Structure-Guided Design and Optimization of Dipeptidyl Inhibitors of Norovirus 3CL Protease. Structure-Activity Relationships and Biochemical, X-ray Crystallographic, Cell-Based, and In Vivo Studies

Anushka C. Galasiti Kankanamalage[^], **Yunjeong Kim**[+], **Pathum M. Weerawarna**[^], **Roxanne Adeline Z. Uy**[^], **Vishnu C. Damalanka**[^], **Sivakoteswara Rao Mandadapu**[^], **Kevin R. Alliston**[^], **Nurjahan Mehzabeen**[@], **Kevin P. Battaile**[$], **Scott Lovell**[@], **Kyeong-Ok Chang**[*,+], and **William C. Groutas**[*,^]

[^]Department of Chemistry, Wichita State University, Wichita, Kansas 67260

[+]Department of Diagnostic Medicine & Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, Kansas 66506

[@]Protein Structure Laboratory, The University of Kansas, Lawrence, Kansas 66047

[$]IMCA-CAT, Hauptman-Woodward Medical Research Institute, APS Argonne National Laboratory, Argonne, IL 60439

Abstract

Norovirus infection constitutes the primary cause of acute viral gastroenteritis. There are currently no vaccines or norovirus-specific antiviral therapeutics available for the management of norovirus infection. Norovirus 3C-like protease is essential for viral replication, consequently, inhibition of this enzyme is a fruitful avenue of investigation that may lead to the emergence of anti-norovirus therapeutics. We describe herein the optimization of dipeptidyl inhibitors of norovirus 3C-like protease using iterative SAR, X-ray crystallographic, and enzyme and cell-based studies. We also demonstrate herein *in vivo* efficacy of an inhibitor using the murine model of norovirus infection.

Introduction

Human noroviruses are the primary cause of sporadic and epidemic acute gastroenteritis in the US and worldwide,[1–3] consequently, they constitute an important public health problem, as well as a potential bioterrorism threat. Although the illness is generally considered to be mild and self-limiting, it can incapacitate infected individuals, including military troops on ships or war zones during the symptomatic phase.[4] Noroviruses are very stable in the environment and refractory to many common disinfectants, with only a few virions required

[*]authors to whom correspondence should be addressed., Department of Chemistry, Wichita State University, Wichita, KS 67260, Tel. (316) 978 7374; Fax: (316) 978 3431, bill.groutas@wichita.edu, Department of Diagnostic Medicine & Pathobiology, Manhattan, KS 66506, Tel. (785) 532 3849; Fax: (785) 532 4039, kchang@vet.ksu.edu.

to initiate virus infection and shedding, which could be a source for further contamination. Therefore, norovirus outbreaks are hard to contain using routine sanitation, and even implementation of aggressive sanitary measures often fails to prevent subsequent outbreaks.[5–6] The problem is further compounded by the current dearth of diagnostics, effective vaccines, and norovirus-specific antiviral therapeutics and/or prophylactics.[7–9]

Human noroviruses are single-stranded, positive sense RNA viruses belonging to the *Caliciviridae* family.[10] Genogroups I, II and IV of the six genogroups (GI-GVI) in the genus *Norovirus* are known to infect humans. The norovirus genome (7–8 kb) consists of three open reading frames that encode a 200 kDa polyprotein (ORF1), a major capsid protein VP1 (ORF2), and a small basic protein VP2 (ORF3).[10–11] The mature polyprotein precursor is processed by a virus-encoded 3C-like protease (3CLpro) to generate six mature non-structural proteins, including the viral protease (3CLpro or NS6[Pro]) and the RNA dependent RNA polymerase (NS7[Pol]).[12] Co- and post-translational processing of the polyprotein by norovirus 3CLpro is essential for virus replication, consequently, norovirus 3CLpro has emerged as a potential druggable target for the discovery of anti-norovirus small molecule therapeutics and prophylactics.[13–14]

Norovirus 3CLpro is a chymotrypsin-like cysteine protease with a Cys-His-Glu catalytic triad and an extended binding site.[11,15] The primary substrate specificity of the protease is for a P1 glutamine residue and a strong preference for a –D/E-F-X-L-Q-G-P-sequence (X is H, Q or V), corresponding to the subsites S_5-S_4-S_3-S_2-S_1-S_1'-S_2'-, respectively.[15–16] Cleavage is at the P_1-P_1' (Q–G) scissile bond. We have recently reported an array of norovirus inhibitors, including acyclic and cyclic sulfamide[17–19] and piperazine[20] derivatives. We have also disclosed for the first time peptidyl transition state (TS) inhibitors,[13a–e] TS mimics,[13f] as well as macrocyclic inhibitors[13g] effective in enzyme and cell based assays. We have furthermore described the first high throughput FRET assay of 3CLpro from GI and GII noroviruses as a screening tool for identifying potential protease inhibitors and have determined high resolution X-ray crystal structures of Norwalk virus (NV, a prototype strain of norovirus) 3CLpro in complex with peptidyl transition state inhibitors,[13c] as well as the first solution structure of the protease using high-field NMR.[13h] Finally, we have demonstrated proof-of-concept using the mouse model of murine norovirus (MNV) infection (*vide infra*).

In continuing our foray in this area, we describe herein the structure-based optimization of a series of dipeptidyl inhibitors of NV 3CLpro represented by structure (I) (Figure 1) using an array of X-ray crystallographic, structure-activity relationship, biochemical, cell-based, and animal studies using the mouse model of murine norovirus (MNV) infection.

Results and Discussion

Inhibitor Design Rationale

We initially focused on the design of peptidyl transition state inhibitors of NV 3CLpro that incorporate in their structure a recognition element (a peptidyl fragment) that is congruent with the known substrate specificity of the enzyme (*vide supra*) and a warhead (aldehyde or α-ketoamide), latent warhead (bisulfite adduct) or transition state mimic (α-

hydroxyphosphonate). In the case of inhibitors incorporating an aldehyde or α-ketoamide functionality in their structure, interaction with the active site cysteine (Cys139) leads to the formation of a reversible adduct (Figure 2).[17a–c]

Furthermore, in previous studies we demonstrated that norovirus 3CLpro shows a strong preference for a P2 cyclohexylalanine and, consequently, a P_2 cyclohexyl alanine residue, as well as a glutamine surrogate,[21] were incorporated in the structures of the inhibitors. The key binding interactions between norovirus 3CLpro and inhibitor were revealed by determining the high resolution X-ray crystal structure of NV 3CLpro with bound inhibitor (I) (R_1=cyclohexylmethyl, R_2=H, X=CH(OH)SO$_3$Na). The co-crystal structure of the complex showed that, under the crystallization conditions used, the bisulfite reverted to the precursor aldehyde which subsequently formed a tetrahedral adduct with the active site cysteine (Cys139) (Figure 3). Inspection of the co-crystal structure revealed opportunities for additional binding interactions with a more efficient use of chemical structure. Specifically, one such opportunity was recognized by observing that there is a particular stretch of residues spanning Ala159, Ala160, Thr161 and Lys162 that are within 4.0 Å from the benzyl ring (Figure 4). Thus, it was envisaged that the incorporation of an appropriate functional group into the phenyl ring could serve as a locus for the formation of a hydrogen bond with the Thr161 or Lys162 backbone and/or side chains, stabilizing the flexible benzyl segment and increasing binding affinity. The formation of a halogen bond also appeared plausible.[22] Thus, a small focused library of compounds was synthesized to evaluate these hypotheses, the ultimate goal being the identification of a dipeptidyl lead candidate suitable for conducting preclinical studies.

Chemistry

The synthesis of compounds **13–44** is outlined in Scheme 1. Refluxing cyclohexylalanine methyl ester hydrochloride (or leucine methyl ester hydrochloride) with trichloromethyl chloroformate yielded the corresponding isocyanate which was reacted with an appropriately substituted benzyl alcohol to yield a carbamate adduct methyl ester that was hydrolyzed to the corresponding acid with lithium hydroxide in aqueous THF. Subsequent coupling with glutamine surrogate methyl ester hydrochloride[21] afforded the desired dipeptidyl ester which was then reduced to the corresponding alcohol with lithium borohydride. Dess-Martin oxidation followed by flash chromatography purification yielded pure dipeptidyl aldehyde. The enantiomeric purity of the aldehyde was consistently high, with the amount of epimerized aldehyde ranging between 0–10%, depending on the structure of the dipeptidyl aldehyde. Further reaction of the aldehyde with diethyl phosphite in the presence of diisopropyl ethyl amine yielded the corresponding α-hydroxyphosphonate as a mixture of epimers.[23] The corresponding bisulfite adducts were readily obtained as white solids by stirring the aldehydes with sodium bisulfite in an ethyl acetate/water mixture.[24] Reaction of the aldehyde with cyclopropyl isonitrile followed by Dess-Martin oxidation of the α-hydroxy cyclopropyl amide yielded the desired α-ketoamides. The synthesized compounds are listed in Table 1.

The synthesized compounds were evaluated for their inhibitory activity against NV 3CLpro, as well as their anti-norovirus activity in a cell-based replicon system, as previously

described.[13] The IC_{50} and ED_{50} values are the average of at least two determinations and are listed in Table 1. The enzyme selectivity of a select number of inhibitors was evaluated against a panel of representative proteases and the results are summarized in Table 2. The methodologies employed in conducting the enzyme assays and inhibition studies were as described previously by us[13] and others.[25] The in vivo efficacy of compound *16* was evaluated in the murine model of norovirus infection.

X-ray crystallography was used to elucidate the nature of the interaction of compound *17* with NV 3CLpro. The X-ray crystal structure of NV 3CLpro revealed the presence of prominent difference electron density with the substructure of *17* that is equivalent to precursor aldehyde inhibitor *16* covalently bound to Cys 139. However, no electron density was observed for the hydroxyphosphonate group that should be present for inhibitor *17* (Figure 5). Instead, the structure of the NV 3CLpro-ligand complex was found to correspond to the covalent adduct of precursor aldehyde inhibitor *16* and NV 3CLpro. In addition, the m-chlorobenzyl ring was partially disordered so the electron density for this region of the inhibitor was somewhat ambiguous. The interactions between NV 3CLpro and inhibitor *16* are shown in Figure 6. The m-chlorophenyl ring of inhibitor *16* occupies a hydrophobic pocket near Ile 109 and Val 168. Van der Waals and electrostatic surface representations of the inhibitor and NV 3CLpro are shown in Figure 7.

Inspection of the results shown in Table 1 reveals that, in general, dipeptidyl inhibitors that incorporate in their structure an aldehyde or aldehyde bisulfite adduct (masked aldehyde) display lower ED_{50} values than the corresponding compounds bearing an α-hydroxyphosphonate or α-ketoamide moiety. Furthermore, the aldehyde and aldehyde bisulfite adduct-derived inhibitors had comparable ED_{50} values. These results are consistent with previous studies which showed that the activity of the aldehyde bisulfite adducts parallels the activity of the precursor aldehydes.[13d,f] Other things being equal, inhibitors having a cyclohexyl alanine P2 residue are more potent than the corresponding t-butyl glycine or leucine inhibitors (compare compounds *16, 21* and *22*). As revealed by the X-ray crystal structure (Figure 7), the cyclohexyl alanine side chain optimally fills the hydrophobic S2 subsite, while the leucine and t-butyl alanine side chains do not. A range of ring substituents were explored to enhance pharmacological activity. Compounds with a halogen at the meta position were particularly effective (compounds *16, 27, 29, 33*). These observations are suggestive of the involvement of a halogen bond between the halogen and the C=O of Ala160, however, this could not be demonstrated with absolute certainty because of the partially disordered m-chlorobenzyl moiety.

In order to evaluate the specific interactions between the enzyme and the α-hydroxyphosphonate inhibitors, inhibitor *17* was incubated with NV 3CLpro in Tris buffer, pH 8.0. Surprisingly, the enzyme-inhibitor complex formed corresponded to the complex formed between aldehyde *16* and 3CL protease. The formation of this complex indicates initial non-enzymatic (buffer)- or enzyme-catalyzed conversion of α-hydroxyphosphonate *17* to the corresponding aldehyde *16*, followed by the formation of a tetrahedral adduct with the active site cysteine (Cys139). Studies aimed at ascertaining the validity of this hypothesis are in progress. The ensemble of hydrogen bonding and hydrophobic interactions

in the enzyme-ligand complex are near identical to those observed with aldehyde inhibitor (I), where R_1=cyclohexylmethyl, R_1=H, X=CHO) (Figure 3).

An initial evaluation of selectivity was carried out using a panel of proteases and four representative inhibitors (*16, 17, 29* and *30*). As shown in Table 2, these compounds displayed minimal activity toward the proteases of the digestive and blood coagulation cascade systems. Selectivity was uniformly lower against the serine protease human neutrophil elastase (HNE).

Compound *16* was used to obtain a preliminary indication of in vivo efficacy. All compounds that were effective against NV in cell culture were also effective against MNV-1 in the range 0.08 to 5 μM. Among them compound *16* was found to be the most effective with an EC_{50} value of 80 nM. It is noted that compound *16* was also most effective against NV (Table 1). In mice, treatment with compound *16* significantly reduced the virus titers in the small and large intestines at 3 days post virus infection by 42.12- and 7.98-fold, respectively, compared to the untreated control group (Figure 8). The overall virus titers in the small intestine were higher than those in the large intestine, and the reduction of viral titers was also greater in the small intestine than the large intestine.

In conclusion, these studies describe the optimization of the pharmacological activity and selectivity of dipeptidyl inhibitors of norovirus 3CL protease by employing iterative medicinal chemistry/SAR studies, X-ray crystallography, *in vitro* and cell-based screening, and *in vivo* efficacy studies. Importantly, these studies have identified lead compounds that show efficacy in the murine model of norovirus infection and, consequently, are suitable for conducting further pre-clinical studies.

Experimental Section

General

Reagents and dry solvents were purchased from various chemical suppliers (Aldrich, Acros Organics, Chem-Impex, TCI America, and Bachem) and were used as obtained. Silica gel (230–450 mesh) used for flash chromatography was purchased from Sorbent Technologies (Atlanta, GA). Thin layer chromatography was performed using Analtech silica gel plates. Visualization was accomplished using UV light and/or iodine. NMR spectra were recorded in $CDCl_3$ or DMSO-d_6 using a Varian XL-400 spectrometer. Melting points were recorded on a Mel-Temp apparatus and are uncorrected. High resolution mass spectrometry (HRMS) was performed at the University of Kansas Mass Spectrometry lab using an LCT Premier mass spectrometer (Waters, Milford, MA) equipped with a time of flight mass analyzer and an electrospray ion source. The purity of the compounds was established using HPLC and was >95%.

Synthesis of amino acid methyl esters 1. General procedure

To a 250 mL RB flask (oven dried and purged with nitrogen) was added absolute methanol (30 mL) and the solution was cooled to 0 °C in an ice bath while kept under a nitrogen atmosphere. Thionyl chloride (8 mL) was added to the cooled methanol with stirring, followed by the addition of the amino acid (100 mmol). The ice bath was replaced by a

water bath and the reaction mixture was heated to ~50 °C for 3 h with stirring. Removal of the solvent left a white residue which was washed with diethyl ether (250 mL) and collected by vacuum filtration to yield the amino acid methyl ester hydrochloride *1* as a white solid.

Synthesis of amino acid methyl ester isocyanates 2. General procedure

Amino acid methyl ester hydrochloride (100 mmol) was placed in a dry 500-mL RB flask and then dried overnight on the vacuum pump. The flask was flushed with nitrogen and dry dioxane (200 mL) was added followed by trichloromethyl chloroformate (29.67 g, 150 mmol), and the reaction mixture was refluxed for 10 h. The solvent was removed on the rotary evaporator and the residue was vacuum distilled to yield pure isocyanate *2* as a colorless oil.

Synthesis of substituted benzyl carbamates 3. General procedure

A solution of substituted benzyl alcohol (20 mmol) in dry acetonitrile (15 mL) was treated with triethylamine (4.05 g, 40 mmol) followed by the amino acid methyl ester isocyanate (20 mmol). The resulting solution was refluxed for 2 h and then allowed to cool to room temperature. The solution was concentrated and the residue was taken up in ethyl acetate (75 mL). The organic layer was washed with 5% HCl (2×20 mL) and brine (20 mL). The organic layer was dried over anhydrous sodium sulfate, filtered and concentrated, leaving a colorless oil (compound *3*).

Synthesis of acids 4. General procedure

A solution of ester *3* (20 mmol) in tetrahydrofuran (30 mL) was treated with 1M LiOH (40 mL). The reaction mixture was stirred for 3 h at room temperature and the disappearance of the ester was monitored by TLC. Most of the solvent was evaporated off and the residue was diluted with water (25 mL). The solution was acidified to pH ~3 using 5% hydrochloride acid (20 mL) and the aqueous layer was extracted with ethyl acetate (3×100 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered, and concentrated to yield compound *4* as a colorless oil.

Synthesis of compounds 5. General Procedure

To a solution of compound *4* (10 mmol) in dry DMF (20 mL) was added EDCI (2.40 g, 12.5 mmol, 1.25 eq), HOBt (1.92 g, 12.5 mmol, 1.25 eq) and the mixture was stirred for 30 minutes at room temperature. In a separate flask, a solution of deprotected glutamine surrogate *12* (2.23 g, 10 mmol) in DMF (15 mL) cooled to 0–5 °C was treated with diisopropylethylamine (DIEA) (9.5 g, 40 mmol, 4 eq), stirred for 30 minutes, and then added to the reaction mixture containing acid. The reaction mixture was stirred for 12 h while monitoring the reaction by TLC. The solvent was removed and the residue was partitioned between ethyl acetate (200 mL) and 10% citric acid (2×40 mL). The ethyl acetate layer was further washed with saturated aqueous $NaHCO_3$ (40 mL), followed by saturated NaCl (50 mL). The organic layer was dried over anhydrous sodium sulfate, filtered and concentrated to yield a yellow-colored oily product. Purification by flash chromatography yielded ester *5* as a white solid.

Synthesis of alcohols 6. General procedure

To a solution of ester **5** (5 mmol) in anhydrous THF (30 mL) was added lithium borohydride (2M in THF, 7.5 mL, 15 mmol) dropwise, followed by absolute ethyl alcohol (15 mL), and the reaction mixture was stirred at room temperature overnight. The reaction mixture was then acidified by adding 5% HCl and the pH adjusted to ~2. Removal of the solvent left a residue which was taken up in ethyl acetate (100 mL). The organic layer was washed with brine (25 mL), dried over anhydrous sodium sulfate, filtered, and concentrated to yield compound **6** as a white solid.

Synthesis of aldehydes 7. General procedure

Compound **6** (5 mmol) was dissolved in anhydrous dichloromethane (50 mL) under a nitrogen atmosphere and cooled to 0°C. Dess-Martin periodinane reagent (3.18 g, 7.5 mmol, 1.5 eq) was added to the reaction mixture with stirring. The ice bath was removed and the reaction mixture was stirred at room temperature for 3 h (monitoring by TLC indicated complete disappearance of the starting material). A solution of 10% aqueous sodium thiosulfate (20 mL) was added and the solution was stirred for another 15 minutes. The aqueous layer was removed and the organic layer was washed with 10% aqueous sodium thiosulfate (20 mL), followed by saturated aqueous sodium bicarbonate (2×20 mL), water (2×20 mL) and brine (20 mL). The organic layer was dried over anhydrous sodium sulfate, filtered and concentrated. The yellow residue was purified by flash chromatography (silica gel/methylene chloride/ethyl acetate/methanol) to yield a white solid **7**.

2-Chlorobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (13)—Yield (78%), mp 58–60 °C. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 9.41 (s, 1H), 8.47 – 8.52 (m, 1H), 7.57 – 7.69 (m, 1H), 7.44 – 7.51 (m, 2H), 7.32 – 7.40 (m, 2H), 5.74 – 5.77 (m, 1H), 5.02 – 5.19 (m, 2H), 4.06 – 4.22 (m, 2H), 3.00 – 3.17 (m, 2H), 2.19 – 2.31 (m, 1H), 2.08 – 2.17 (m, 1H), 1.84 – 1.94 (m, 1H), 1.54 – 1.76 (m, 8H), 1.39 – 1.52 (m, 3H), 1.08 – 1.19 (m, 2H), 0.81 – 0.93 (m, 2H). HRMS (ESI) calcd for C$_{24}$H$_{32}$ClN$_3$O$_5$Na: [M+Na]$^+$: 500.1928. Found: 500.1914.

3-Chlorobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (16)—Yield (86%), mp 54–56 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.46 (s, 1H), 8.35 (d, J = 5.47 Hz, 1H), 7.33 (br. s., 1H), 7.27 – 7.31 (m, 1H), 7.22 – 7.27 (m, 1H), 7.16 – 7.21 (m, 1H), 6.36 – 6.49 (m, 1H), 5.60 (d, J = 8.20 Hz, 1H), 5.05 (s, 2H), 4.17 – 4.41 (m, 2H), 3.15 – 3.38 (m, 2H), 2.24 – 2.50 (m, 2H), 1.84 – 2.06 (m, 2H), 1.74 – 1.83 (m, 2H), 1.45 – 1.73 (m, 6H), 1.30 – 1.43 (m, 1H), 1.02 – 1.27 (m, 3H), 0.76 – 1.01 (m, 2H). HRMS (ESI) calcd for C$_{24}$H$_{32}$ClN$_3$O$_5$Na: [M+Na]$^+$: 500.1928. Found: 500.1917.

3-Chlorobenzyl ((S)-1-cyclohexyl-2-oxo-2-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)ethyl)carbamate (21)—Yield (81%), mp 51–53 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.49 (s, 1H), 8.49 – 8.56 (m, 1H), 7.38 (br. s., 1H), 7.27 – 7.30 (m, 2H), 7.23 (d, J = 4.30 Hz, 1H), 5.71 – 5.78 (m, 1H), 5.44 – 5.51 (m, 1H), 5.08 (d, J = 13.28 Hz, 2H), 4.25 – 4.33 (m, 1H), 4.14 – 4.21 (m, 1H), 3.37 (d, J = 9.37 Hz, 2H), 2.36 – 2.52 (m, 2H), 1.95 (d, J = 6.64 Hz, 1H), 1.88 (br. s., 1H), 1.76 (d, J = 10.94 Hz, 2H), 1.68 (br. s.,

2H), 1.58 (s, 4H), 1.21 – 1.34 (m, 2H), 1.13 (d, J = 12.89 Hz, 2H). HRMS (ESI) calcd for $C_{23}H_{30}ClN_3O_5Na$: $[M+Na]^+$: 486.1772. Found: 486.1758.

3-Chlorobenzyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (22)—Yield (82%), mp 53–55 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.45 (s, 1H), 8.51 – 8.54 (m, 1H), 7.90 – 7.94 (m, 1H), 7.36 – 7.43 (m, 1H), 7.13 – 7.33 (m, 3H), 5.26 – 5.32 (m, 1H), 5.02 – 5.15 (m, 2H), 4.25 – 4.37 (m, 2H), 3.29 – 3.41 (m, 2H), 2.39 – 2.47 (m, 1H), 1.83 – 1.99 (m, 2H), 1.66 – 1.78 (m, 3H), 1.57 (m, 2H), 0.97 (d, J = 6.16 Hz, 6H).. HRMS (ESI) calcd for $C_{21}H_{28}ClN_3O_5Na$: $[M+Na]^+$: 460.1615. Found: 460.1611.

4-Chlorobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (24)—Yield (67%), mp 56–58 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.48 (s, 1H), 8.31 (d, J = 5.47 Hz, 1H), 7.27 – 7.35 (m, 4H), 6.01 (br. s., 1H), 5.34 (d, J = 8.20 Hz, 1H), 5.03 – 5.11 (m, 2H), 4.28 – 4.39 (m, 2H), 3.35 (t, J = 8.20 Hz, 2H), 2.35 – 2.51 (m, 2H), 1.90 – 1.98 (m, 1H), 1.79 – 1.90 (m, 2H), 1.60 – 1.76 (m, 5H), 1.53 (ddd, J = 5.47, 8.98, 14.06 Hz, 1H), 1.23 – 1.30 (m, 2H), 1.08 – 1.22 (m, 2H), 0.81 – 1.03 (m, 3H). HRMS (ESI) calcd for $C_{24}H_{32}ClN_3O_5Na$: $[M+Na]^+$: 500.1928. Found: 500.1917.

2-Fluorobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (25)—Yield (81%), mp 61–63 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.49 (s, 1H), 8.35 – 8.53 (m, 1H), 8.24 – 8.30 (m, 1H), 7.40 (t, J = 7.23 Hz, 1H), 7.27 – 7.34 (m, 1H), 7.10 – 7.17 (m, 1H), 7.05 (t, J = 9.18 Hz, 1H), 5.82 (br. s., 1H), 5.27 – 5.34 (m, 1H), 5.14 – 5.25 (m, 2H), 4.25 – 4.43 (m, 2H), 3.26 – 3.44 (m, 2H), 2.31 – 2.57 (m, 2H), 1.93 – 2.00 (m, 2H), 1.79 – 1.91 (m, 2H), 1.47 – 1.76 (m, 5H), 1.38 (d, J = 2.73 Hz, 1H), 1.07 – 1.29 (m, 3H), 0.84 – 1.03 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{32}FN_3O_5Na$: $[M+Na]^+$: 484.2224. Found: 484.2214.

3-Fluorobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (27)—Yield (66%), mp 53–55 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.48 (s, 1H), 8.40 (d, J = 5.08 Hz, 1H), 7.27 – 7.35 (m, 1H), 7.06 – 7.14 (m, 1H), 6.99 (dt, J = 2.34, 8.59 Hz, 2H), 5.97 (br s, 1H), 5.41 (d, J = 8.20 Hz, 1H), 5.06 – 5.14 (m, 2H), 4.25 – 4.41 (m, 2H), 3.29 – 3.41 (m, 2H), 2.27 – 2.54 (m, 2H), 1.90 – 2.00 (m, 1H), 1.79 – 1.89 (m, 2H), 1.60 – 1.78 (m, 6H), 1.48 – 1.58 (m, 1H), 1.32 – 1.46 (m, 1H), 1.08 – 1.30 (m, 3H), 0.85 – 1.04 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{32}FN_3O_5Na$: $[M+Na]^+$: 484.2224. Found: 484.2207.

3-Bromobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (29)—Yield (83%), mp 52–55 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.48 (s, 1H), 8.24 – 8.67 (m, 1H), 7.53 (br s, 1H), 7.35 – 7.47 (m, 1H), 7.25 – 7.29 (m, 2H), 7.22 (d, J = 7.42 Hz, 1H), 5.66 – 5.74 (m, 1H), 5.08 (br s, 2H), 4.23 – 4.40 (m, 2H), 3.36 (d, J = 8.59 Hz, 2H), 2.32 – 2.53 (m, 2H), 1.95 (br. s., 1H), 1.84 (br. s., 2H), 1.69 (d, J = 9.37 Hz, 5H), 1.57 (br. s., 3H), 1.32 – 1.46 (m, 1H), 1.08 – 1.28 (m,

2H), 0.84 – 1.05 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{32}BrN_3O_5Na$: [M+Na]$^+$: 544.1423. Found: 544.1410.

3-Iodobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (33)—Yield (80%), mp 56–58 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.48 (s, 1H), 8.38 – 8.44 (m, 1H), 7.73 (br s, 1H), 7.64 (d, J = 7.81 Hz, 1H), 7.31 (d, J = 7.81 Hz, 1H), 7.05 – 7.11 (m, 1H), 5.86 (br s, 1H), 5.34 (d, J = 7.81 Hz, 1H), 5.05 (s, 2H), 4.27 – 4.39 (m, 2H), 3.31 – 3.40 (m, 2H), 2.32 – 2.53 (m, 2H), 1.91 – 1.99 (m, 2H), 1.79 – 1.90 (m, 2H), 1.69 (d, J = 7.81 Hz, 6H), 1.47 – 1.58 (m, 1H), 1.38 (br s, 1H), 1.09 – 1.29 (m, 2H), 0.84 – 1.04 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{32}IN_3O_5Na$: [M+Na]$^+$: 592.1284. Found: 592.1295.

2-Methoxybenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (35)—Yield (64%), mp 56–58 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.49 (br. s., 1H), 8.16 – 8.24 (m, 1H), 7.30 – 7.37 (m, 2H), 7.29 (s, 1H), 6.93 (t, J = 7.42 Hz, 1H), 6.88 (d, J = 8.20 Hz, 1H), 5.80 (br s, 1H), 5.11 – 5.24 (m, 2H), 4.28 – 4.40 (m, 2H), 3.83 (s, 3H), 3.27 – 3.39 (m, 2H), 2.40 (d, J = 12.89 Hz, 2H), 1.96 (br s, 2H), 1.78 – 1.90 (m, 2H), 1.64 – 1.76 (m, 4H), 1.60 (s, 1H), 1.53 (dd, J = 4.10, 9.57 Hz, 1H), 1.39 (br s, 1H), 1.08 – 1.30 (m, 3H), 0.85 – 1.04 (m, 2H). HRMS (ESI) calcd for $C_{25}H_{35}N_3O_6Na$: [M+Na]$^+$: 496.2424. Found: 496.2432.

3-Methoxybenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (36)—Yield (76%), mp 60–62 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.48 (s, 1H), 8.32 (d, J = 5.95 Hz, 1H), 7.22 – 7.26 (m, 1H), 6.88 – 6.95 (m, 2H), 6.84 (d, J = 8.24 Hz, 1H), 6.48 (br s, 1H), 5.56 (d, J = 8.39 Hz, 1H), 5.08 (s, 2H), 4.22 – 4.44 (m, 2H), 3.80 (s, 3H), 3.24 – 3.36 (m, 2H), 2.29 – 2.51 (m, 2H), 1.87 – 2.08 (m, 2H), 1.76 – 1.86 (m, 1H), 1.59 – 1.75 (m, 6H), 1.46 – 1.58 (m, 1H), 1.38 (br. s., 1H), 1.06 – 1.28 (m, 3H), 0.82 – 1.04 (m, 2H). HRMS (ESI) calcd for $C_{25}H_{35}N_3O_6Na$: [M+Na]$^+$: 496.2424. Found: 496.2418.

3-Cyanobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (37)—Yield (73%), mp 58–60 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.45 (s, 1H), 8.53 (d, J = 5.47 Hz, 1H), 7.65 – 7.76 (m, 2H), 7.54 – 7.62 (m, 1H), 7.43 – 7.50 (m, 1H), 6.24 – 6.50 (m, 1H), 5.60 – 5.75 (m, 1H), 5.02 – 5.35 (m, 2H), 4.45 – 4.56 (m, 1H), 4.24 – 4.41 (m, 1H), 3.18 – 3.43 (m, 2H), 2.29 – 2.59 (m, 2H), 2.07 – 2.25 (m, 2H), 1.49 – 1.99 (m, 6H), 1.30 – 1.46 (m, 1H), 1.08 – 1.23 (m, 4H), 0.83 – 1.01 (m, 3H). HRMS (ESI) calcd for $C_{25}H_{32}N_4O_5Na$: [M+Na]$^+$: 491.2270. Found: 491.2258.

3-((tert-Butoxycarbonyl)amino)benzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (38)—Yield (65%), mp 106–110 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.52 (s, 1H), 9.41 – 9.46 (m, 1H), 8.40 – 8.46 (m, 1H), 7.32 – 7.38 (m, 1H), 7.24 (br s, 1H), 7.22 (s, 1H), 7.20 (s, 1H), 6.92 – 6.98 (m, 1H), 5.42 – 5.51 (m, 1H), 4.98 (s, 2H), 4.22 – 4.41 (m, 2H), 3.30 (d, J = 7.20 Hz, 2H), 2.22 – 2.50 (m, 2H), 2.03 (s, 1H), 1.86 – 1.94 (m, 1H), 1.76 (d, J = 8.09 Hz, 3H), 1.56

– 1.72 (m, 4H), 1.48 (s, 9H), 1.30 – 1.40 (m, 1H), 1.04 – 1.25 (m, 4H), 0.91 (br s, 2H). HRMS (ESI) calcd for $C_{29}H_{42}N_4O_7Na$: $[M+Na]^+$: 581.2951. Found: 581.2957.

3-Nitrobenzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (39)—Yield (70%), mp 52–54 °C. ^{1}H NMR (400 MHz, CDCl₃) δ 9.45 (s, 1H), 8.67 (d, J = 5.08 Hz, 1H), 8.30 (s, 1H), 8.16 (d, J = 8.20 Hz, 1H), 7.66 (d, J = 7.81 Hz, 1H), 7.50 – 7.55 (m, 1H), 5.85 (br s, 1H), 5.38 (d, J = 8.20 Hz, 1H), 5.11 – 5.30 (m, 2H), 4.31 – 4.40 (m, 1H), 4.24 (d, J = 6.25 Hz, 1H), 3.37 (dd, J = 6.05, 8.79 Hz, 2H), 2.33 – 2.52 (m, 2H), 1.86 – 1.97 (m, 2H), 1.47 – 1.84 (m, 6H), 1.31 – 1.44 (m, 2H), 1.06 – 1.28 (m, 3H), 0.82 – 1.04 (m, 3H). HRMS (ESI) calcd for $C_{24}H_{32}N_4O_7Na$: $[M+Na]^+$: 511.2169. Found: 511.2153.

Benzyl ((S)-4-methyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)pentan-2-yl)carbamate (40)—Yield (83%), mp 57–59 °C. ^{1}H NMR (400 MHz, CDCl₃) δ 9.53 (br s, 1H), 8.32 – 8.45 (m, 1H), 7.32 – 7.46 (m, 5H), 5.78 – 5.85 (m, 1H), 5.34 (br s, 1H), 5.09 – 5.18 (m, 2H), 4.35 (br s, 2H), 3.30 – 3.45 (m, 2H), 2.30 – 2.59 (m, 2H), 1.99 (d, J = 6.64 Hz, 1H), 1.83 – 1.93 (m, 1H), 1.69 – 1.81 (m, 1H), 1.58 – 1.69 (m, 2H), 1.23–1.30 (m, 1H), 0.94 (d, J= 6.15 Hz, 6H). HRMS (ESI) calcd for $C_{21}H_{29}N_3O_5Na$: $[M+Na]^+$: 426.2005. Found: 426.1997.

Benzyl ((S)-3-cyclohexyl-1-oxo-1-(((S)-1-oxo-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)propan-2-yl)carbamate (43)—Yield (74%), mp 53–56 °C. ^{1}H NMR (400 MHz, CDCl₃) δ 9.48 (s, 1H), 8.31 (d, J = 5.86 Hz, 1H), 7.73 (br s, 1H), 7.34 (br s, 5H), 5.50 – 5.59 (m, 1H), 5.04 – 5.15 (m, 2H), 4.20 – 4.53 (m, 2H), 3.20 – 3.41 (m, 2H), 2.23 – 2.51 (m, 2H), 1.88 – 2.03 (m, 1H), 1.74 – 1.86 (m, 2H), 1.66 (d, J = 5.86 Hz, 6H), 1.45 – 1.58 (m, 1H), 1.30 – 1.43 (m, 1H), 1.18 (d, J = 16.12 Hz, 3H), 0.79 – 1.03 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{33}N_3O_5Na$: $[M+Na]^+$: 466.2318. Found: 466.2299.

Synthesis of α-Hydroxyphosphonates *8*. General procedure

To a solution of diethylphosphite (135 mg, 1 mmol) in dry dichloromethane (3.5 mL) was added DIEA (129 mg, 1 mmol) followed by a solution of aldehyde *7* (1 mmol) in dichloromethane (5 mL). The reaction mixture was stirred at room temperature for 20 h. The reaction mixture was diluted with dichloromethane (100 mL) and washed with 5% HCl (2 × 20 mL), and brine (30 mL). The organic layer was dried over anhydrous sodium sulfate, filtered and concentrated, leaving a light yellow solid which was purified by flash chromatography to yield compound *8* as a white solid.

2-Chlorobenzyl ((2S)-3-cyclohexyl-1-(((2S)-1-(diethoxyphosphoryl)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-1-oxopropan-2-yl)carbamate (14)—Yield (79%), mp 103–106 °C. ^{1}H NMR (400 MHz, CDCl₃) δ 7.66 – 7.71 (m, 1H), 7.41 (d, J = 3.12 Hz, 1H), 7.33 – 7.37 (m, 1H), 7.28 – 7.32 (m, 1H), 7.21 – 7.26 (m, 3H), 5.49 – 5.57 (m, 1H), 5.15 – 5.24 (m, 2H), 4.06 – 4.21 (m, 4H), 3.97 – 4.04 (m, 1H), 3.86 – 3.95 (m, 1H), 3.41 – 3.52 (m, 1H), 1.77 – 1.86 (m, 2H), 1.56 – 1.71 (m, 6H), 1.50 – 1.56 (m, 1H), 1.25 – 1.37 (m, 7H), 1.23 (s, 2H), 1.17 (dt, J = 2.15, 7.13 Hz, 4H), 0.86

(br s, 4H). HRMS (ESI) calcd for $C_{28}H_{43}ClN_3O_8PNa$: $[M+Na]^+$: 638.2374. Found: 638.2379.

3-Chlorobenzyl ((2S)-3-cyclohexyl-1-(((2S)-1-(diethoxyphosphoryl)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-1-oxopropan-2-yl)carbamate (17)—Yield (81%), mp 58–61 °C. ^{1}H NMR (400 MHz, CDCl3) δ 7.77 – 7.83 (m, 1H), 7.30 – 7.38 (m, 1H), 7.24 – 7.29 (m, 2H), 7.16 – 7.23 (m, 1H), 6.39 – 6.48 (m, 1H), 5.86 – 5.94 (m, 1H), 5.33 – 5.42 (m, 1H), 4.98 – 5.12 (m, 2H), 4.22 – 4.42 (m, 2H), 4.07 – 4.20 (m, 4H), 3.91 – 4.01 (m, 1H), 3.17 – 3.35 (m, 2H), 2.30 – 2.47 (m, 2H), 2.10 – 2.23 (m, 2H), 1.66 (d, J = 11.33 Hz, 6H), 1.41 – 1.54 (m, 2H), 1.28 – 1.36 (m, 6H), 1.05 – 1.22 (m, 3H), 0.79 – 1.00 (m, 3H). HRMS (ESI) calcd for $C_{28}H_{43}ClN_3O_8PNa$: $[M+Na]^+$: 638.2374. Found: 638.2367.

2-Fluorobenzyl ((2S)-3-cyclohexyl-1-(((2S)-1-(diethoxyphosphoryl)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-1-oxopropan-2-yl)carbamate (26)—Yield (73%), mp 98–101 °C. ^{1}H NMR (400 MHz, DMSO-d6) δ 7.80 (br s, 1H), 7.65 (d, J = 8.98 Hz, 1H), 7.49 – 7.57 (m, 1H), 7.33 – 7.48 (m, 2H), 7.14 – 7.25 (m, 2H), 5.90 – 5.97 (m, 1H), 5.71 – 5.84 (m, 1H), 5.07 (d, J = 6.35 Hz, 2H), 3.95 – 4.09 (m, 5H), 3.75 – 3.93 (m, 1H), 2.96 – 3.24 (m, 3H), 2.00 – 2.26 (m, 2H), 1.50 – 1.75 (m, 7H), 1.36 – 1.47 (m, 2H), 1.19 – 1.33 (m, 6H), 1.01 – 1.15 (m, 3H), 0.84 (d, J = 7.91 Hz, 3H). HRMS (ESI) calcd for $C_{28}H_{43}FN_3O_8PNa$: $[M+Na]^+$: 622.2670. Found: 622.2658.

3-Bromobenzyl ((2S)-3-cyclohexyl-1-(((2S)-1-(diethoxyphosphoryl)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propan-2-yl)amino)-1-oxopropan-2-yl)carbamate (30)—Yield (78%), mp 48–52 °C. ^{1}H NMR (400 MHz, CDCl3) δ 8.35 – 8.40 (m, 1H), 7.46 (d, J = 3.32 Hz, 1H), 7.33 – 7.39 (m, 1H), 7.21 (br s, 1H), 7.11 – 7.18 (m, 2H), 5.88 (s, 1H), 5.26 – 5.31 (m, 1H), 5.01 (d, J = 3.32 Hz, 2H), 4.02 – 4.15 (m, 6H), 3.27 (d, J = 6.44 Hz, 3H), 1.88 (br. s., 1H), 1.76 (d, J = 9.37 Hz, 2H), 1.60 (br s, 7H), 1.46 (d, J = 1.95 Hz, 1H), 1.23 – 1.34 (m, 6H), 1.19 (s, 1H), 0.98 – 1.16 (m, 3H), 0.75 – 0.96 (m, 3H). HRMS (ESI) calcd for $C_{28}H_{43}BrN_3O_8PNa$: $[M+Na]^+$: 682.1869. Found: 682.1896.

Synthesis of bisulfite adducts 9. General procedure—To a solution of aldehyde **7** (5 mmol) in dry ethyl acetate (20 mL) was added absolute ethanol (12 mL) with stirring, followed by a solution of sodium bisulfite (540 mg; 5 mmol) in water (5 mL). The reaction mixture was stirred for 3 h at 50 °C. The reaction mixture was allowed to cool to room temperature and then vacuum filtered. The solid was thoroughly washed with absolute ethanol and the filtrate was dried over anhydrous sodium sulfate, filtered, and concentrated to yield yellowish oil. The oily product was treated with ethyl ether (2×50 mL) to form white solid. The white solid was stirred with ethyl ether (30 mL) and ethyl acetate (15 mL) for 5 minutes. Careful removal of the solvent using a pipette left compound **9** as a white solid.

Sodium (2S)-2-((S)-2-((((2-chlorobenzyl)oxy)carbonyl)amino)-3-cyclohexylpropana mido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (15)—Yield (60%), mp 128–132 °C. ^{1}H NMR (400 MHz, CDCl3) δ 7.51 – 7.53 (m, 1H), 7.40 – 7.45 (m, 1H), 7.35 – 7.39 (m, 1H), 7.26 (s, 3H), 6.99 – 7.00 (m, 1H), 5.18 –

5.29 (m, 2H), 3.68 – 3.78 (m, 2H), 3.48 (d, $J = 7.03$ Hz, 3H), 3.29 – 3.40 (m, 1H), 1.92 – 1.98 (m, 1H), 1.80 – 1.91 (m, 2H), 1.71 (br s, 7H), 1.24 – 1.28 (m, 5H), 0.79 – 0.96 (m, 2H), 0.68 – 0.76 (m, 1H). HRMS (ESI) calcd for $C_{24}H_{33}ClN_3O_8S$: [M-]: 558.1677. Found: 558.1648.

Sodium (2S)-2-((S)-2-((((3-chlorobenzyl)oxy)carbonyl)amino)-3-cyclohexylpropana mido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (18)—Yield (67%), mp 124–126 °C. ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.66 (d, $J = 8.98$ Hz, 1H), 7.58 (dd, $J = 9.18, 11.52$ Hz, 1H), 7.49 (d, $J = 8.20$ Hz, 1H), 7.42 (d, $J = 4.30$ Hz, 1H), 7.40 (d, $J = 1.95$ Hz, 1H), 7.30 – 7.39 (m, 2H), 5.41 (d, $J = 5.86$ Hz, 1H), 5.25 (d, $J = 5.86$ Hz, 1H), 5.03 (s, 2H), 4.17 – 4.26 (m, 1H), 3.89 – 3.99 (m, 2H), 3.81 – 3.87 (m, 1H), 3.10 (t, $J = 9.96$ Hz, 1H), 2.95 – 3.05 (m, 1H), 2.01 – 2.25 (m, 3H), 1.51 – 1.78 (m, 6H), 1.37 – 1.48 (m, 1H), 1.28 (br s, 1H), 1.15 – 1.21 (m, 1H), 1.08 – 1.14 (m, 2H), 0.74 – 0.94 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{33}ClN_3O_8S$: [M-]: 558.1677. Found: 558.1656.

Sodium(2S)-2-((S)-2-((((3-bromobenzyl)oxy)carbonyl)amino)-3-cyclohexylpropana mido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (31)—Yield (86%), mp 132–135 °C. ^{1}H NMR (400 MHz, DMSO-d_6) 0δ 7.68 (d, $J = 8.98$ Hz, 1H), 7.60 (d, $J = 9.37$ Hz, 1H), 7.53 – 7.58 (m, 1H), 7.50 (d, $J = 7.81$ Hz, 1H), 7.45 (d, $J = 4.30$ Hz, 1H), 7.29 – 7.41 (m, 2H), 5.54 (d, $J = 5.47$ Hz, 1H), 5.36 (d, $J = 5.08$ Hz, 1H), 5.03 (s, 2H), 4.18 – 4.27 (m, 1H), 3.91 – 4.10 (m, 1H), 2.96 – 3.15 (m, 2H), 2.05 – 2.23 (m, 2H), 1.89 – 2.05 (m, 1H), 1.49 – 1.75 (m, 6H), 1.38 – 1.47 (m, 1H), 1.23 – 1.35 (m, 1H), 1.02 – 1.21 (m, 5H), 0.75 – 0.95 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{33}BrN_3O_8S$: [M-]: 602.1172. Found: 602.1168.

Sodium (2S)-2-((S)-3-cyclohexyl-2-((((3-iodobenzyl)oxy)carbonyl)amino)propane mido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (34)—Yield (78%), mp 135–137 °C. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.48 – 8.52 (m, 1H), 7.73 (br s, 1H), 7.66 (d, $J = 7.62$ Hz, 1H), 7.51 – 7.59 (m, 1H), 7.45 (d, $J = 3.52$ Hz, 1H), 7.36 (br s, 1H), 7.13 – 7.21 (m, 1H), 5.45 – 5.52 (m, 1H), 5.28 – 5.35 (m, 1H), 4.92 – 5.14 (m, 2H), 4.13 – 4.29 (m, 1H), 3.88 – 4.08 (m, 1H), 2.97 – 3.19 (m, 2H), 1.91 (s, 3H), 1.37 – 1.78 (m, 7H), 1.02 – 1.24 (m, 5H), 0.73 – 0.97 (m, 3H). HRMS (ESI) calcd for $C_{24}H_{33}IN_3O_8S$: [M-]: 650.1033. Found: 650.1019.

Sodium (2S)-2-((S)-2-(((benzyloxy)carbonyl)amino)-4-methylpentanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-sulfonate (41)—Yield (73%), mp 109–113 °C. ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.57 – 7.62 (m, 1H), 7.47 – 7.52 (m, 1H), 7.26 – 7.40 (m, 5H), 5.47 – 5.51 (m, 1H), 5.31 – 5.36 (m, 1H), 4.96 – 5.05 (m, 3H), 4.19 – 4.28 (m, 1H), 3.83 – 4.06 (m, 2H), 3.07 – 3.16 (m, 1H), 2.97 – 3.06 (m, 1H), 2.04 – 2.23 (m, 2H), 1.71 – 1.81 (m, 1H), 1.52 – 1.68 (m, 2H), 1.39 – 1.50 (m, 2H), 0.85 (ddd, $J = 3.03, 6.52, 9.40$ Hz, 6H). HRMS (ESI) calcd for $C_{21}H_{30}N_3O_8S$: [M-]: 484.1754. Found: 484.1754.

Sodium (2S)-2-((S)-2-(((benzyloxy)carbonyl)amino)-3-cyclohexylpropanamido)-1-hydroxy-3-((S)-2-oxopyrrolidin-3-yl)propane-1-

sulfonate (44)—Yield (78%), mp 131–133 °C. ^{1}H NMR (400 MHz, DMSO-d[6]) δ 7.62 – 7.69 (m, 1H), 7.43 – 7.51 (m, 1H), 7.27 – 7.40 (m, 5H), 5.82 – 5.86 (m, 1H), 5.52 (d, J = 6.25 Hz, 1H), 5.36 (d, J = 5.96 Hz, 1H), 5.02 (s, 2H), 3.83 – 4.08 (m, 2H), 2.97 – 3.20 (m, 3H), 1.86 – 2.25 (m, 3H), 1.51 – 1.82 (m, 6H), 1.37 – 1.48 (m, 2H), 1.29 (br s, 1H), 1.02 – 1.22 (m, 3H), 0.75 – 0.94 (m, 2H). HRMS (ESI) calcd for $C_{24}H_{34}N_3O_8S$: [M-]: 524.2067. Found: 524.2045.

Synthesis of compound 10. General procedure

A solution of aldehyde **7** (5 mmol) in ethyl acetate (30 mL) kept at 0 °C was treated with acetic acid (0.34 g; 5.75 mmol) followed by cyclopropyl isocyanide (0.37 g; 5.5 mmol), and the reaction mixture was stirred at room temperature for 18 h. The solution was concentrated in vacuo and the residue was dissolved in methanol (30 mL) and treated with a solution of K_2CO_3 (1.72 g; 12.5 mmol) in water (25 mL). The reaction mixture was stirred at room temperature for 2 h. Methanol was evaporated off and the aqueous layer was extracted with ethyl acetate (3 × 100 mL). The combined organic layers were washed with 5% HCl (2 × 50 mL) and brine (75 mL). The organic layer was dried over anhydrous sodium sulfate and the solvent was removed to yield compound **10** as a white solid. This was used in the next step without further purification.

Synthesis of α-ketoamides 11. General procedure

To a solution of compound **10** (5 mmol) in anhydrous dichloromethane (50 mL) cooled to 0 °C and kept under a nitrogen atmosphere was added Dess-Martin periodinane reagent (3.18 g, 7.5 mmol, 1.5 eq) with stirring. The ice bath was removed and the reaction mixture was stirred at room temperature for 3 h. The reaction was monitored by TLC until the starting material disappeared. A solution of 10% aqueous sodium thiosulfate (20 mL) was added and the solution was stirred for 15 minutes. The solution was poured into a separatory funnel and the aqueous layer was removed. The organic layer was washed with 10 % aqueous sodium thiosulfate (20 mL), followed by saturated aqueous sodium bicarbonate (2 × 20 mL), water (2 × 20 mL) and brine (20 mL). The organic layer was dried over anhydrous sodium sulfate, filtered and concentrated leaving a yellow solid which was purified by flash chromatography (silica gel/methylene chloride/ethyl acetate/methanol) to yield **11** as a white solid.

3-Chlorobenzyl ((S)-3-cyclohexyl-1-(((S)-4-(cyclopropylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-1-oxopropan-2-yl)carbamate (19)—
Yield (88%), mp 135–138 °C. ^{1}H NMR (400 MHz, CDCl[3]) δ 8.42 (br s, 1H), 7.36 (br s, 1H), 7.27 – 7.29 (m, 2H), 7.19 – 7.24 (m, 2H), 6.97 (br s, 1H), 5.97 (br s, 1H), 5.08 (d, J = 4.30 Hz, 2H), 4.24 – 4.42 (m, 2H), 3.27 – 3.40 (m, 2H), 2.76 (dd, J = 3.52, 7.42 Hz, 1H), 2.35 – 2.60 (m, 2H), 2.07 (dd, J = 3.32, 5.27 Hz, 1H), 1.89 – 1.99 (m, 2H), 1.77 – 1.87 (m, 2H), 1.59 – 1.75 (m, 4H), 1.43 – 1.56 (m, 1H), 1.38 (br s, 1H), 1.09 – 1.28 (m, 3H), 0.86 – 1.04 (m, 2H), 0.81 – 0.85 (m, 2H), 0.56 – 0.66 (m, 2H). HRMS (ESI) calcd for $C_{28}H_{37}ClN_4O_6Na$: [M+Na]+: 583.2299. Found: 583.2293.

3-Chlorobenzyl ((S)-1-(((S)-4-(cyclopropylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (23)—Yield (72%), mp 66–68 °C. ^{1}H NMR (400 MHz, CDCl[3]) δ 8.40 (d, J = 5.47 Hz, 1H),

7.24 (br s, 1H), 7.16 – 7.22 (m, 1H), 7.11 (br s, 2H), 6.75 (br s, 1H), 5.63 – 5.75 (m, 1H), 5.10 – 5.21 (m, 1H), 4.93 – 5.05 (m, 2H), 4.22 – 4.36 (m, 2H), 3.15 – 3.30 (m, 2H), 2.69 (d, J = 3.12 Hz, 1H), 2.39 – 2.51 (m, 1H), 2.31 (d, J = 3.91 Hz, 1H), 1.80 – 2.01 (m, 5H), 1.50 – 1.72 (m, 3H), 1.38 – 1.49 (m, 2H), 0.86 (d, J = 4.69 Hz, 2H), 0.74 (d, J = 7.42 Hz, 2H), 0.54 (br s, 2H). HRMS (ESI) calcd for $C_{25}H_{33}ClN_4O_6Na$: [M+Na]$^+$: 543.1986. Found: 543.1989.

3-Fluorobenzyl ((S)-3-cyclohexyl-1-(((S)-4-(cyclopropylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-1-oxopropan-2-yl)carbamate (28)—
Yield (74%), mp 95–98 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.41 – 8.49 (m, 1H), 7.28 – 7.36 (m, 2H), 7.25 (d, J = 2.34 Hz, 1H), 7.10 (t, J = 7.42 Hz, 2H), 6.90 – 7.02 (m, 1H), 5.83 (br s, 1H), 5.04 – 5.13 (m, 2H), 4.25 – 4.39 (m, 2H), 3.36 (dd, J = 4.49, 8.79 Hz, 2H), 2.76 (dd, J = 3.52, 7.42 Hz, 1H), 2.51 – 2.61 (m, 1H), 2.39 – 2.50 (m, 1H), 2.06 – 2.14 (m, 1H), 1.94 (d, J = 10.94 Hz, 2H), 1.80 (br s, 1H), 1.58 – 1.74 (m, 6H), 1.48 (br s, 1H), 1.31 – 1.42 (m, 1H), 1.07 – 1.29 (m, 3H), 0.79 – 1.03 (m, 3H), 0.61 (ddd, J = 1.56, 4.00, 5.76 Hz, 2H). HRMS (ESI) calcd for $C_{28}H_{37}FN_4O_6Na$: [M+Na]$^+$: 567.2595. Found: 567.2592.

3-Bromobenzyl ((S)-3-cyclohexyl-1-(((S)-4-(cyclopropylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-1-oxopropan-2-yl)carbamate (32)—
Yield (70%), mp 124–127 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.39 – 8.45 (m, 1H), 7.52 (br s, 1H), 7.43 (d, J = 7.62 Hz, 1H), 7.27 (br s, 1H), 7.22 (d, J = 7.62 Hz, 2H), 6.90 – 6.95 (m, 1H), 5.71 – 5.75 (m, 1H), 5.18 – 5.27 (m, 1H), 5.07 (s, 2H), 4.28 – 4.35 (m, 1H), 3.35 (dd, J = 4.10, 8.98 Hz, 3H), 2.72 – 2.81 (m, 2H), 2.06 – 2.15 (m, 2H), 1.86 – 2.01 (m, 2H), 1.69 (d, J = 10.94 Hz, 5H), 1.58 (s, 3H), 1.25 (s, 2H), 0.87 – 1.02 (m, 2H), 0.84 (d, J = 7.42 Hz, 2H), 0.55 – 0.67 (m, 2H). HRMS (ESI) calcd for $C_{28}H_{37}BrN_4O_6Na$: [M+Na]$^+$: 627.1794. Found: 627.1791.

Benzyl ((S)-1-(((S)-4-(cyclopropylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (42)—Yield (78%), mp 75–78 °C. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.43 (d, J = 5.68 Hz, 1H), 7.33 (s, 5H), 7.08 (d, J = 3.48 Hz, 1H), 6.57 (s, 1H), 5.53 (d, J = 8.79 Hz, 1H), 5.19 – 5.26 (m, 1H), 5.08 (s, 2H), 4.40 (dt, J = 5.49, 8.88 Hz, 1H), 3.22 – 3.36 (m, 2H), 2.71 – 2.80 (m, 1H), 2.48 – 2.58 (m, 1H), 2.35 – 2.46 (m, 1H), 1.88 – 2.10 (m, 3H), 1.59 – 1.79 (m, 2H), 1.45 – 1.55 (m, 1H), 0.95 (d, J = 6.41 Hz, 6H), 0.79 – 0.86 (m, 2H), 0.57 – 0.66 (m, 2H). HRMS (ESI) calcd for $C_{25}H_{34}N_4O_6Na$: [M+Na]$^+$: 509.2376. Found: 509.2364.

Enzyme assays and inhibition studies against NV with 3CLpro and cell-based replicon system

These studies were carried out as described previously.[13c]

X-ray Crystallographic studies. Crystallization and Data Collection

Purified NV 3CLpro in 100 mM NaCl, 50 mM PBS pH 7.2, 1 mM DTT at a concentration of 10 mg/mL was used for preparation of the norovirus 3CL protease-compound *17* complex. A stock solution of 100 mM compound *17* or *44* was prepared in DMSO and each NV 3CLpro:inhibitor complex was prepared by mixing 12 μL of *17* or *44* (3 mM) with 388 μL (0.49 mM) of NV 3CLpro and incubating on ice for 1 h. The mixture was loaded onto a

Superdex 75 10/300 GL column equilibrated with 100 mM NaCl, 20 mM Tris pH 8.0 and the sample was concentrated to 8.0 mg/mL in a Vivaspin-20 (Vivaproducts, Inc.) concentrator for crystallization screening. All crystallization experiments were conducted Compact Jr. (Rigaku Reagents) sitting drop vapor diffusion plates at 20 °C using equal volumes of protein and the crystallization solution were equilibrated against 75 μL of the latter. Crystals of NV 3CLpro-ligand *17* that displayed a needle morphology were obtained overnight from the Proplex HT screen (Molecular Dimensions) condition D4 (20% (w/v) PEG 5000 MME, 100 mM Tris pH 7.5, 200 mM ammonium sulfate). Samples were transferred to a fresh drop composed of 80% crystallization solution and 20% PEG glycerol and stored in liquid nitrogen. Two crystal forms of the NV 3CLpro:*44* complex that displayed a prismatic morphology were observed in 2 days from the Wizard 3 & 4 screen (Emerald Biosystems). The primitive hexagonal (NV 3CLpro:*44*-h) form was obtained from condition A1 (20% w/v PEG3350, 200 mM ammonium citrate) and a primitive orthorhombic (NV 3CLpro:*44*-o) from A10 (20% w/v PEG3350, 200 mM sodium thiocyanate). Samples were transferred to a fresh drop composed of 80% crystallization solution and 20% PEG 400 and stored in liquid nitrogen prior to data collection. X-ray diffraction data were collected at the Advanced Photon Source beamline 17-ID using a Dectris Pilatus 6M pixel array detector.

Structure Solution and Refinement

Intensities were integrated using XDS[26–27] and the Laue class analysis and data scaling were performed with Aimless[28] which suggested that the highest probability Laue class was *6/mmm* and possible space groups was $P6_122$ or $P6_522$ for NV 3CLpro, NV 3CLpro:*17*, and NV 3CLpro:*44*-h. The Matthew's coefficient[29] suggested that there was a single molecule in the asymmetric unit (Vm=2.8 Å^3/Da, % solvent=56%). Structure solution was conducted by molecular replacement with Phaser[30] using a previously determined structure of inhibitor bound NV 3CLpro (PDB: 3UR9)[13c] as the search model and all space groups with 622 point symmetry were tested. The top solution was obtained in the space group $P6_522$ which was used from this point forward. NV 3CLpro:*44*-o was isomorphous with PDB 3UR9. Structure refinement using and manual model building were conducted with Phenix[31] and Coot[32] respectively. Disordered side chains were truncated to the point for which electron density could be observed. Structure validation was conducted with Molprobity[33] and figures were prepared using the CCP4MG package[34].

Study with MNV.[35]

RAW264.7 cells, mouse macrophage cell line, were maintained in Dulbecco's Minimum Essential Medium (DMEM) containing antibiotics of chlortetracycline (25 μg/ml), penicillin (250 U/ml), and streptomycin (250 μg/ml) in the presence of 2~5% fetal bovine serum. MNV-1 was obtained from H. Virgin at Washington University (St. Louis, MO). Antiviral activity of each compound against MNV-1 in cell culture was determined with RAW264.7 cells. Each compound was dissolved in DMSO to make a 10 mM stock solution. The amount of DMSO in cell culture did not exceed 0.5%. The stock solution was serially diluted with cell culture media prior to addition to confluent monolayer of RAW264.7 cells in 24-well plates. Cells were immediately infected with virus at a multiplicity of infection (MOI) of 0.05 ~0.1. After further incubation of the cells at 37°C for 48 h, cells were frozen

and thawed and virus titers were determined by the 50% tissue culture infectious dose ($TCID_{50}$) method. The 50% effective concentration (EC_{50}) values were determined by GraphPad Prism (GraphPad Software, San Diego, CA) by nonlinear regression analyses of dose-response curves of the inhibition of virus replication against log inhibitor concentration (variable slope).

Animal experiments.[36]

The animal study was performed in accordance with a protocol approved by the Institutional Animal Care and Use Committee (IACUC) at Kansas State University. BALB/c mice were purchased from Charles River Lab (Wilmington, MA). Seven to eight week-old female BALB/c mice were inoculated intraperitoneally with MNV-1 at 2×10^6 $TCID_{50}$/mouse. Mice were peritoneally given 50 μl of drug vehicle (10% EtOH and 90% PEG400) or compound *16* (20 mg/kg/day) divided into two doses per day. Compound administration started 4 h prior to virus infection and continued daily until mice were euthanized. At 3 days post infection (dpi), mice were sacrificed and the small and large intestinal tracts were collected. The intestines were placed in PBS and the intestinal contents were gently removed by squeezing. The tissues were homogenized in PBS (5 or 3 mL for small and large intestines, respectively) and frozen at −70C until ready to use. Virus titers of the homogenized supernatant were determined by the $TCID_{50}$ method and statistical analysis was performed on log-transformed $TCID_{50}$ values using a two-tailed Student t-test ($p<0.05$). Fold changes in the geometric mean liver virus titers in each group were calculated by dividing virus titers in control group by the treated group.

Acknowledgements

References

1. Arias M, Emmott E, Vashist S, Goodfellow I. Progress Towards the Prevention and Treatment of Norovirus Infections. Future Microbiol. 2014; 8:1475–1487. [PubMed: 24199805]

2. Iturriza-Gomara M, Lopman B. Norovirus in Healthcare Settings. Curr. Opin. Infect. Dis. 2014; 27:437–443. [PubMed: 25101555]

3. Bok K, Green KY. Norovirus Gastroenteritis in Immunocompromised Patients. N. Engl. J. Med. 2012; 367:2126–2132. [PubMed: 23190223]

4. Thornton SA, Sherman SS, Farkas T, Zhong W, Torres P, Jiang X. Gastroenteritis in United States Marines During Operation Iraqi Freedom. Clin. Infect. Dis. 2004; 40:519–525. [PubMed: 15712073]

5. Tung G, Macinga D, Arbogast J, Jaykus LA. Efficacy of Commonly Used Disinfectants for Inactivation of Human Noroviruses and Their Surrogates. J. Food Prot. 2013; 76:1210–1217. [PubMed: 23834796]

6. Barclay L, Park GW, Vega E, Hall A, Parashar U, Vinje J, Lopman B. Infection Control for Norovirus. Clin. Microbiol. Infect. 2014 May 11.

7. Rocha-Pereira J, Neyts J, Jochmans D. Norovirus: Targets and Tools in Antiviral Drug Discovery. Biochem. Pharmacol. 2014; 91:1–11. [PubMed: 24893351]

8. Kaufman SS, Green KY, Korba BE. Treatment of Norovirus Infections: Moving Antivirals from the Bench to the Bedside. Antiviral Res. 2014; 105:80–91. [PubMed: 24583027]

9. Ramani S, Atmar RL, Estes MK. Epidemiology of Human Noroviruses and Updates on Vaccine Development. Curr. Opin. Gastroenterol. 2014; 30:25–33. [PubMed: 24232370]

10. Green, KY. Caliciviridae: The Noroviruses in *Green's Virology*. Knipe, DM.; Howley, PM., editors. Vol. 1. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 949-979.

11. Blakeney SJ, Cahill A, Reilly PA. Processing of Norwalk Virus Nonstructural Proteins by a 3C–like Cysteine Proteinase. Virology. 2003; 308:216–224. [PubMed: 12706072]

12. Thorne LG, Goodfellow IG. Norovirus Gene Expression and Replication. J. Gen. Virol. 2014; 95:278–291. [PubMed: 24243731]

13. (a) Tiew K-C, He G, Aravapalli S, Mandadapu SR, Gunnam MR, Alliston KR Lushington GH, Kim Y Chang K-O, Groutas WC. Design, Synthesis, and Evaluation of Inhibitors of Norwalk Virus 3C Protease. Bioorg. Med. Chem. Lett. 2011; 21:5315–5319. [PubMed: 21802286] (b) Mandadapu SR, Weerawarna PM, Gunnam MR, Alliston KR, Lushington GH, Kim Y, Chang K-O, Groutas WC. Potent Inhibition of Norovirus 3CL Protease by Peptidyl α-Ketoamides and α-Ketoheterocycles. Bioorg. Med. Chem. Lett. 2012; 22:4820–4826. [PubMed: 22698498] (c) Kim Y, Lovell S, Tiew K-C, Mandadapu SR, Alliston KR, Battaille KP, Groutas WC, Chang K-O. Broad Spectrum Antivirals Against 3C or 3C–Like Proteases of Picornaviruses, Noroviruses and Coronaviruses. J. Virol. 2012; 6:1754–11762.(d) Mandadapu SR, Gunnam MR, Tiew K-C, Uy RAZ, Prior AM, Alliston KR, Hua DH, Kim Y, Chang K-O, Groutas W. C Inhibition of Norovirus 3CL Protease by Bisulfite Adducts of Transition State Inhibitors. Bioorg. Med. Chem. Lett. 2013; 23:62–65. [PubMed: 23218713] (e) Prior AM, Kim Y, Weerasekara S, Moroze M, Alliston KR, Uy RAZ, Groutas WC, Chang K-O, Hua DH. Design, Synthesis, and Bioevaluation of Viral 3C and 3C–like Protease Inhibitors. Bioorg. Med. Chem. Lett. 2013; 23:6317–6320. [PubMed: 24125888] (f) Mandadapu SR, Gunnam MR, Kankanamalage ACG, Uy RAZ, Alliston KR, Lushington GH, Kim Y, Chang K-O, Groutas WC. Potent Inhibition of Norovirus by Dipeptidyl α-Hydroxyphosphonate Transition State Mimics. Bioorg. Med. Chem. Lett. 2013; 23:5941–5944. [PubMed: 24054123] (g) Mandadapu SR, Weerawarna PM, Prior AM, Uy RAZ, Aravapalli S, Alliston KR, Lushington GH, Kim Y, Hua DH, Chang K-O, Groutas WC. Macrocyclic Inhibitors of 3C and 3C–like Proteases of Picornavirus, Norovirus and Coronavirus. Bioorg. Med. Chem. Lett. 2013; 23:3709–3712. [PubMed: 23727045] (h) Takahashi D, Kim Y, Lovell S, Prakash O, Groutas WC, Chang K-O. Structural and Inhibitor Studies of Norovirus 3C–Like Proteases. Virus Res. 2013; 178:437–444. [PubMed: 24055466]

14. Deng L, Muhaxhiri Z, Estes MK, Palzkill T, Prasad Venkataram BV. Song, Y. Synthesis, Activity and Structure-Activity Relationship of Noroviral Protease Inhibitors. MedChemComm. 2013; 4:1354–1359.

15. Hussey RJ, Coates L, Gill RS, Erskine PT, Coker S-F, Mitchell E, Cooper JB, Wood S, Broadbridge R, Clarke IN, Lambden PR, Shoolingin-Jordan PM. A Structural Study of Norovirus 3C Protease Specificity: Binding of a Designed Active Site-Directed Peptide Inhibitor. Biochemistry. 2011; 50:240–249. [PubMed: 21128685]

16. Nomenclature used is that of Schecter, I.; Berger, A. Biochem. Biophys. Res. Comm. 1967; 27:157. [PubMed: 6035483]

17. Dou D, Tie K-C.; He G, Mandadapu SR.; Aravapalli S, Alliston KR.; Kim Y, Chang K-O.; Groutas, WC. Potent Inhibition of Norwalk Virus by Cyclic Sulfamide Derivatives Bioorg. Med. Chem. 2011; 19:5975–5983.

18. Dou D, Mandadapu SR, Alliston KR, Kim Y, Chang K-O, Groutas WC. Cyclosulfamide-based Derivatives as Inhibitors of Noroviruses. Eur. J. Med. Chem. 2012; 47:59–64. [PubMed: 22063754]

19. Dou D, Tiew K-C, Mandadapu SR, Gunnam MR, Alliston KR, Kim Y, Chang K-O, Groutas WC. Potent Norovirus Inhibitors Based on the Acyclic Sulfamide Scaffold. Bioorg. Med. Chem. 2012; 20:2111–2118. [PubMed: 22356738]

20. (a) Dou D, Mandadapu SR, Alliston KR, Kim Y, Chang K-O, Groutas WC. Design and Synthesis of Potent Inhibitors of Noroviruses by Scaffold Hopping. Bioorg. Med. Chem. 2011; 19:5749–

5755. [PubMed: 21893416] (b) Dou D, He G, Mandadapu SR, Aravapalli S, Kim Y, Chang K-O, Groutas WC. Inhibition of Noroviruses by Piperazine Derivatives. Bioorg. Med. Chem. Lett. 2012; 22:377–379. [PubMed: 22119464]

21. Webber SE, Okano K, Little TL, Reich SH, Xin Y, Fuhrman SA, Matthews DA, Love RA, Hendrickson TF, Patick AK, Meador JW, Ferre RA, Brown EL, Ford CE, Binford SL, Worland ST. Tripeptide Aldehyde Inhibitors of Human Rhinovirus 3C Protease: Design, Synthesis, Biological Evaluation, and Cocrystal Structure Solution of P1 Glutamine Isosteric Replacements. J. Med. Chem. 1998; 41:2786–2805. [PubMed: 9667969]

22. (a) Lu Y, Liu Y, Xu Z, Li H, Liu H, Zhu W. Halogen Bonding for Rational Drug Design and New Drug Discovery. J. Med. Chem. 2012; 7:375–383.(b) Scholfield MR, Vander Zanden CM, Carter M, Shing Ho P. Halogen Bonding (X-bonding): A Biological Perspective. Protein Sci. 2013; 22:139–152. [PubMed: 23225628] (c) Zhou P, Huang J, Tian F. Specific Noncovalent Interactions at Protein-Ligand Interface: Implications for Rational Drug Design. Curr. Med. Chem. 2012; 19:226–238. [PubMed: 22320300]

23. Patel DV, Rielly-Gauvin K, Ryono DE, Free CA, Rogers WL, Smith SA, DeForrest JM, Oehl RS, Petrillo EW. α-Hydroxy Phosphinyl-Based Inhibitors of Human Renin. J. Med. Chem. 1995; 38:4557–4569. [PubMed: 7473584]

24. Kjell DP, Slattery BJ, Semo MJ. A Novel Nonaqueous Method for Regeneration of Aldehydes from Bisulfite Adducts. J. Org. Chem. 1999; 64:5722–5724. [PubMed: 11674650]

25. De Fenza A, Hiene A, Koert U, Klebe G. *Trypsin*. Understanding the Binding Selectivity Toward Trypsin and Factor Xa. ChemMedChem. 2007; 2:297–308. [PubMed: 17191291] Sidhu PS, Liang A, Mehta AY, Abdel Aziz MH, Zhou Q, Desai UR. *Thrombin*. Rational Design of potent, Small, Synthetic Allosteric Inhibitors of Thrombin. J. Med. Chem. 2011; 54:5522–5531. [PubMed: 21714536] Taggart CC, Lowe GJ, Greene CM, Mulgrew AT, O'Neill SJ, Levine RL, McElvaney NGCathepsin B, L SCleave. *Cathepsins:* and Inactivate Secretory Leukoprotease Inhibitor. J. Biol. Chem. 2001; 276:33345–33352. [PubMed: 11435427] Groutas WC, Kuang R, Venkataraman R, Epp JB, Ruan S, Prakash O. *Human neutrophil elastase*. Structure-Based Design of a General Class of Mechanism-Based Inhibitors of the Serine Proteinases Employing a Novel Amino Acid-Derived Heterocyclic Scaffold. Biochemistry. 1997; 36:4739–4750. [PubMed: 9125494] Tanaka Y, Grapsas I, Dakoji S, Cho YJ, Mobashery S. *Carboxypeptidase A:* Conscripting the Active Site Zinc Ion in Carboxypeptidase A in Inactivation Chemistry by a New Type of Irreversible Enzyme Inactivator. J. Am. Chem. Soc. 1994; 116:7475–7480.

26. Kabsch W. Automatic indexing of rotation diffraction patterns. J. Appl. Crystallogr. 1988; 21:67–72.

27. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. [PubMed: 20124692]

28. Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr. D. Biol. Crystallogr. 2011; 67:282–292. [PubMed: 21460446]

29. Matthews BW. Solvent content of protein crystals. J. Mol. Biol. 1968; 33:491–497. [PubMed: 5700707]

30. McCoy AJ, et al. Phaser crystallographic software. J. Appl. Cryst. 2007; 40:658–674. [PubMed: 19461840]

31. Adams PD, et al. PHENIX: a Comprehensive Python-based System for Macromolecular Structure Solution. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:213–221. [PubMed: 20124702]

32. Emsley P, et al. Features and Development of Coot. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:486–501. [PubMed: 20383002]

33. Chen VB, et al. MolProbity: All-atom Structure Validation for Macromolecular Crystallography. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:12–21. [PubMed: 20057044]

34. Potterton L, et al. Developments in the CCP4 Molecular Graphics Project. Acta Crystallogr. D Biol. Crystallogr. 2004; 60:2288–94. [PubMed: 15572783]

35. Wobus CE, Karst SM, Krug A, Chang KO, Sosnovtsev SV, Belliott G, Mackenzie JM, Green KY, Virgin HW. Replication of a Norovirus in Cell Culture Reveals a tropism for Dendritic Cells and Macrophages. PLoS Biol. 2(12):e432. [PubMed: 15562321]

36. Perry JW, Ahmed M, Chang KO, Donato NJ, Showalter HD, Wobus CE. A Small Molecule Deubiquitinase Inhibitors has Anti-Viral Effects Through Induction of the Unfolded Protein Response. PLoS Pathog. 8(7):e1002783. [PubMed: 22792064]

37. Evans P. Scaling and Assessment of Data Quality. Acta Crystallogr. D Biol Crystallogr. 2006; 62:72–82. [PubMed: 16369096]

38. Diederichs K, Karplus PA. Improved R-factors for Diffraction Data Analysis in Macromolecular Crystallography. Nat Struct Biol. 1997; 4:269–275. [PubMed: 9095194]

39. Weiss MS. Global Indicators of X-ray Data Quality. J. Appl. Crystallogr. 2001; 34:130–135.

40. Karplus PA, K. Diederichs K. Linking Crystallographic Model and Data Quality. Science. 2012; 336:1030–1033. [PubMed: 22628654]

41. Evans P. Biochemistry. Resolving Some Old Problems in Protein Crystallography. Science. 2012; 336:986–987. [PubMed: 22628641]
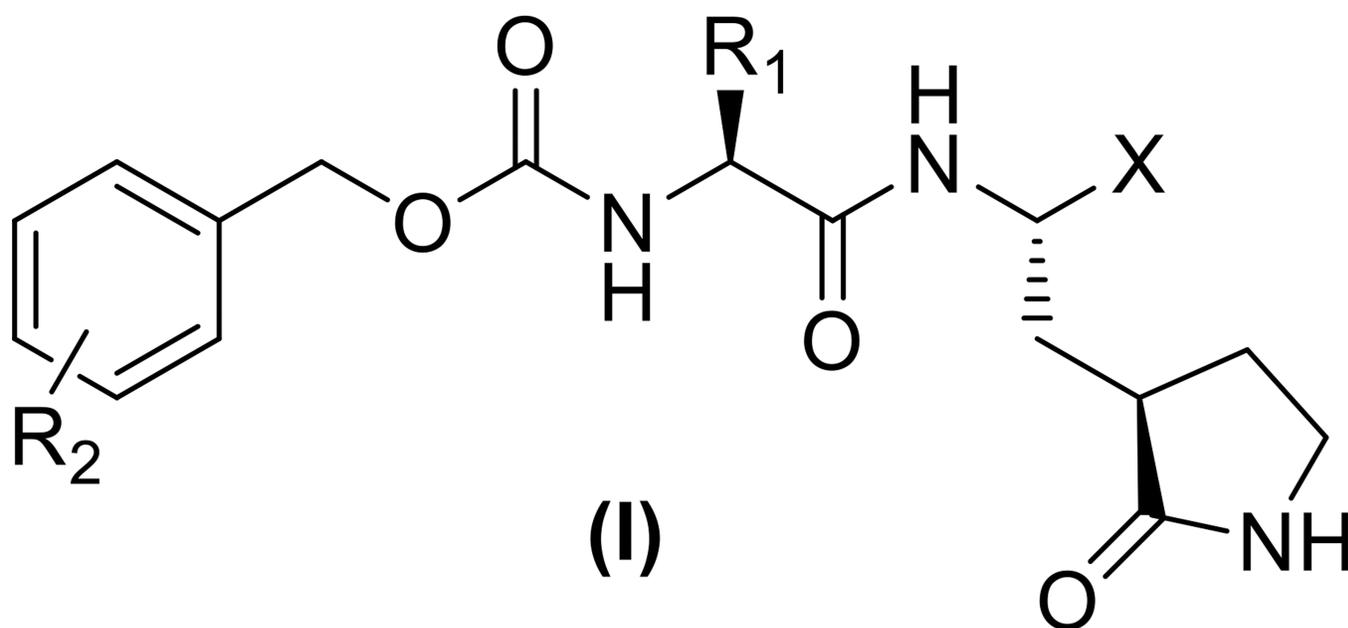
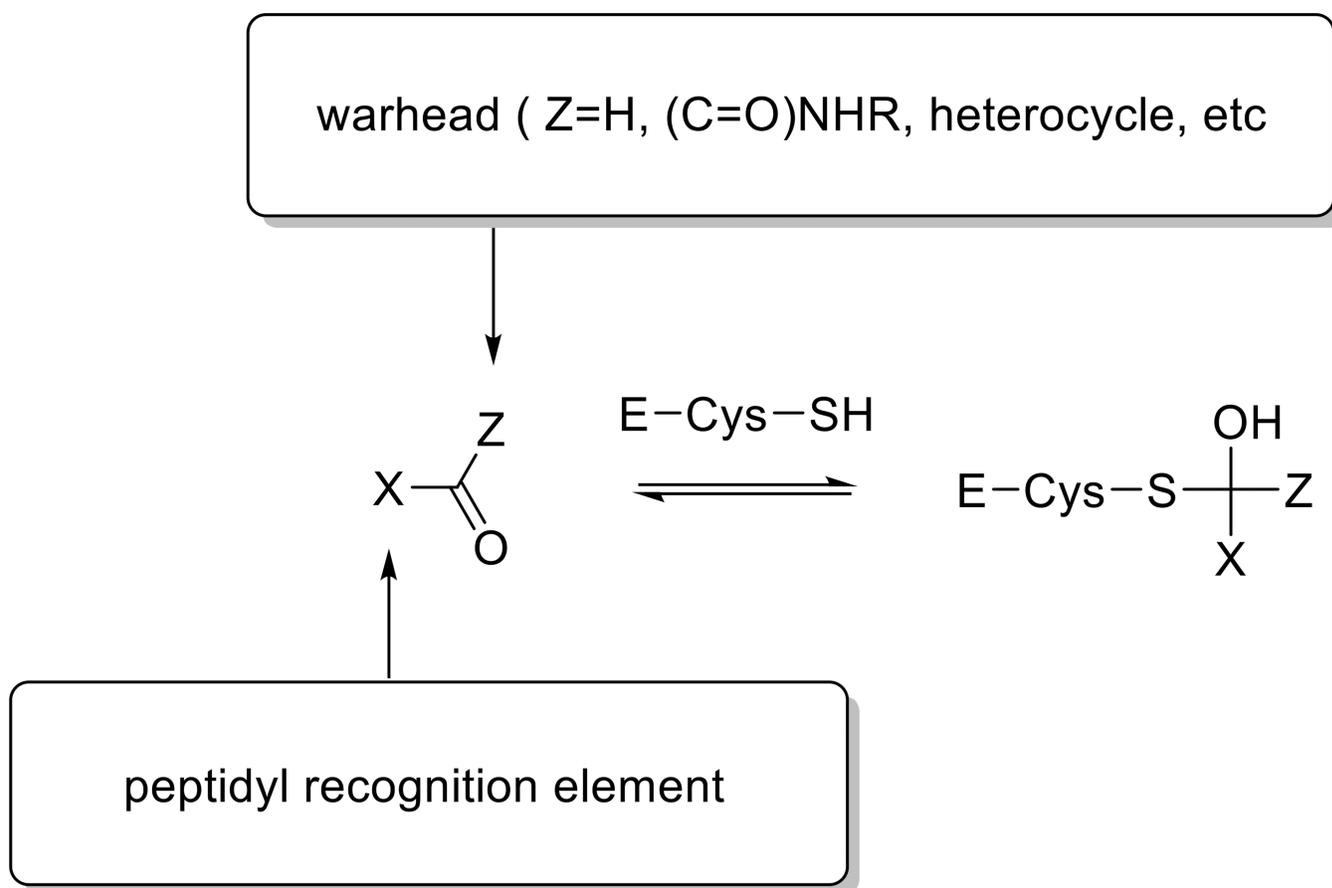**Figure 1.**
General structure of dipeptidyl inhibitor (I)

Figure 2.
Interaction of a cysteine protease (E-Cys-SH) with a peptidyl transition state inhibitor.
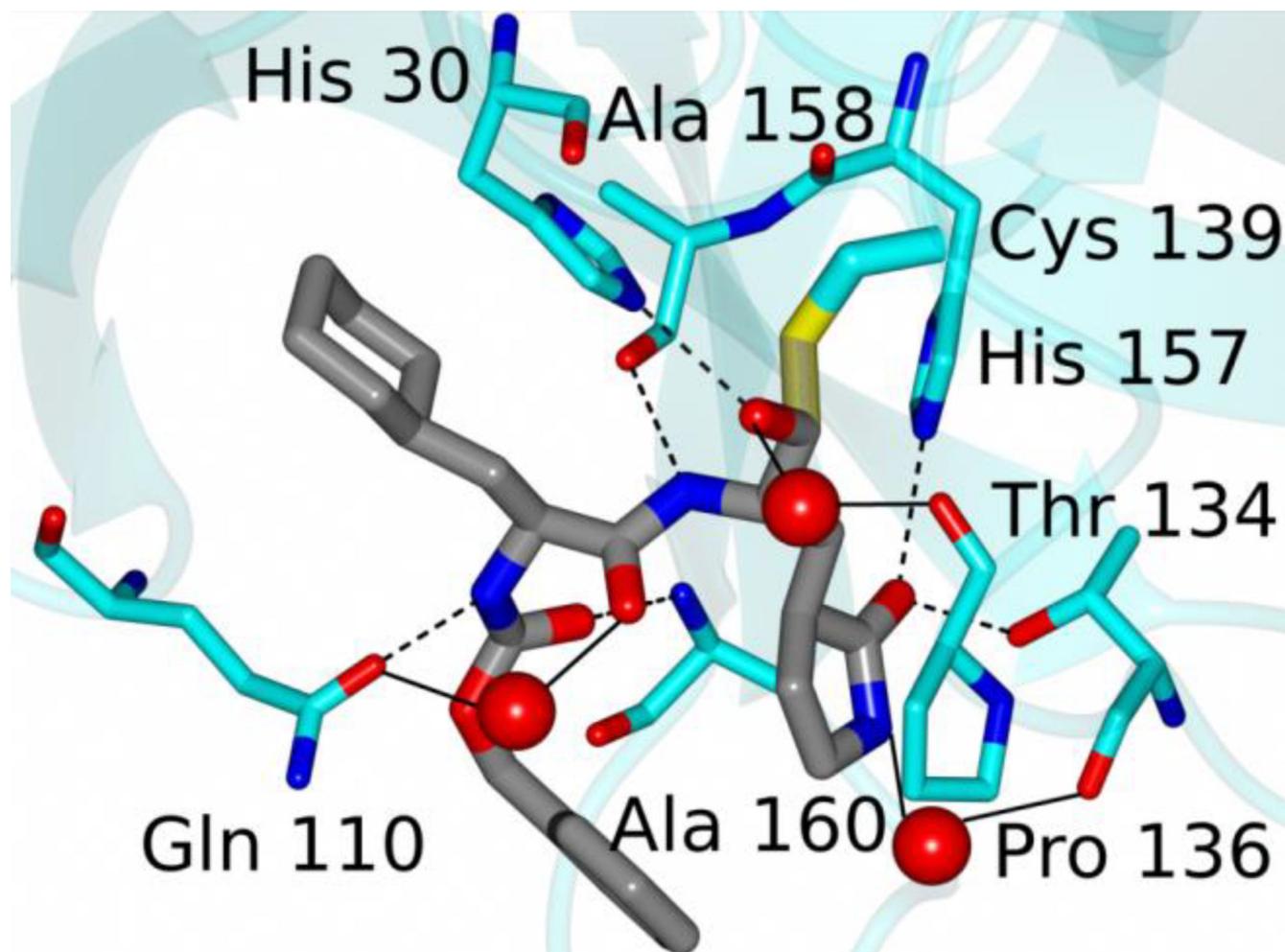
Figure 3.
Hydrogen bonding interactions between the inhibitor and chain B for NV 3CLpro:ligand (precursor aldehyde of inhibitor *44*) Side chain residues are colored cyan, inhibitor is colored gray, and water molecules are represented as red spheres. Hydrogen bonds are represented as dashed lines and water mediated contacts as solid lines.
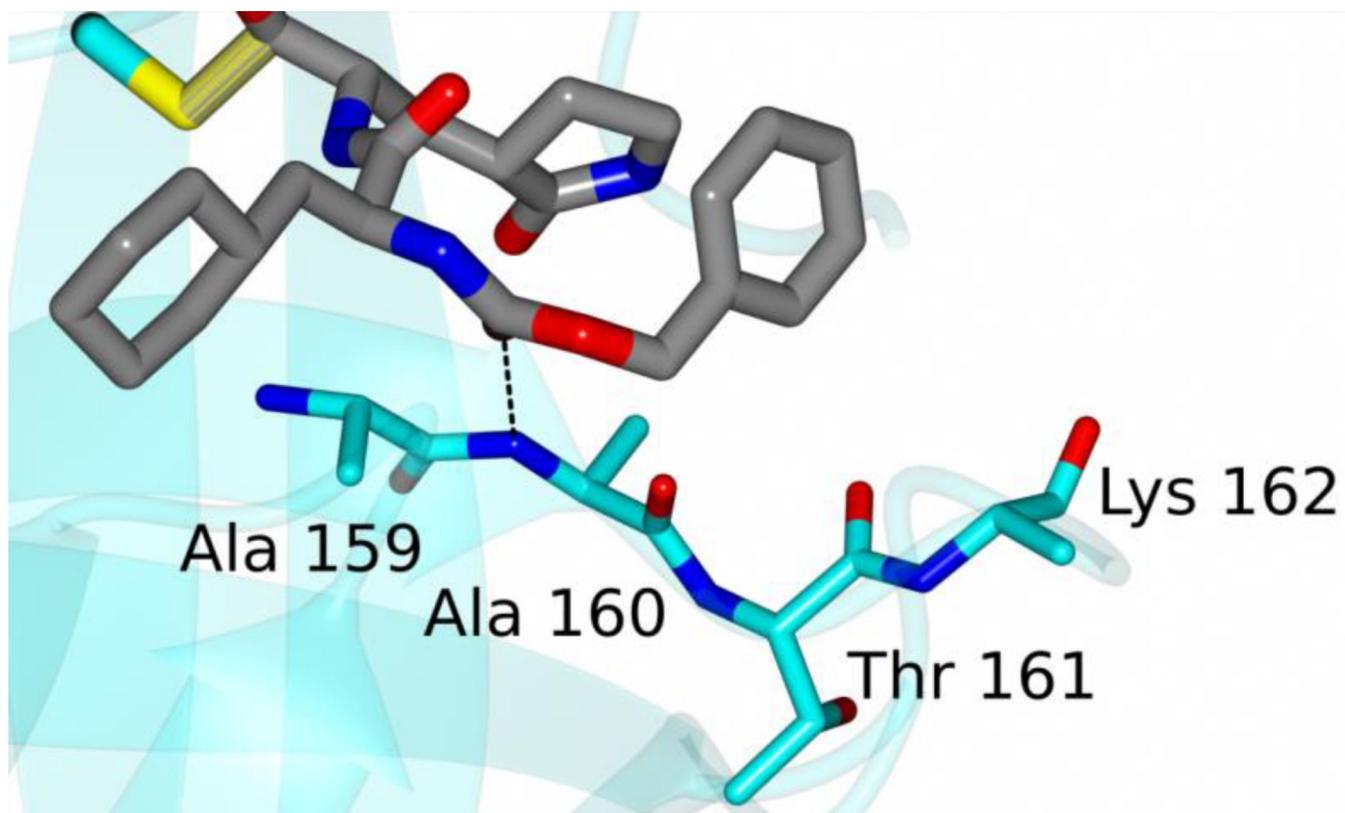
Figure 4.
Residues spanning Ala159-Lys162 located near the benzyl ring of precursor aldehyde of inhibitor **44**.
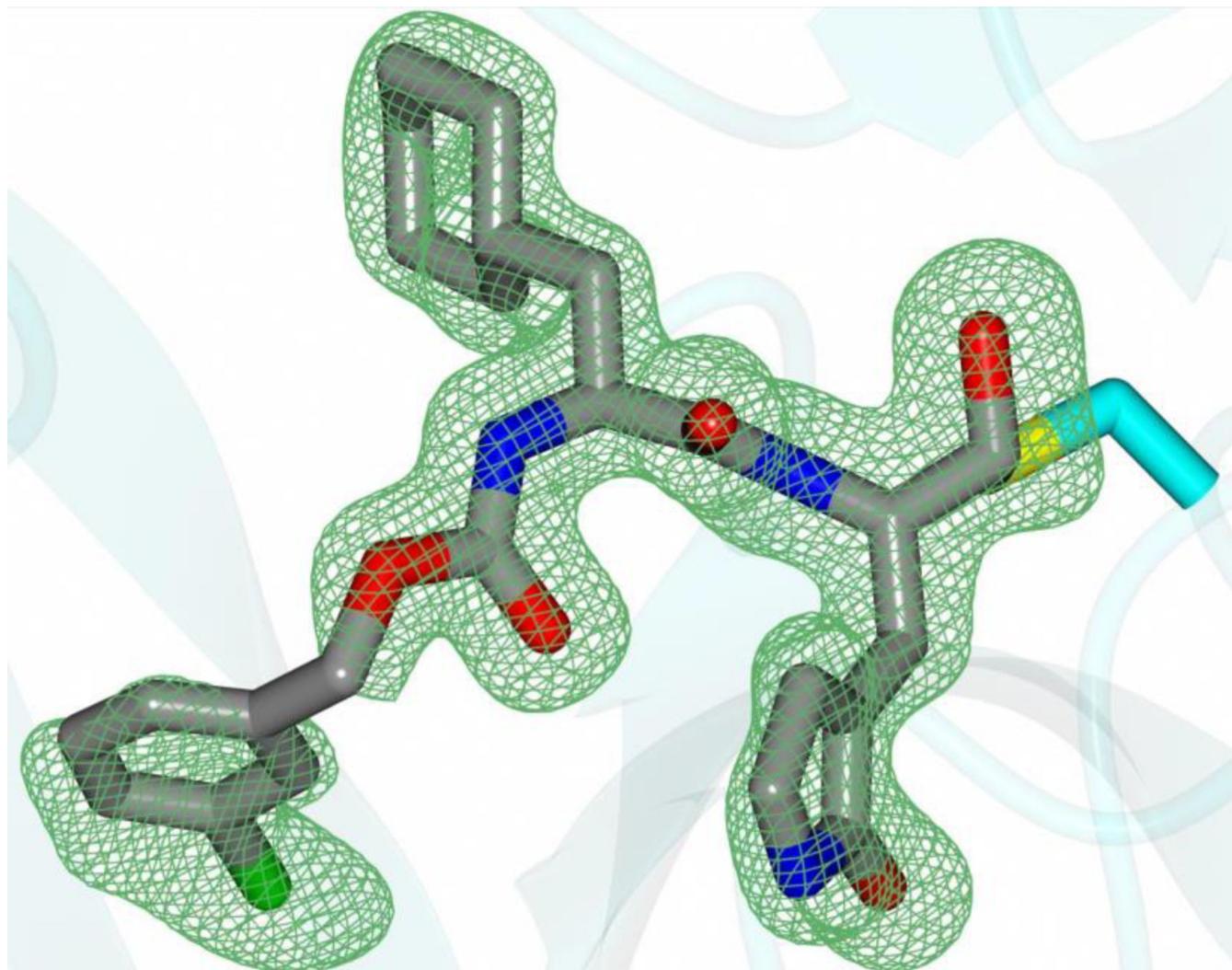
Figure 5.
Fo – Fc map (green mesh) of precursor aldehyde derived from compound *17* bound to NV
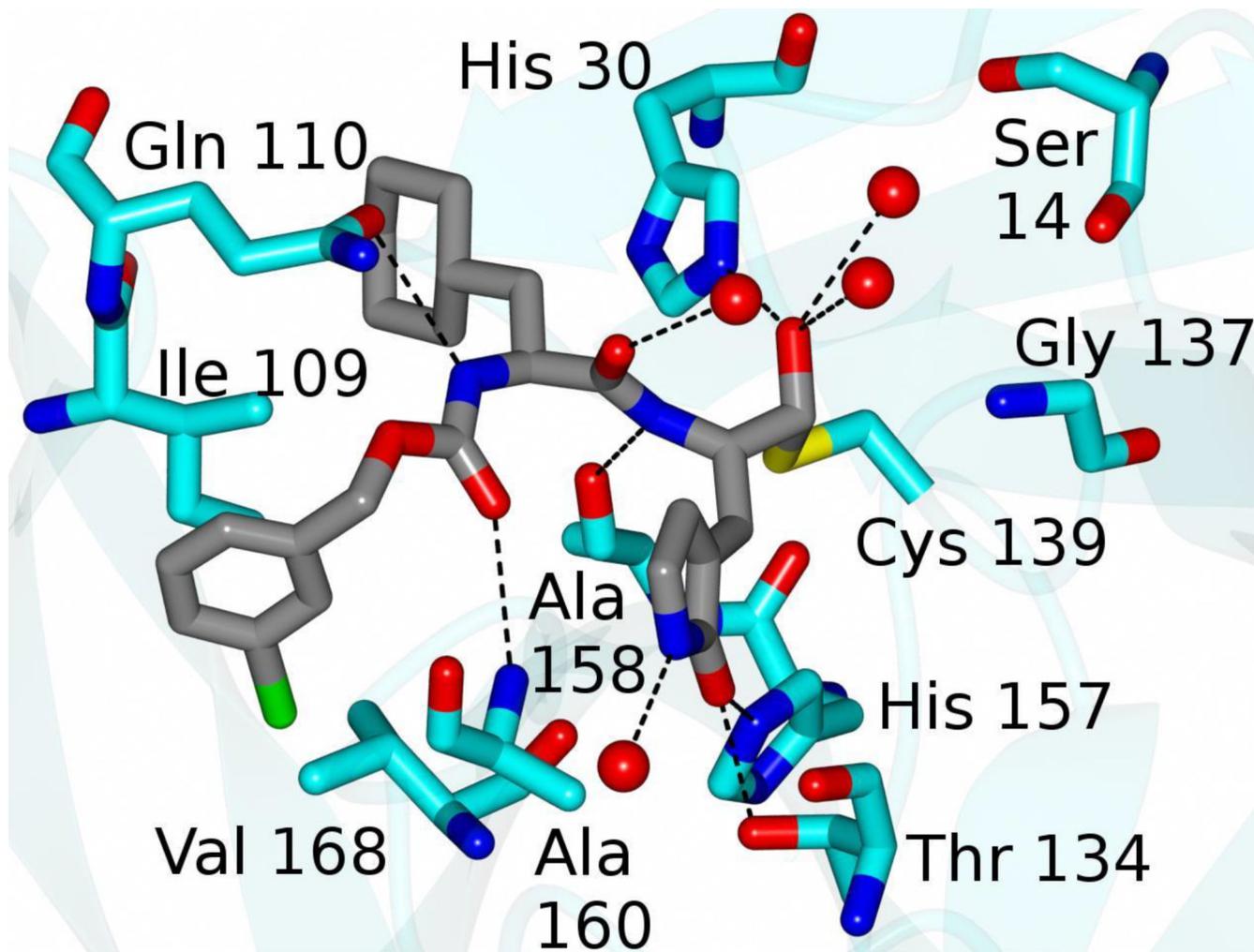3CLpro contoured at 3σ.

Figure 6.
Hydrogen bond interactions between NV 3CLpro and precursor aldehyde derived from compound *17*. Hydrogen bonds are represented as dashed lines and water mediated contacts are shown as solid lines.
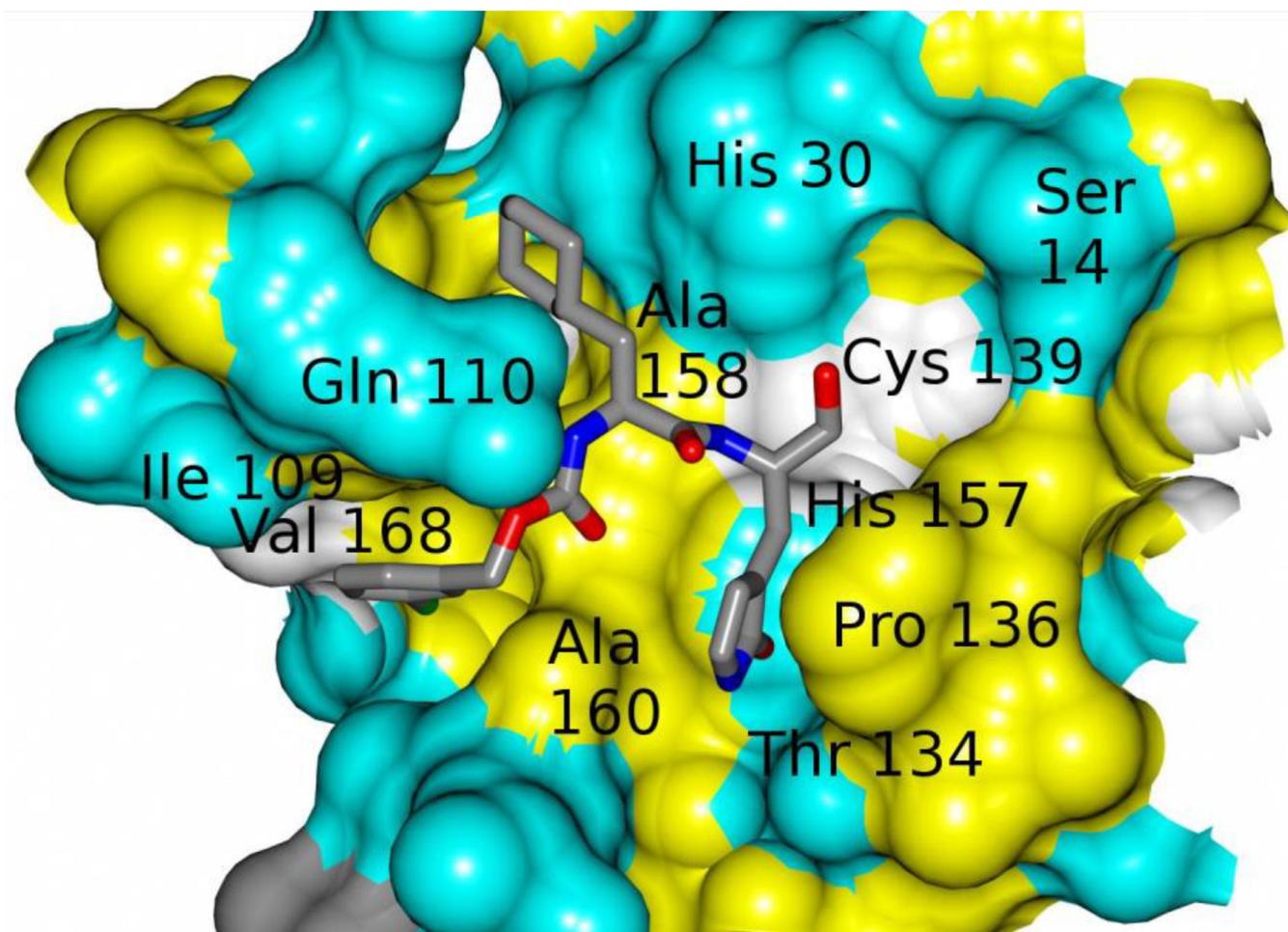
Figure 7.
Surface representation of precursor aldehyde of compound *17* bound to NV 3CLpro.
Neighboring residues of NV 3CLpro are colored yellow (nonpolar), white (weakly polar)
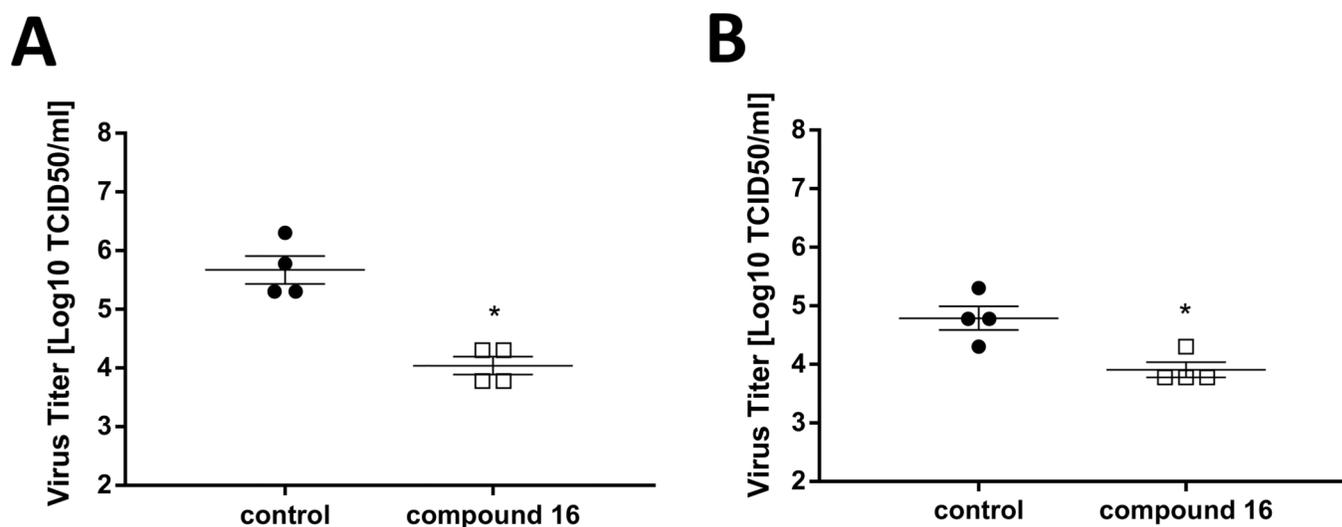and cyan (polar).

Figure 8.
Protease inhibitor suppresses norovirus replication in the intestinal tract in mice (effects of a 3CLpro inhibitor treatment on murine norovirus titers in the intestinal tracts of mice). Balb/c mice were orally infected with MNV-1 at 2×10^6 TCID$_{50}$ and intraperitoneally given compound *16* at 10 mg/kg twice daily with the first dose starting 4 hrs prior to virus infection. At 72 h post virus infection, the intestinal tract tissues were harvested and processed for determination for virus titers by the TCID$_{50}$ assay. Mean and the standard error of the mean of virus titers in the small intestine (A) or large intestine (B) at 3 day post MNV-1 infection are shown. Each filled circle indicates a control mouse that was given drug vehicle and each empty box indicates a mouse given compound *16*. Asterisk indicates *P*<0.01.
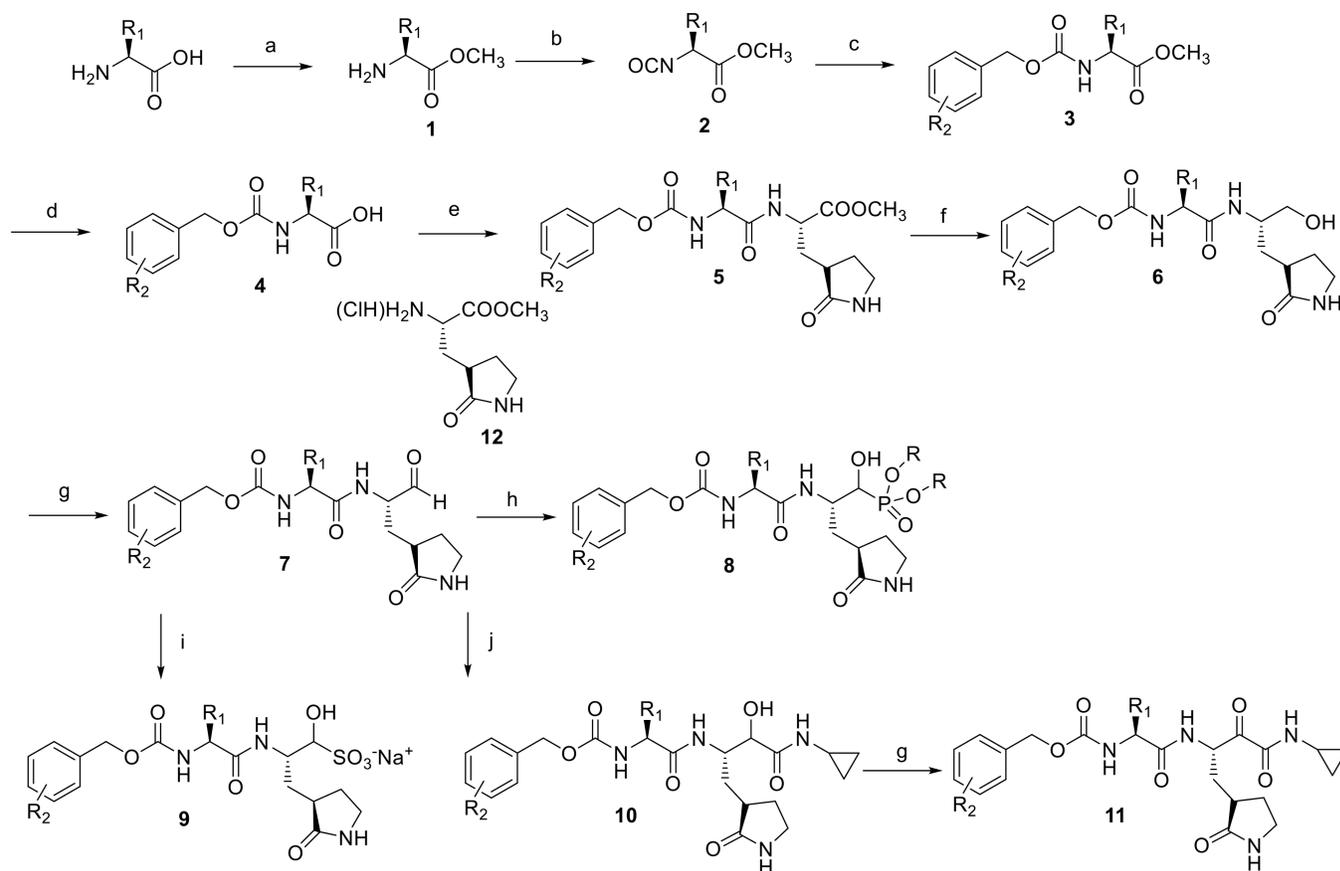
Scheme 1.
Synthesis of inhibitors *13–44*.
[a]CH$_3$OH/SOCl$_2$/45 °C/ 3h; [b]CCl$_3$O(CO)Cl/ dioxane/ reflux/ 12h; [c] ArCH$_2$OH/ TEA/ acetonitrile/ reflux/ 2h; [d]1M LiOH(aq)/ THF/ RT/ 3h; [e]EDCl/ HOBt/ DIEA/ DMF; [f]2M LiBH$_4$/ THF/ CH$_3$OH; [g]Dess-Martin periodinane/ DCM; [h]Diethylphosphite/ DCM/ DIEA; [i]C$_2$H$_5$OH/ EtOAc/ HaHSO$_3$; [j]EtOAc/ HOAc/ cyclopropyl isocyanide/ K$_2$CO$_3$(aq)/ CH$_3$OH/ 18h

Table 1

Activity of compounds *13–44* against NV 3CL protease and norovirus in cell-based replicon cells.

Compound	R₁	R₂	X	IC50 (μM)	ED50 (μM)
13			CHO	0.8	0.08
14	o-Cl	Cha	CH(OH)P(O)(OC₂H₅)₂	>10	0.2
15			CH(OH)(SO₃–Na⁺)	0.7	0.07
16			CHO	0.1	0.02
17			CH(OH)P(O)(OC₂H₅)₂	>10	0.06
18		Cha	CH(OH)(SO₃–Na⁺)	0.1	0.02
19	m-Cl		C(O)C(O)NH(Cyc-Prop)	0.25	0.085
20			C(O)NHSO₂CH₃	>10	>2
21		Chg	CHO	5.1	0.15
22		Leu	CHO	0.9	0.1
23			C(O)C(O)NH(Cyc-Prop)	5.5	0.3
24	p-Cl	Cha	CHO	0.71	0.08
25	o-F	Cha	CHO	0.9	0.08
26			CH(OH)P(O)(OC₂H₅)₂	>10	>2
27	m-F	Cha	CHO	1.2	0.09
28			C(O)C(O)NH(Cyc-Prop)	3.5	0.3
29			CHO	0.3	0.03
30			CH(OH)P(O)(OC₂H₅)₂	6.5	0.15
31	m-Br	Cha	CH(OH)(SO₃–Na⁺)	0.15	0.03
32			C(O)C(O)NH(Cyc-Prop)	1.5	0.2
33			CHO	0.35	0.05
34	m-I	Cha	CH(OH)(SO₃–Na⁺)	0.15	0.04
35	o-OCH3	Cha	CHO	>10	0.1
36	m-OCH3	Cha	CHO	1.4	0.15
37	m-CN	Cha	CHO	>10	0.2
38	m-NHBoc	Cha	CHO	4.5	1.5

Compound	R_1	R_2	X	IC50 (μM)	ED50 (μM)
39	m-NO2	Cha	CHO	0.6	0.15
40			CHO	0.6	0.2
41		Leu	CH(OH)(SO3−Na+)	0.8	0.3
42	H		C(O)C(O)NH(Cyc-Prop)	3.4	1.1
43			CHO	0.3	0.06
44		Cha	CH(OH)(SO3−Na+)	0.4	0.06

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Selectivity of selected compounds against a panel of proteases.

Enzyme[25]	I/E	Compound (% inhibition)				[I]$_{if}$ (μM)
		16	17	29	30	
HNE	50	15	86.5	26.5	63	15.4
Chymotrypsin	250	0	2.4	0	0	2.5
Trypsin	250	5	8.8	0	1	125
Thrombin	250	0	0	0	0	2.75
Factor Xa	250	1	0	0	6	1.23
Plasmin	250	0	15	7	13	2.5
Carboxypeptidase A	250	12	14.3	11.6	0	43

Table 3

Crystallographic data for NV 3CLpro:inhibitor structures.

	NV 3CLpro:17	NV 3CLpro:44-h	NV 3CLpro:44-o
Data Collection			
Unit-cell parameters (Å, °)	a=124.23, c=49.97	a=121.85, c=51.50	a=37.47 b=66.91 c=126.4
Space group	$P6_522$	$P6_522$	$P2_12_12_1$
Resolution (Å)[1]	45.32-1.85	46.28-1.60	45.94-1.45
	(1.89-1.85)	(1.63-1.60)	(1.48-1.45)
Wavelength (Å)	1.0000	1.0000	1.0000
Temperature (K)	100	100	100
Observed reflections	384,657	579,016	371,208
Unique reflections	19,899	30,219	57,354
$<I/\sigma(I)>$[1]	14.9 (2.4)	16.3 (2.0)	18.9 (1.8)
Completeness (%)[1]	100 (100)	100 (100)	100 (100)
Multiplicity[1]	19.3 (18.5)	19.2 (16.9)	6.5 (6.5)
Rmerge (%)[1,2]	19.7 (150.5)	16.6 (169.2)	4.7 (105.4)
R_{meas} (%)[1,4]	20.3 (154.7)	17.0 (174.5)	5.1 (114.5)
R_{pim} (%)[1,4]	4.6 (35.8)	3.9 (42.3)	2.0 (44.4)
CC1/2 [1,5]	0.998 (0.763)	99.9 (71.1)	100 (63.5)
Refinement			
Resolution (Å) [1]	45.32-1.85	36.85-1.60	35.65-1.45
Reflections (working/test)[1]	18,907/968	28,660/1,530	54,369/2,905
R_{factor} / R_{free} (%)[1,3]	16.3/19.7	16.3/17.5	17.3/19.9
No. of atoms	1,236/33/129	1,319/32/160	2,444/58/218
(Protein/Ligand/Water)			
Model Quality			
R.m.s deviations			
Bond lengths (Å)	0.009	0.009	0.008
Bond angles (°)	1.229	1.234	1.020
Average B-factor (Å^2)			
All Atoms	21.1	17.2	24.3
Protein	19.9	15.6	23.6
Ligand	25.7	18.5	25.4
Water	29.9	19.0	32.1
Coordinate error	0.16	0.18	0.15
(maximum likelihood) (Å)			
Ramachandran Plot			
Most favored (%)	98.8	98.3	98.5
Additionally allowed (%)	1.2	1.7	1.5

[1] Values in parenthesis are for the highest resolution shell.

[2] $R_{\text{merge}} = \Sigma_{hki}\Sigma_i \mid I_i(hkl) - <I(hkl)> \mid / \Sigma_{hki}\Sigma_i I_i (hkl)$, where $I_i(hkl)$ is the intensity measured for the i th reflection and $<I(hkl)>$ is the average intensity of all reflections with indices hkl.

[3] $R_{\text{facor}} = \Sigma_{hkl} \mid\mid F_{\text{obs}}(hkl) \mid - \mid F_{\text{calc}}(hkl) \mid\mid / \Sigma_{hkl} \mid F_{\text{obs}}(hkl) \mid$; Rfree is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.

[4] R_{meas} = redundancy-independent (multiplicity-weighted) R_{merge}[28,37]. R_{pim} = precision-indicating (multiplicity-weighted) R_{merge}[38, 39].

[5] $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data [40,41].