₃ were condensed via reductive amination to give a diastereo-mixture of compound **e**. This diastereo-mixture was separated and purified by flash column chromatography into **7e** and **8e** corresponding in stereochemistry to **7** and **8**, respectively. Trt and Boc groups of **7e** and **8e** were removed and purified by a silica column. Finally, Weinreb amides were converted to aldehyde and then purified by RP-HPLC. Yields were 3% and 1% for **7** and **8**, respectively, based on the enantio-mixture of olefin **d**.

For measurements of IC₅₀s of compounds **7** and **8**, peptide substrate, H-Thr-Ser-Ala-Val-Leu-Gln-Ser-Gly-Phe-Arg-Lys-NH₂, (111 μ M) in a reaction solution (25 μ L of 20 mM Tris-HCl buffer pH7.5 containing 7 mM DTT) was incubated with the R188I SARS 3CL^{pro} (56 nM)²³ at 37°C for 60 min in the presence of various inhibitor concentrations at 37°C for 60 min. The cleavage reaction was monitored by analytical HPLC, a linear gradient of CH₃CN (10-20%) in

an aq. 0.1% TFA over 30 min, and the cleavage rates were calculated from the decrease in the substrate peak area. Each IC₅₀ value was obtained from the sigmoidal dose–response curve. The chemical structure and potency are listed in Table I.

Compound **7e**: $[\alpha]_{\text{D}}^{28} -63.6$ (*c* 1.1, CHCl₃); ¹H NMR (400 MHz): $\delta = 7.40\text{--}7.06$ (m, 20H), 6.58–6.55 (s, 2H), 4.71–4.66 (m, 1H), 4.33–4.30 (m, 1H), 4.01–3.82 (m, 4H), 3.60 (s, 2H), 3.53 (s, 1H), 3.09 (s, 3H), 2.81–2.56 (m, 4H), 2.44–2.39 (m, 1H), 2.17–2.11 (m, 1H), 1.75–1.47 (m, 7H), 1.42 (s, 9H), 1.37–0.84 (m, 4H); ¹³C NMR (100 MHz): $\delta = 169.8, 169.2, 154.5, 154.4, 142.51, 142.46, 142.39, 140.3, 138.2, 138.0, 137.7, 137.3, 129.74, 129.71, 128.9, 128.68, 128.64, 128.54, 127.92, 127.90, 127.85, 126.89, 126.85, 125.95, 125.85, 119.3, 119.2, 80.2, 80.1, 77.2, 75.02, 75.0, 61.5, 57.8, 57.5, 53.7, 47.8, 47.3, 47.1, 47.07, 47.04, 46.2, 42.4, 41.7, 36.1, 36.0, 34.2, 32.89, 32.88, 32.6, 32.4, 32.1, 31.9, 29.8, 29.6, 28.3, 26.1, 25.8, 25.7$; HRMS (EI) Calcd. For C₅₁H₆₂N₆O₅ [M]⁺: 838.4782. Found: 838.4787.

Compound **8e**: $[\alpha]_{\text{D}}^{26} -174$ (*c* 0.33, CHCl₃); ¹H NMR (400 MHz): $\delta = 7.38\text{--}7.36$ (m, 6H), 7.34–7.25 (m, 9H), 7.20–7.13 (m, 5H), 7.08–7.06 (m, 2H), 4.90–4.62 (m, 1H), 3.87–3.77 (m, 3H), 3.64–3.41 (m, 4H), 3.19 (s, 3H), 3.01–2.30 (m, 7H), 1.73–1.55 (m, 3H), 1.38 (d, 5.6 Hz, 9H), 1.32–0.84 (m, 8H); ¹³C NMR (100 MHz): $\delta = 169.9, 169.4, 154.5, 142.5, 142.4, 138.2, 138.0, 137.5, 137.3, 129.7, 128.7, 128.6, 127.92, 127.89, 127.84, 126.9, 126.8, 125.92, 125.85, 119.2, 80.2, 80.1, 77.2, 75.01, 75.00, 61.5, 58.0, 57.7, 53.9, 42.6, 42.2, 41.7, 36.44, 36.37, 34.45, 33.4, 32.9, 32.6, 32.4, 32.3, 32.2, 29.8, 29.85, 29.80, 29.3, 28.3, 26.14, 26.08, 25.8, 25.7$; HRMS (EI) Calcd. For C₅₁H₆₂N₆O₅ [M]⁺: 838.4782. Found: 838.4773.

Compound **7**: (S)–2–[({(3S,4aR,8aS)–2–[3–(phenylamino)–propan-1-carbonyl]decahydroisoquinolin-3-yl)methyl) amino]–3–(1H-imidazol-4-yl)–propanal

¹H NMR (400 MHz): $\delta = 8.70$ (s, 1H), 7.43 (m, 1H), 7.21–7.17 (m, 2H), 6.80–6.79 (m, 2H), 5.09 (m, 1H), 4.75–4.64 (m, 2H), 3.71–3.61 (m, 1H), 3.50–3.46 (m, 2H), 3.21–3.19 (m, 2H), 3.12–3.12 (m, 2H), 2.91–2.84 (m, 1H), 2.68–2.62 (m, 1H), 1.74–0.91 (m, 12H); HRMS (EI) Calcd. For C₂₅H₃₃N₅O [M–H₂O]⁺: 419.2685. Found: 419.2689.

Compound **8**: (S)–2–[({(3R,4aS,8aR)–2–[3–(phenylamino)–propan-1-carbonyl]decahydroisoquinolin-3-yl)methyl) amino]–3–(1H-imidazol-4-yl)–propanal

¹H NMR (400 MHz): $\delta = 8.51$ (s, 1H), 7.34–7.33 (m, 1H), 7.17–7.12 (m, 2H), 6.73–6.68 (m, 2H), 5.10 (m, 1H), 4.75–4.68 (m, 2H), 3.74–3.51 (m, 1H), 3.48–3.40 (m, 2H), 3.21–3.19 (m, 2H), 3.18–3.14 (m, 2H), 2.90–2.76 (m, 1H), 2.64–2.62 (m, 1H), 1.73–0.91 (m, 12H); HRMS (EI) Calcd. For C₂₅H₃₃N₅O [M–H₂O]⁺: 419.2685. Found: 419.2682.

X-Ray Crystallography

The purified SARS 3CL^{Pro} in 20 mM Bis-Tris pH5.5, 10 mM NaCl, and 1 mM DTT was concentrated to 8 mg/mL.⁸ Crystals of SARS 3CL^{Pro} were grown at 4°C using a sitting-drop vapor diffusion method by mixing it with an equal volume of reservoir solution containing 100 mM MES pH 6.2, 5–10% PEG20000, and 5 mM DTT. Cubic-shaped crystals with dimensions of 0.3 mm × 0.3 mm × 0.3 mm grew within 3 days. The crystals were then soaked for 24 hours with a reservoir-based solution of 100 mM MES pH 6.2, 5–8% PEG20000, and 5 mM DTT containing 3 mM of **7** or **8**. Crystals were then transferred into a cryobuffer of 100 mM MES pH 6.2, 10% PEG20000, 5 mM DTT, 15% ethylene glycol containing 3 mM of **7** or **8**, and

flash-frozen in a nitrogen stream at 100 K. X-ray diffraction data of SARS 3CL^{Pro} in complexes with inhibitor **7** or **8** was collected with a SPring-8, beamline BL44XU with a Rayonix MX300HE CCD detector at a wavelength of 0.900 Å. The structures of SARS 3CL^{Pro} in complex with inhibitors were determined by molecular replacement using the Molrep⁴³ program with a R188I SARS 3CL^{Pro} structure (PDB code 3AW1⁸) as the search model. Rigid body refinement and subsequent restrained refinement protocols were performed with the program Refmac⁵⁴⁴ of the CCP package.⁴⁵ The Coot program⁴⁶ was used for manual model rebuilding. Water molecules were added using Coot only after refinement of the protein structures had converged. Ligands generated on JLigand⁴⁷ software were directly built into the corresponding difference electron density, and the model was then subjected to an additional round of refinement. The figures of structural representation in this paper were generated on Pymol⁴⁸ or UCSF-Chimera⁴⁹ software. Crystallographic data and results of refinement are summarized in Table II.

gCOMBINE Analysis

PDB⁵⁰ files of the complex (corresponding to PDB ID: 3AW0, 3AVZ, 3ATW, 4TWY, 4TWW, 4WY3, 5C5O and 5C5N)^{8,24} were prepared as containing a single complex with a ligand. The complexes were imposed on the structure of SARS 3CL^{Pro} without a ligand (PDB ID: 3AW1⁸) on the Match-Maker⁵¹ program based on the main chain C α of the protease. Each complex was protonated and minimized under an amberFF14SB⁵² force field powered by MMTK⁵³ and inhibitors were assigned AM1-BCC charges⁵⁴ and fully optimized at the AM1 level using the MOPAC program⁵⁵ on UCSF-Chimera.⁴⁹ The proteases and inhibitors were saved separately.

The parametrization for the gCOMBINE²⁶ was performed on the tLEaP module in AMBERTOOL14 program⁵⁶ suits. The PDB files of the receptor protease were cleaned by pdb4amber.⁵⁷ The antechamber suite⁵⁸ has been developed to be used with the general AMBER force field for small molecules. The PDB files of the inhibitor were converted to mol2 files by antechamber, and frmod and lib files for each inhibitor were generated by parmchk. Each file of the protease and inhibitor was combined and neutralized. For each complex, parameter and topology files were generated.

The COMBINE program was used to decompose the interaction energy between the inhibitor and protein in each complex. That is, this program was used to calculate the Lennard-Jones and electrostatic interactions between the inhibitor and each protein residue on a per residue basis. gCOMBINE is a graphical interface that runs COMBINE. The gCOMBINE automatically arranges and constructs an X matrix in which the rows represent the different compounds studied, and the columns contain the residue-based energy information, which is separated into two blocks (van der Waals and electrostatic), plus an additional column (Y matrix) containing the experimental binding affinities. This X matrix was then projected onto a small number of orthogonal LVs using partial least-squares (PLS) analysis, and the original energy terms were given weights, w_i , according to their importance in the model, in the form of PLS pseudo coefficients. The higher these coefficients are, the more significant they are for explaining the variance in the experimental data. Thus, in this study, the van der Waals interactions, u_i^{vdw} , and the electrostatic interactions, u_i^{ele} , between the inhibitor and each protein residue were selected to estimate the pIC₅₀ value. The variables that were unimportant for activities were discarded, and the remaining variables were used to build

Table II Data Collection and Refinement Statistics for 3CL^{PRO} and Its Inhibitor in Complexes With Compounds 7 and 8

PDB ID	5C5O	5C5N
Inhibitor name	7	8
Space group	C ₁₂₁	P ₁
Unit cell parameters		
Length <i>a</i>	107.80	54.66
Length <i>b</i>	82.09	58.66
Length <i>c</i>	53.27	68.10
Angle α	90	93.73
Angle β	104.35	103.40
Angle γ	90	106.51
Resolution	1.65	1.50
Observations		
Unique observations	52746	119208
Redundancy	4.0	2.2
Completeness	84.8	95.0
Mean I/sigma(I)	2.41 (at 1.65 Å)	2.36 (at 1.50 Å)
R _{merge}	0.08	0.06
Resolution range	37.59–1.69	30.2–1.50
R _{cryst}	0.29	0.25
R _{free}	0.32	0.27
RMSZ from ideal		
Bond length	0.87	0.99
Bond angle	0.92	1.01

the final PLS model. As a result of COMBINE analysis, gCOMBINE provided a summary of PLS weights, coefficients, rank, and files for PDB with coefficients. In the present study, the PLS ranks were calculated by 5.

RESULTS AND DISCUSSION

Design, Preparation, and Inhibitory Activities of Inhibitor 7 and 8

The compounds **7** and **8** were designed based on our previous structural analysis of 3CL^{PRO} compounds, **1**, **2**, **3**, **4**, **5**, and **6**. Figure 1 shows a summary of our previous structural analysis of the complex with **1-6**. The chemistry of the *aza*-decaline derived inhibitors including **7** and **8** was divided into three portions: histidine aldehyde (His-al), *aza*-decaline, and an *N*-acyl part (Figure 1B). The His-al part took an almost identical conformation and similarly interacted with the 3CL^{PRO} over the complexes examined. Briefly, His-al is located at the active center of 3CL^{PRO} as aldehyde and an imidazole ring interacted with the catalytic dyad, thiol of Cys, and imidazole of His, respectively. Thus, the His-al part functioned as a pharmacophore in these compounds. In contrast, the *N*-acyl part of the *aza*-decaline based inhibitors, **4**, **5**, and **6** were directed outward of 3CL^{PRO} and then opposite to the P₃ to P₄ sites of peptide mimic

inhibitors, **1**, **2**, and **3**. Based on these observations, compounds **7** and **8** were designed as making an aromatic ring of the *N*-acyl part feasible to access the S₃ or S₄ pocket of 3CL^{PRO} via a relatively flexible β -alanine linker and are expected to provide greater potency. The stereochemistry of compounds **7** and **8** corresponded to compounds **4** and **5**, respectively.

As for the template for the synthesis of the designed compounds, *aza*-decaline is an attractive scaffold which is restricted in configuration and can provide distinct and relative positions and directions with substituents. Considering this ability of extension, among the three parts of the basic inhibitor design, an *aza*-decaline was constructed by two cyclization steps utilizing different chemistries, Diels-Alder and diastereo-selective cyclization by PdCl₂(AcCN)₂ for cyclohexyl and *aza*-cyclohexyl on the cyclohexyl, respectively. Since the second cyclization is feasible to react with a wide variety of substrates, the acylation is scheduled in the early steps. On the other hand, since the His-al part contains a labile group, aldehyde, the reductive amination is scheduled in the late steps. In the present study, the synthetic scheme worked well at least for synthesis of inhibitors **4-8**. Production of a wide variety of inhibitors containing various stereochemistries is to be examined elsewhere based on these key reactions.^{42,59,60}

Although compounds **7** and **8** were designed based on the structural analysis mentioned earlier, the IC₅₀s were not improved in comparison with those of the corresponding compounds **4** and **5**, respectively. Compound **7** and **8** inhibited hydrolysis of SARS 3CL^{PRO} in IC₅₀, 275 μ M and 215 μ M, respectively.

Structural Analysis of R188I SARS 3CL Protease Complexed with Compounds 7 and 8

To investigate the reasons for the deterioration in IC₅₀, crystallographic analyses of the SARS 3CL^{PRO} in complexes with compounds **7** and **8** were performed. Data collection and refinement statistics for the R188I SARS 3CL protease with compounds **7** and **8** are summarized in Table II.

The overall structure of 3CL^{PRO} in complex with inhibitors **7** and **8** was similar to the other *aza*-decaline-based inhibitors (Figure 2A). Briefly, the aldehyde group and imidazole ring of His-al of these inhibitors took an almost identical conformation and similarly interacted with 3CL^{PRO}. In contrast, the direction of the *aza*-substituents of the *aza*-decaline based inhibitors, **7** and **8** (Figure 2B), were outward from 3CL^{PRO} and opposite to the P₃ to P₄ sites of peptide mimic inhibitors **1**, **2**, and **3** as in our previous inhibitors, **4**, **5**, and **6** (Figure 1A). Thus, crystallographic analyses of the SARS 3CL^{PRO} in complexes with compounds **7** and **8** revealed that in our case the bAla was not functioning as a flexible linker to reach the Ph group of the

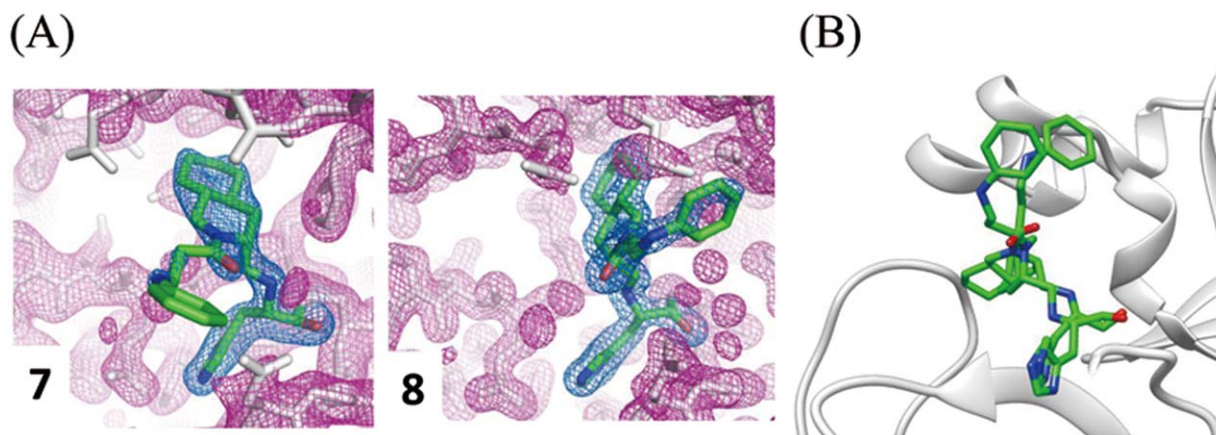


FIGURE 2 (A) Interactions of **7** (PDB code 5C5O, left panel) and **8** (PDB code 5C5N, right panel) at the active center of 3CL^{Pro}. Red and blue meshes represent observed electron density corresponding to 3CL^{Pro} and inhibitors, respectively. The inhibitors are shown in the stick model. The S₂ pocket is located at the top, and the S₁ pocket at the bottom interact with the *aza*-decaline and His-al part of the inhibitor, respectively. (B) The *N*-acyl-parts of *aza*-decaline-based inhibitors **7** and **8** are directed outward from 3CL^{Pro}. This figure is the same compositional arrangement as that of Figure 1(A).

hydrophobic S₃ pocket. These features were in good accordance with the deteriorated IC₅₀s of **7** and **8**, which may be due to a higher entropic discordance of **7** and **8** than of **4** and **5**, respectively.

The *trans aza*-decaline scaffolds of **7** and **8** were adapted in the S₂ pocket as deep as **4** and **5**, respectively. Based on these structural analyses focusing on the chemical structure of the inhibitors, several points to increase potency were revealed. The *aza*-decaline structure is sufficient to be adapted in the S₂ pocket, and stereochemistry on the *aza*-decaline is a determinant for orientation in the S₂ pocket (Figure 3). Interestingly, the orientation of *aza*-decaline in the S₂ pocket was not critical but was a determinant for relative potency; **4** and **7** were better than **5** and **8**, respectively, in which these scaffolds almost crossed perpendicular to each other when complex structures were superimposed by the whole protein. Thus, the *trans aza*-decaline scaffold is suitable for embedding into the S₂ pocket. In comparison with peptide aldehyde inhibitors (Figure 4), to make the *N*-acyl-group of *aza*-decaline based inhibitors feasible to interact with the S₃ and S₄ sites, regioisomers such as 3, 3-disubstituent or 3, 4-substituted *trans aza*-decaline may be necessary.

Paasche et al.⁶¹ have pointed out that the low inhibition potencies of known covalently interacting inhibitors may, at least in part, be attributed to insufficient fostering of the proton-transfer reaction based on MM/MQ analysis of SARS 3CL^{Pro}. Unfortunately, our method of study was incapable of accessing this charge state. Cleary and colleagues⁶² showed that

acetal or hemiacetal formation can be one of the components of the inhibitory potency for chymotrypsin. Among the eight complexes used in this study, structural refinements resulted in distances between the carbonyl carbon and sulfide ranging within 1.70–2.54 Å, which included distances judged as chemical bond formation by modeling software. There was, however, no correlation between this distance and potency ($r = -0.16$).

Quantitative Structure–Activity Relationship of R188I SARS 3CL Pro and Inhibitors

To investigate the structural basis of SARS 3CL^{Pro} based on our results, a chemometrical analysis was applied to the crystallographic structures of the complex with inhibitors **1–8**. The ligands, including peptide aldehyde, **1**, **2**, and **3**, which are large enough to cover nonprime sites from P₁ to P₅ and two pairs of diastereoisomers of *trans aza*-decaline **4** and **7** and **5** and **8**, respectively, were expected to give fine structural diversity around the active site. The inputs for the COMBINE program are structural data including topology files, coordination files, and pharmacological activity correlated with ligand binding. In this study, the IC₅₀ values were converted to negative logarithmic values, pIC₅₀s, which correlated with the binding free energy^{63,64} of the inhibitors. In Table I, values ranging from 3.56 - 7.19 are listed with the inhibitor and the ligand chemical structure.

Based on the input static structures, the program computes residue-based van der Waals and electrostatic interactions

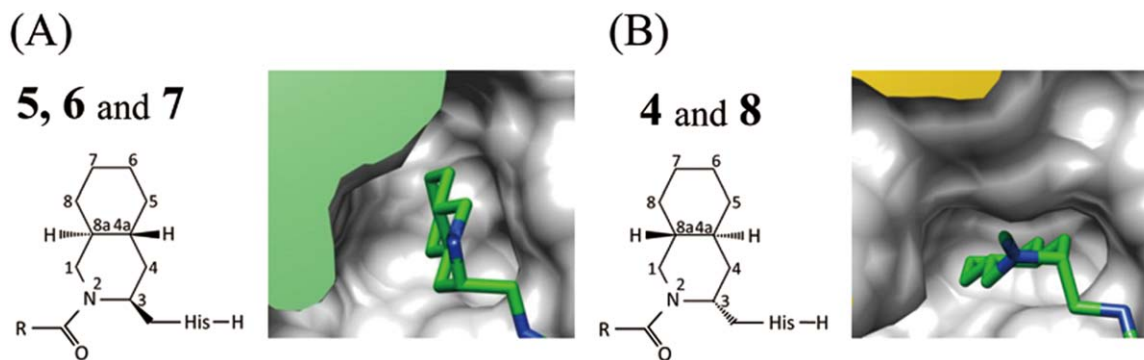


FIGURE 3 Comparison of the interaction mode of (3*S*,4*aR*,8*aS*)- and (3*R*,4*aS*,8*aR*)-types. The *aza*-decaline part (3*S*,4*aR*,8*aS*)-type includes 5, 6, and 7 (A) and (3*R*,4*aS*,8*aR*)-types include 4 and 8 (B). The left and right panels are in the same compositional arrangement in related 3CL^{PRO} and are represented as a surface model.

between the residues in each complex and ligand for each complex (Figure 5). Ligand-binding pockets from S_1 to S_4 were clearly detected. Some of the pockets were located over the structural domains. The figure revealed that S_1 and S_2 contribute to inhibitor binding more than other pockets. Extensively, even in the region of S_1 , in which all the examined inhibitors interacted with an identical His aldehyde structure, the calculated energies were not identical. This result illustrates that substitutions can affect fine interactions of the pharmacophore. These energy descriptors for each complex and the corresponding pIC_{50} comprise the X and Y matrix for the gCOMBINE analysis, respectively. The COMBINE program does PLS analysis such that the composed X matrix is then projected onto a small number of orthogonal LVs in an optimized manner for protein–ligand interaction.

Table III shows a chemometric summary of the analysis for the indices of 5 LVs. The optimal dimensionality of the PLS

models was determined by monitoring the cross-validation indexes as a function of the number of LVs extracted. The cross-validation procedure employed the leave-one-out method. The predictive ability of the resulting models was reported by both the cross-validated correlation coefficient (q^2) and the standard deviation of error in the predictions. The q^2 value served as the criterion to determine the optimal dimensionality of the PLS model. Thus, for further analysis, we decided to take ranks by 2 as our model. At this rank, the linear regression analysis reached $r^2 = 0.97$ (correlation coefficient). Similar to the primary component analysis, this indicates that the first and second LVs contribute 87% and 10% of the explanation of pIC_{50} s in this system, respectively. These results indicate that the first two LVs can explain the pIC_{50} s of complexes to a considerable extent.

The COMBINE program gave several files corresponding to each rank of LV. In fact, for each complex, predicted pIC_{50} by

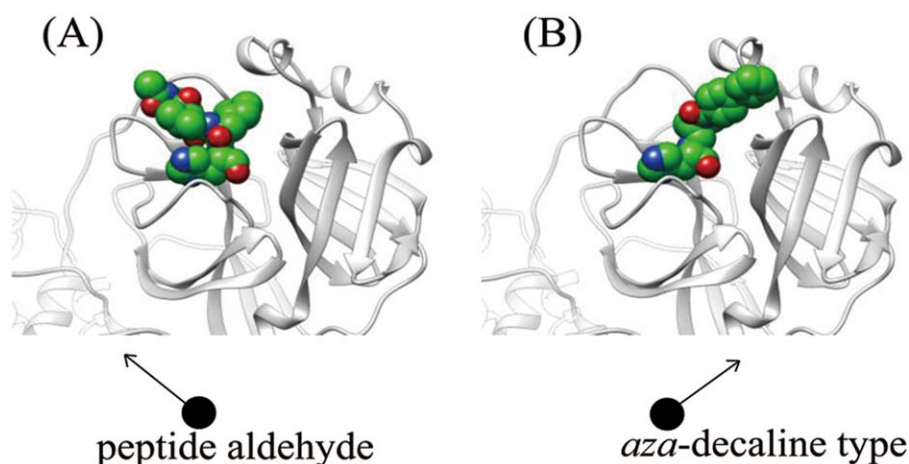


FIGURE 4 Comparison of directions of the P_3 and P_4 regions of peptide aldehyde inhibitors (A) and the *N*-acyl-part of *aza*-decaline-based inhibitors (B).

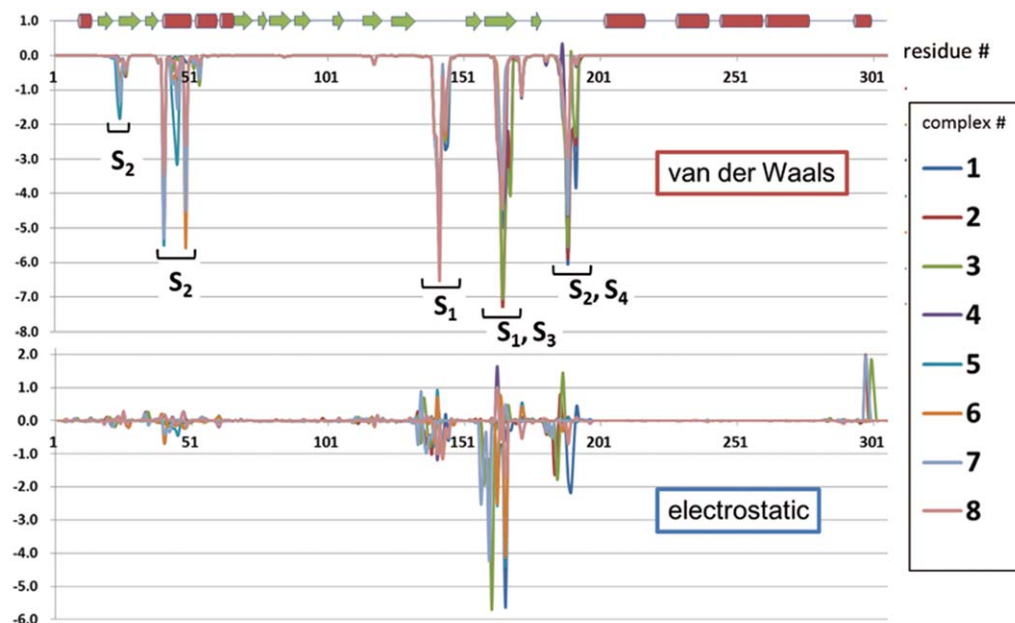


FIGURE 5 Interactions of 3CL^{Pro} and the inhibitor in each complex by residual resolution. Vertical and horizontal axes represent the computed energy of van der Waals (upper graph) and electrostatic (lower graph) interactions and the residue number of 3CL^{Pro}, respectively. At the top, the schematic drawing of the secondary structure of 3CL^{Pro} along with the residue number is presented.

the first two LVs and experimental pIC₅₀ were in good correlation and clearly correlated with the corresponding experimental pIC₅₀ as shown in Figure 6. As a consequence, these 8 complexes composed a wide range of pIC₅₀s in the order of 3.5.

P matrix (**X**-loading) in PLS to each LV reflects which and how much residue in a protein is contributing. Figure 7 shows a scatter plot of the part of the residues composing the first and second LVs in the horizontal and vertical axes, respectively. Among these elements which reflect the features of the first two LVs, the top 6 contributing residues for interaction were Pro¹⁶⁸, Glu¹⁶⁶, Met¹⁶⁵, Leu¹⁶⁷, Tyr¹⁶¹, and Gln¹⁹². On the other hand, Met⁴⁹ prominently contributed a large portion of the second LV.

For the first LV, residues mainly composing the S₁ pocket and residues covering a wide range of binding sites were extracted. Thus, the first LV can be interpreted as it stands. Briefly, the first LV for inhibition reflects the fact that the phar-

macophore interacts with active site fastening, and parts other than the pharmacophore interact with a wide range of binding sites as seen with peptide mimic inhibitors.

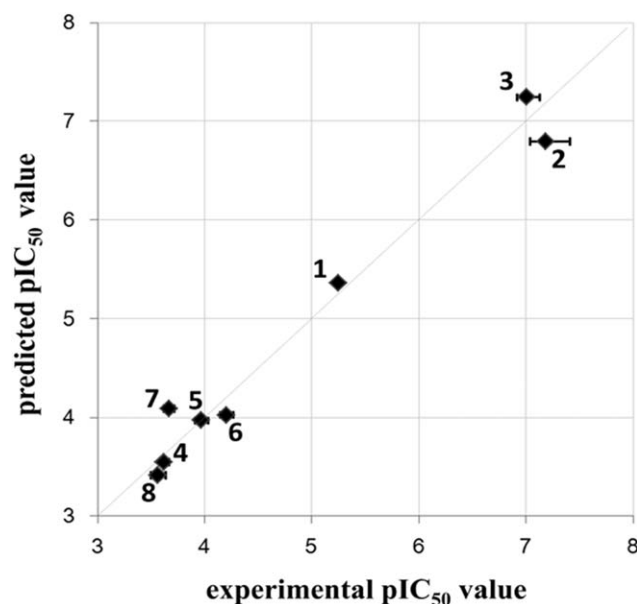


FIGURE 6 Scatter plot comparing experimental vs. predicted activities in COMBINE models for the eight complexes. The two LV model is represented. The straight line is (the diagonal) included in this plot as reference.

Table III Extracted Statistical Summary of Chemometrical Analysis of 3CL^{Pro} Complexes

Numbers of LVs	r^2	q^2
1	0.863	0.602
2	0.971	0.632
3	0.993	0.577
4	1.000	0.570
5	1.000	0.571

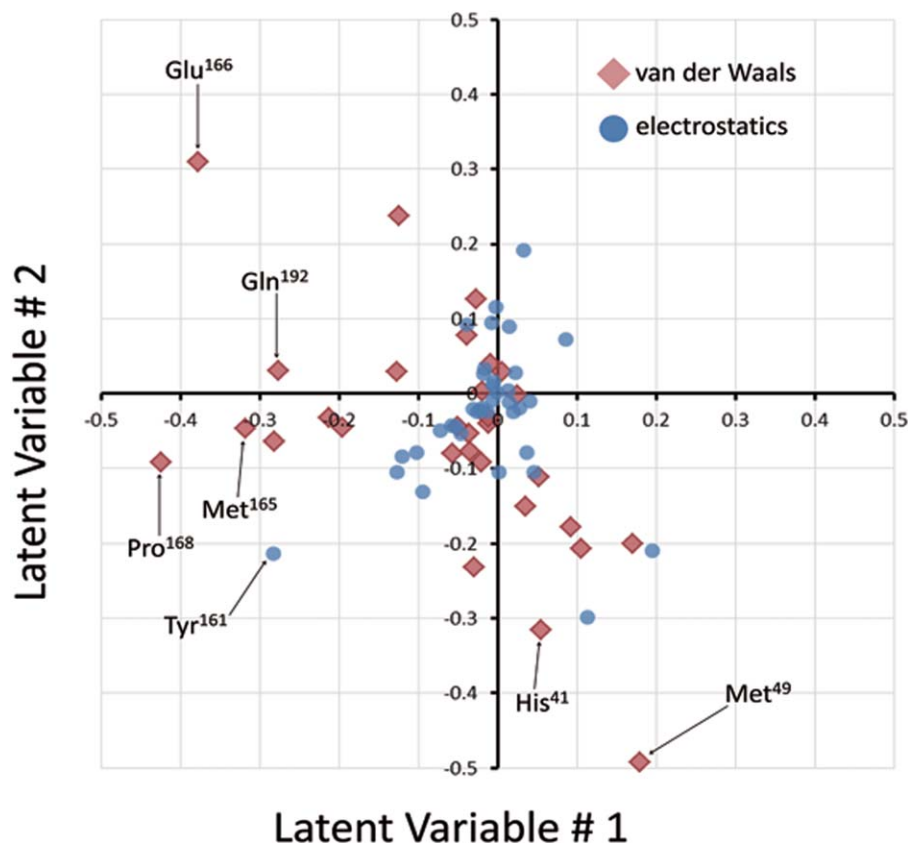


FIGURE 7 Scatter plot of amino acid residues contributing to the LVs. The vertical and horizontal axes correspond to the first and second LVs, respectively. The relevant energy descriptors have been labeled.

Next, the S_2 pocket was highlighted in the examining residues related with the second LV. Among these residues, it was revealed that configuration of the dominant Met⁴⁹ is connected to the inhibitory effect of ligand binding. Figure 8A shows that the active site superimposed on the total 3CL^{PRO} structures 1-9. Structure 9 contains no inhibitor. The catalytic Cys-His dyads were well overlapped. On the other hand, the configuration of Met⁴⁹ side chains was separated into two groups. In one group, no ligand (9), Leu side chain (1), and (3R,4aS,8aR) *aza*-decaline (4 and 8) were involved. In the other group, the Cha side chain (2 and 3) and (3S,4aR,8aS) *aza*-decaline (5, 6 and 7) were involved. Inhibitors in the latter group had greater potency in comparison with the corresponding chemical structures. Thus, the S_2 pocket is enlarged on the potent inhibitors by reorientation of the Met⁴⁹ side chain. Figure 8B shows the correlation between the X score of each complex corresponding to the second LV model and each RMSD of the Met⁴⁹ side chain heavy atom of complexes from that of structure 9. These values were strongly correlated with $r = 0.77$. Thus, the main contribution to the second LV was the orientation of the Met⁴⁹ side chain followed by enlarging the S_2 pocket. The isobutyl

group of the Leu side chain was not sufficient for inducing this reorientation. Even if a large functional group such as *aza*-decaline occupied the S_2 pocket, the angle between the plane consisting of the chair-chair configuration of *aza*-decaline and to Met⁴⁹ was critical for the configurational change, which can be controlled by the stereochemistry of position 3 of the *aza*-decaline scaffold. These analyses revealed that the *aza*-decaline scaffold with stereochemistry of *S* at position 3 is a novel scaffold for designing an inhibitor to 3CL^{PRO}. As mentioned in inhibitor-based analysis, optimized regioisomers need to be elucidated. The key feature of the interaction in the S_2 pocket gives a novel strategy for developing a new inhibitor.

It is an important and unique strategy in COMBINE that multiple structures are involved in the analysis. This COMBINE strategy may, at least in part, represent the flexibility⁶⁵⁻⁶⁷ of the protein upon ligand binding. In this research, not protein dynamic structural changes, but configurational change of side chains by inhibitor binding and their substantial control of inhibitory potency were detected. This fact highlights the advantages of the COMBINE method. Together with ligand-based analysis and the COMIBINE analysis focusing on

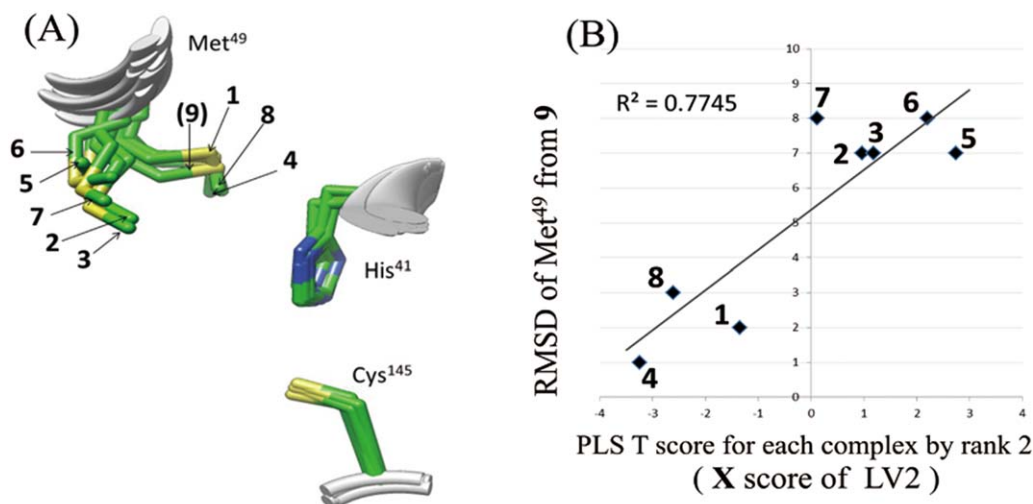


FIGURE 8 (A) Relative occupation and orientation of the catalytic dyad and Met⁴⁹ side chains. Structure of 3CL^{Pro} complexes with 1-8 and 9 were superimposed with the whole protein region. The side chains of the three residues are shown in the stick model. (B) Scatter plot showing the correlation between the second LV and orientation of the Met⁴⁹ side chain. The vertical and horizontal axes correspond to the RMSD of Met⁴⁹ from that of structure 9 and the X score of LV 2 for each complex.

protein structure, a peptide mimetic inhibitor and development of a small molecular weight molecule is feasible. These sets of training on COMBINE will be adapted to modeled complexes such as docking simulations.

In conclusion, peptidic protease inhibitors were able to increase potency when designed based on structural analysis at each designing step. For *aza*-decaline based inhibitors, the potencies decreased during design. Structural analysis revealed that orientation of the substituent on the *aza*-decaline scaffold was not suitable for interacting with the protein. The COMBINE analysis revealed that the volume of the S₂ pocket perturbed by stereochemistry of the ligand is important for inhibitor potency. This is derived as the second LV (10% explanation of the X matrix). Further design and synthesis of inhibitors are expected.

PDB ID Codes of Entry for This Article

5C5O and 5C5N

The authors thank Ms. C. Teruya of Kyoto Pharmaceutical University for obtaining EI Mass spectra.

REFERENCES

- Lee, N.; Hui, D.; Wu, A.; Chan, P.; Cameron, P.; Joynt, G. M.; Ahuja, A.; Yung, M. Y.; Leung, C. B.; To, K. F.; Lui, S. F.; Szeto, C. C.; Chung, S.; Sung, J. J. Y. *New Engl J Med* 2003, 348, 1986–1994.
- Drosten, C.; Günther, S.; Preiser, W.; van der Werf, S.; Brodt, H. R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A. M.; Berger, A.; Burguière, A. M.; Cinatl, J.; Eickmann, M.; Escriou, N.; Grywna, K.; Kramme, S.; Manuguerra, J. C.; Müller, S.; Rickerts, V.; Stürmer, M.; Vieth, S.; Klenk, H. D.; Osterhaus, A. D. M. E.; Schmitz, H.; Doerr, H. W. *New Engl J Med* 2003, 348, 1967–1976.
- Ksiazek, T. G.; Erdman, D.; Goldsmith, C. S.; Zaki, S. R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J. A.; Lim, W.; Rollin, P. E.; Dowell, S. F.; Ling, A. E.; Humphrey, C. D.; Shieh, W. J.; Guarner, J.; Paddock, C. D.; Rota, P.; Fields, B.; DeRisi, J.; Yang, J. Y.; Cox, N.; Hughes, J. M.; LeDuc, J. W.; Bellini, W. J.; Anderson, L. J. *New Engl J Med* 2003, 348, 1953–1966.
- Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J. H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H.; Zhang, J.; McEachern, J.; Field, H.; Daszak, P.; Eaton, B. T.; Zhang, S.; Wang, L. F. *Science* 2005, 310, 676–679.
- Lau, S. K. P.; Woo, P. C. Y.; Li, K. S. M.; Huang, Y.; Tsoi, H. W.; Wong, B. H. L.; Wong, S. S. Y.; Leung, S. Y.; Chan, K. H.; Yuen, K. Y. P. *Proc Natl Acad Sci USA* 2005, 102, 14040–14045.
- Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. *Science* 2003, 300, 1763–1767.
- Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, B.; Pei, J.; Liu, Y.; Chen, J.; Lai, L. *J Biol Chem* 2004, 279, 1637–1642.
- Akaji, K.; Konno, H.; Mitsui, H.; Teruya, K.; Shimamoto, Y.; Hattori, Y.; Ozaki, T.; Kusunoki, M.; Sanjoh, A. *J Med Chem* 2011, 54, 7962–7973.
- Thanigaimalai, P.; Konno, S.; Yamamoto, T.; Koivai, Y.; Taguchi, A.; Takayama, K.; Yakushiji, F.; Akaji, K.; Kiso, Y.; Kawasaki, Y.; Chen, S. E.; Naser-Tavakolian, A.; Schön, A.; Freire, E.; Hayashi, Y. *Eur J Med Chem* 2013, 65, 436–447.
- Chuck, C. P.; Chen, C.; Ke, Z.; Chi-Cheong Wan, D.; Chow, H. E.; Wong, K. B. *Eur J Med Chem* 2013, 59, 1–6.

11. Regnier, T.; Sarma, D.; Hidaka, K.; Bacha, U.; Freire, E.; Hayashi, Y.; Kiso, Y. *Bioorg Med Chem Lett* 2009, 19, 2722–2727.
12. Ghosh, A. K.; Xi, K.; Grum-Tokars, V.; Xu, X.; Ratia, K.; Fu, W.; Houser, K. V.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D. *Bioorg Med Chem Lett* 2007, 17, 5876–5880.
13. Park, J. Y.; Kim, J. H.; Kwon, J. M.; Kwon, H. J.; Jeong, H. J.; Kim, Y. M.; Kim, D.; Lee, W. S.; Ryu, Y. B. *Bioorgan Med Chem* 2013, 21, 3730–3737.
14. Ryu, Y. B.; Park, S. J.; Kim, Y. M.; Lee, J. Y.; Seo, W. D.; Chang, J. S.; Park, K. H.; Rho, M. C.; Lee, W. S. *Bioorg Med Chem Lett* 2010, 20, 1873–1876.
15. Wen, C. C.; Kuo, Y. H.; Jan, J. T.; Liang, P. H.; Wang, S. Y.; Liu, H. G.; Lee, C. K.; Chang, S. T.; Kuo, C. J.; Lee, S. S.; Hou, C. C.; Hsiao, P. W.; Chien, S. C.; Shyur, L. F.; Yang, N. S. *J Med Chem* 2007, 50, 4087–4095.
16. Ikejiri, M.; Saijo, M.; Morikawa, S.; Fukushi, S.; Mizutani, T.; Kurane, I.; Maruyama, T. *Bioorg Med Chem Lett* 2007, 17, 2470–2473.
17. Cho, J. H.; Bernard, D. L.; Sidwell, R. W.; Kern, E. R.; Chu, C. K. *J Med Chem* 2006, 49, 1140–1148.
18. Biot, C.; Daher, W.; Chavain, N.; Fandeur, T.; Khalife, J.; Dive, D.; De Clercq, E. *J Med Chem* 2006, 49, 2845–2849.
19. Liu, W.; Zhu, H. M.; Niu, G. J.; Shi, E. Z.; Chen, J.; Sun, B.; Chen, W. Q.; Zhou, H. G.; Yang, C. *Bioorgan Med Chem* 2014, 22, 292–302.
20. Lee, H.; Mittal, A.; Patel, K.; Gatuz, J. L.; Truong, L.; Torres, J.; Mulhearn, D. C.; Johnson, M. E. *Bioorgan Med Chem* 2014, 22, 167–177.
21. Jacobs, J.; Grum-Tokars, V.; Zhou, Y.; Turlington, M.; Saldanha, S. A.; Chase, P.; Egglar, A.; Dawson, E. S.; Baez-Santos, Y. M.; Tomar, S.; Mielech, A. M.; Baker, S. C.; Lindsley, C. W.; Hodder, P.; Mesecar, A.; Stauffer, S. R. *J Med Chem* 2013, 56, 534–546.
22. Ramajayam, R.; Tan, K. P.; Liu, H. G.; Liang, P. H. *Bioorgan Med Chem* 2010, 18, 7849–7854.
23. Akaji, K.; Konno, H.; Onozuka, M.; Makino, A.; Saito, H.; Nosaka, K. *Bioorgan Med Chem* 2008, 16, 9400–9408.
24. Shimamoto, Y.; Hattori, Y.; Kobayashi, K.; Teruya, K.; Sanjoh, A.; Nakagawa, A.; Yamashita, E.; Akaji, K. *Bioorgan Med Chem* 2015, 23, 876–890.
25. Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. *J Med Chem* 1995, 38, 2681–2691.
26. Gil-Redondo, R.; Klett, J.; Gago, F.; Morreale, A. *Proteins* 2010, 78, 162–172.
27. Coderch, C.; Tang, Y.; Klett, J.; Zhang, S. E.; Ma, Y. T.; Shaorong, W.; Matesanz, R.; Pera, B.; Canales, A.; Jiménez-Barbero, J.; Morreale, A.; Díaz, J. F.; Fang, W. S.; Gago, F. *Org Biomol Chem* 2013, 11, 3046
28. Coderch, C.; Klett, J.; Morreale, A.; Díaz, J. F.; Gago, F. *ChemMedChem* 2012, 7, 836–843.
29. Guo, J.; Hurley, M. M.; Wright, J. B.; Lushington, G. H. *J Med Chem* 2004, 47, 5492–5500.
30. Kim, H. J.; Chae, C. H.; Yi, K. Y.; Park, K. L.; Yoo, S. *Bioorg Med Chem* 2004, 12, 1629–1641.
31. Liu, S.; Fu, R.; Cheng, X.; Chen, S. P.; Zhou, L. H. *BMC Struct Biol* 2012, 12, 21
32. Murcia, M.; Morreale, A.; Ortiz, A. R. *J Med Chem* 2006, 49, 6241–6253.
33. Nakamura, S.; Nakanishi, I.; Kitauro, K. *Bioorg Med Chem Lett* 2006, 16, 6334–6337.
34. Ortiz, A. R.; Pastor, M.; Palomer, A.; Cruciani, G.; Gago, F.; Wade, R. C. *J Med Chem* 1997, 40, 1136–1148.
35. Peón, A.; Coderch, C.; Gago, F.; González-Bello, C. *ChemMedChem* 2013, 8, 740–747.
36. Pérez, C.; Pastor, M.; Ortiz, A. R.; Gago, F. *J Med Chem* 1998, 41, 836–852.
37. Rodríguez-Barrios, F.; Gago, F. *J Am Chem Soc* 2004, 126, 2718–2719.
38. Rotili, D.; Samuele, A.; Tarantino, D.; Ragno, R.; Musmuca, I.; Ballante, F.; Botta, G.; Morera, L.; Pierini, M.; Cirilli, R.; Nawrozki, M. B.; Gonzalez, E.; Clotet, B.; Artico, M.; Esté, J. A.; Maga, G.; Mai, A. *J Med Chem* 2012, 55, 3558–3562.
39. Silvestri, L.; Ballante, F.; Mai, A.; Marshall, G. R.; Ragno, R. *J Chem Inf Model* 2012, 52, 2215–2235.
40. Wang, T.; Wade, R. C. *J Med Chem* 2001, 44, 961–971.
41. Wang, T.; Tomic, S.; Gabdouliline, R. R.; Wade, R. C. *Biophys J* 2004, 87, 1618–1630.
42. Makabe, H.; Kong, L. K.; Hirota, M. *Org Lett* 2003, 5, 27–29.
43. Vagin, A.; Teplyakov, A. *Acta Crystallogr D* 2010, 66, 22–25.
44. Murshudov, G. N.; Vagin, A. A.; Dodson, E. *J Acta Crystallogr D* 1997, 53, 240–255.
45. 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. *Acta Crystallogr D* 2011, 67, 235–242.
46. Emsley, P.; Cowtan, K. *Acta Crystallogr D* 2004, 60, 2126–2132.
47. Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. *Acta Crystallogr D* 2012, 68, 431–440.
48. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC
49. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J Comput Chem* 2004, 25, 1605–1612.
50. Berman, H.; Henrick, K.; Nakamura, H. *Nat Struct Mol Biol* 2003, 10, 980–980
51. Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Huang, C. C.; Ferrin, T. E. *BMC Bioinform* 2006, 7, 339
52. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* 2006, 65, 712–725.
53. Hinsen, K. *J Comput Chem* 2000, 21, 79–85.
54. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J Comput Chem* 2002, 23, 1623–1641.
55. Stewart, J. J. P. *J Mol Model* 2012, 19, 1–32.
56. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J Comput Chem* 2005, 26, 1668–1688.
57. Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J Mol Biol* 1999, 285, 1735–1747.
58. Silva, A. W. S.; da, Vranken, W. F. *BMC Res Notes* 2012, 5, 367
59. Heapy, A. M.; Brimble, M. A. *Tetrahedron* 2010, 66, 5424–5431.

60. Crane, S. N.; Black, W. C.; Palmer, J. T.; Davis, D. E.; Setti, E.; Robichaud, J.; Paquet, J.; Oballa, R. M.; Bayly, C. I.; McKay, D. J.; Somoza, J. R.; Chauret, N.; Seto, C.; Scheigetz, J.; Wesolowski, G.; Massé, F.; Desmarais, S.; Ouellet, M. *J Med Chem* 2006, 49, 1066–1079.
61. Paasche, A.; Zipper, A.; Schäfer, S.; Ziebuhr, J.; Schirmeister, T.; Engels, B. *Biochemistry* 2014, 53, 5930–5946.
62. Cleary, J. A.; Doherty, W.; Evans, P.; Malthouse, J. P. G. *BBA-Proteins Proteom* 2014, 1844, 1119–1127.
63. Cheng, H. C. *J Pharmacol Toxicol* 2001, 46, 61–71.
64. Cheng, H. C. *Pharmacol Res* 2004, 50, 21–40.
65. Pastor, M.; Pérez, C.; Gago, F. *J Mol Graph Model* 1997, 15, 364–371.
66. Mou, T. C.; Gille, A.; Suryanarayana, S. R. *Mol Pharmacol* 2006, 70, 878–886.
67. Tomic, S.; Nilsson, L.; Wade, R. C. *J Med Chem* 2000, 43, 1780–1792.