Structural and Biochemical Studies of Inhibitor Binding to Human Cytomegalovirus Protease†

Reza Khayat,‡ Renu Batra,‡,⊥ Chungeng Qian,§ Teddy Halmos,|| Murray Bailey,|| and Liang Tong*,‡

Department of Biological Sciences, Columbia University, New York, New York 10027, Boehringer Ingelheim Pharmaceuticals, Inc., Department of Medicinal Chemistry, 900 Ridgebury Road, Ridgefield, Connecticut 06877, and Boehringer Ingelheim (Canada) Ltd., Research and Development, 2100 Cunard Street, Laval, Quebec H7S 2G5, Canada

Received October 21, 2002; Revised Manuscript Received November 26, 2002

ABSTRACT: Herpesvirus protease is required for the life cycle of the virus and is an attractive target for the design and development of new anti-herpes agents. The protease belongs to a new class of serine proteases, with a novel backbone fold and a unique Ser−His−His catalytic triad. Here we report the crystal structures of human cytomegalovirus protease in complex with two peptidomimetic inhibitors. The structures reveal a new hydrogen-bonding interaction between the main chain carbonyl of the P_5 residue and the main chain amide of amino acid 137 of the protease, which is important for the binding affinity of the inhibitor. Conformational flexibility was observed in the S_3 pocket of the enzyme, and this is supported by our characterization of several mutants in this pocket. One of the structures is at 2.5 Å resolution, allowing us for the first time to locate ordered solvent molecules in the inhibitor complex. The presence of two solvent molecules in the active site may have implications for the design of new inhibitors against this enzyme. Favorable and stereospecific interactions have been established in the S_1' pocket for one of these inhibitors.

Members of the *herpesviridae* family infect most species throughout the animal kingdom, and nine human herpesviruses are currently known (*1*). Herpesviruses are large, double-stranded DNA viruses that vary greatly in their biological properties, such as cell tropism, immunological response, genome size, and genome content. On the basis of these differences, the viruses are divided into three subfamilies designated as α, *β*, and *γ*. The herpes simplex virus type 1 (HSV-1), HSV-2, and varicellar-zoster virus (VZV) belong to the α sub-family. The *β* sub-family includes the human cytomegalovirus (HCMV), human herpesvirus 6A (HHV6A), HHV6B, and HHV7. The *γ* sub-family includes the Epstein−Barr virus (EBV) and Kaposi-sarcoma associated herpesvirus (KSHV, also known as HHV8). Herpesvirus infections in humans cause a variety of disorders ranging in severity from childhood chickenpox (VZV infection) to potentially fatal complications of HCMV infection in immunocompromised or immunosuppresed patients.

Herpesviruses encode a protease that is essential for the production of infectious virions (*2*, *3*). This protease catalyzes the maturational processing of the herpesvirus assembly protein, which forms a scaffold that enables the assembly of the herpesvirus capsid. Biochemical studies identified the protease as a serine protease. Solution studies showed that the protease exists in a monomer−dimer equilibrium, with the dimer being the active species (*3*−*5*). The first structural information, on the free enzyme of HCMV protease, revealed the protease to possess a new polypeptide backbone fold and a novel catalytic triad (Ser132−His63−His157) (*6*−*10*).

Herpesvirus protease is an attractive target for the development of new anti-herpes agents (*2*, *3*). A series of peptidomimetic inhibitors has been developed based on the substrate preference of HCMV protease (*11*). A crystal structure of the protease in complex with one such compound, inhibitor 1 (Figure 1), shows that the inhibitor binds in an extended conformation, forming an antiparallel *β*-sheet with the protease (*12*). The catalytic Ser132 residue attacks the activated carbonyl group (α-ketoamide) in the inhibitor (Figure 1), forming a (reversible) covalent complex. There are major conformational changes in the protease upon inhibitor binding, suggesting that HCMV protease is an induced-fit enzyme (*10*, *12*−*14*). HCMV protease shares 26% amino acid sequence identity with that of HSV-1 protease. These enzymes have similar structures and substrate preferences (*2*, *3*, *10*), and structural information obtained on inhibitor complexes of HCMV protease will be applicable to the design of inhibitors against HSV-1 protease as well. To further understand the molecular basis for the observed structure−activity relationships (SAR) of these compounds, here we present the crystal structures of HCMV protease in complex with two other peptidomimetic inhibitors and in different crystal forms. Our structures demonstrate the importance of a hydrogen bond from the P_5 residue of the inhibitor and reveal conformational flexibility in the S_3 pocket of the protease. In addition, we have performed mutagenesis, biochemical, and structural studies to assess the functional roles of residues in the S_3 pocket.

† This work was supported in part by the National Institutes of Health (Grant AI46139 to L.T.). R.K. was supported by a NIH training program in molecular biophysics (GM08281).

* Corresponding author. Phone: (212) 854-5203. Fax: (212) 854-5207. E-mail: tong@como.bio.columbia.edu.

‡ Columbia University.

§ Boehringer Ingelheim Pharmaceuticals, Inc.

|| Boehringer Ingelheim (Canada) Ltd.

⊥ Current address: Institut fur Biologische Informationsverarbeitung, Forschungszentrum Julich, D-52425 Julich, Germany.

10.1021/bi027045s CCC: $25.00
Published on Web 01/09/2003

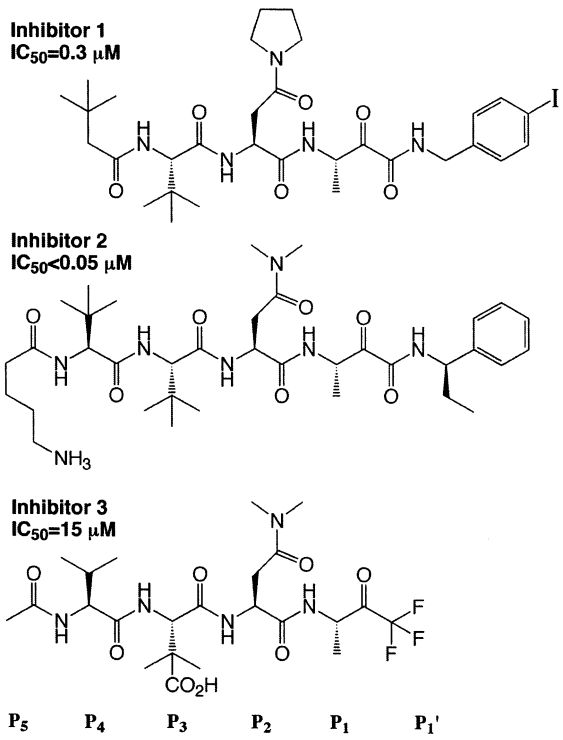

FIGURE 1: Chemical structures of peptidomimetic inhibitors of HCMV protease. The residues are identified from P_5 to P_1', and the inhibitory potencies are indicated as well.

MATERIALS AND METHODS

Preparation of Enzyme. The protease samples used in these studies contain the A143Q mutation, which eliminates the auto-proteolysis at this position (*15*). This mutant maintains the catalytic activity of the native form of the enzyme, and protease samples carrying the A143Q mutation will be referred to as the wild type here. The protease samples were expressed in the *Escherichia coli* strain BL21(DE3) pLysS, and the purification followed the protocol described earlier (*6, 12*). Site-directed mutagenesis was performed with the QuikChange kit (Stratagene), and the presence of the desired mutation was confirmed by sequencing. The mutants were expressed and purified following the same protocol as that for the A143Q sample.

Crystallization, Data Collection, and Data Processing. Crystals were grown at 21 °C using the hanging-drop vapor diffusion method from a solution containing 1 μL of 7 mg/mL protein in 20 mM HEPES (pH 7.0), 80 mM Na_2SO_4, 40 mM NaCl, 2 mM DTT, 2 mM EDTA, and 2 mM of inhibitor in DMF (2.5% final concentration in protein solution) and 1 μL of the reservoir solution (16−18% PEG 4000, 0.1 M HEPES (pH 7.0), 10−12% glycerol, 0.3 M NaCl, 5 mM spermine-HCl, 2 mM DTT, and 2 mM EDTA). The protease−inhibitor mixture was incubated for 40 min prior to crystallization setup. For inhibitor 2, both orthorhombic and tetragonal crystals were obtained (Table 1). Crystals were flash-frozen in liquid propane for data collection at 100 K. X-ray diffraction data up to 2.5 Å resolution were collected at synchrotron radiation sources, including National Synchrotron Light source (NSLS) beam lines X4A and X12C and Advanced Photon Source (APS) beam line 19BM (Table 1). The diffraction images were processed with the HKL package (Table 1) (*16*).

Structure Determination and Refinement. The structures of the protease−inhibitor complexes were determined, when necessary, by molecular replacement with the COMO program (*17*), using the structure in complex with inhibitor 1 as the search model (*12*). The molecular replacement solution was subjected to rigid-body refinement using reflections between 5.0 and 3.5 Å resolution. Structure factors were calculated for all reflections to the highest resolution of the observed data using the atomic model after rigid-body refinement, and the calculated phases were applied to the observed structure factor amplitudes. The phase information was improved by 4-fold noncrystallographic symmetry averaging for the orthorhombic structures and 2-fold noncrystallographic symmetry averaging for the tetragonal structure with the program DM (*18*). The atomic model was rebuilt against the resulting electron density map with the program O (*19*). The structure refinement was carried out with the program CNS (*20*). The crystallographic information is summarized in Table 1.

Kinetic Measurements. The k_{cat} and K_m values for each enzyme were measured by monitoring the hydrolysis of an internally quenched peptide substrate (Bachem) at 30 °C (λ_{ex} = 355 nm and λ_{em} = 495 nm) (*21*). The reaction mixture was composed of 100 mM MOPS (pH7.2), 0.5 M Na_2SO_4, 10% glycerol, 10 mM DTT, and 2 mM EDTA. The fluorescence signal of the cleaved peptide was monitored using a spectrofluorimeter (Photon Technologies Inc.). The parameters for the wild type, E31S, and R137E enzymes were obtained from classical Michaelis−Menten kinetics. However, the E31R mutant exhibited a significantly reduced catalytic activity, and enzyme concentrations up to 1 μM were required for reliable measurements of the reaction rates. Therefore, a different protocol was used to obtain the kinetic parameter of this mutant (*4, 22*). The kinetic parameters were extracted from the experimental observations using nonlinear least-squares fitting, with the program Origin (Microcal Software, Inc.)

RESULTS AND DISCUSSION

Overall Structures. Crystal structures of HCMV protease (A143Q mutant) in complex with the peptidomimetic inhibitor 2 (Figure 1) have been determined in an orthorhombic and a tetragonal crystal form at 2.7 Å resolution (Table 1). For inhibitor 3, the structure of the complex in the orthorhombic crystal form has been determined at 2.5 Å resolution (Table 1). The diffraction data set for this inhibitor complex is of sufficient quality to allow the placement of ordered solvent waters in the active site region of the protease (see below). Finally, we have also determined the crystal structure of inhibitor 3 in complex with the E31R mutant of the protease at 2.6 Å resolution. This mutant is among several that are studied here to characterize the functional contributions of residues lining the S_3 pocket (see below).

All the atomic structures have good *R* factors and geometric parameters. The crystallographic information is summarized in Table 1. The atomic coordinates have been deposited at the Protein Data Bank (accession numbers 1NJU, 1NJT, 1NKK, and 1NKM).

Binding Mode of Inhibitor 2 in the Orthorhombic Crystal Form. The orthorhombic crystal form of the inhibitor 2 complex is isomorphous to that of the inhibitor 1 complex reported by us earlier, with two dimers in the crystallographic

Table 1: Summary of Crystallographic Information

inhibitor	2	2	3	3
protease	wild-type (A143Q)	wild-type (A143Q)	wild-type (A143Q)	E31R (A143Q)
crystal form	orthorhombic	tetragonal	orthorhombic	orthorhombic
cell parameters (*a*, *b*, *c*) (Å)	105.5, 213.5, 52.3	73.9, 73.9, 216.4	107.1, 213.3, 52.8	105.3, 215.2, 52.3
maximum resolution (Å)	2.7	2.7	2.5	2.6
number of observations	143 394	66 771	182 204	103 436
R_{merge} (%)[a]	7.0	6.3	6.0	9.9
resolution range for refinement	20−2.7	20−2.7	20−2.5	20−2.6
number of reflections	27 260	13 923	37 470	30 941
completeness (%)	82	81	88	83
R/free *R* factor[b] (%)	22.3/27.1	23.6/29.8	22.8/27.3	23.3/28.0
rms deviation in bond lengths (Å)	0.007	0.008	0.007	0.007
rms deviation in bond angles (deg)	1.3	1.4	1.3	1.3

[a] $R_{merge} = \sum_h\sum_i|I_{hi} - \langle I_h\rangle|/\sum_h\sum_i I_{hi}$. [b] $R = \sum_h|F_h^o - F_h^c|/\sum_h F_h^o$.

asymmetric unit (*12*). The overall structure of the protease and the binding mode of the inhibitors are similar between the two complexes. The rms distance between equivalent Cα atoms of any pair of monomers of the two complexes is 0.4 Å. Similarly, the rms distance between equivalent Cα atoms of the two dimers is also 0.4 Å, suggesting that the organization of the two dimers are similar as well.

Inhibitor 2 is one of the most potent peptidomimetic inhibitors against HCMV protease, with IC_{50} of less than 50 nM (below the detection limit) (Figure 1) (*11*), whereas inhibitor 1 has an IC_{50} of about 300 nM (*12*). Structure−activity relationship (SAR) studies with the inhibitors suggest that the presence of the des-amino lysine residue at P_5 produces a significant enhancement in potency. In addition, the introduction of an ethyl substituent in the P_1' moiety is also beneficial for the inhibitory activity of inhibitor 2 as compared to inhibitor 1 (Figure 1) (*11*).

The crystal structure of the inhibitor 2 complex provides an explanation at the molecular level for the additional binding affinity of this inhibitor. The bound conformation of the inhibitor is clearly defined by the electron density map (Figure 2A). Like inhibitor 1, inhibitor 2 is complexed to the protease in an extended conformation, forming an antiparallel β-sheet with strand β5 of the protease (Figure 3A), with hydrogen bonds from the P_3 (amide and carbonyl) and P_1 (amide) residues of the inhibitor (Figure 3B). In contrast to inhibitor 1, inhibitor 2 shows an additional hydrogen bond to the protease, between the P_5 carbonyl and the main chain amide of Arg137 (Figure 3B). On the other hand, the side chain of the P_5 Lys residue is mostly disordered in the structure (Figure 2A) and has essentially no interaction with the protease. Therefore, the structural analysis shows that the enhancement in potency by the introduction of the P_5 residue is due mostly to the formation of the additional hydrogen bond to the protease. It has been estimated that an additional backbone hydrogen bond in a protease−inhibitor complex can donate approximately −1.5 kcal/mol of binding energy (*23*).

SAR observations showed that introducing a P_6 residue only marginally increases the potency of inhibitors (*11*). On the basis of our current structure model, there is a sharp change in the direction of the polypeptide in the protease after Arg137, and it is unlikely for the main chain of the P_6 residue to establish hydrogen-bonding interactions with the protease.

The other difference between inhibitor 2 and inhibitor 1 is the introduction of an ethyl group at the P_1' position of

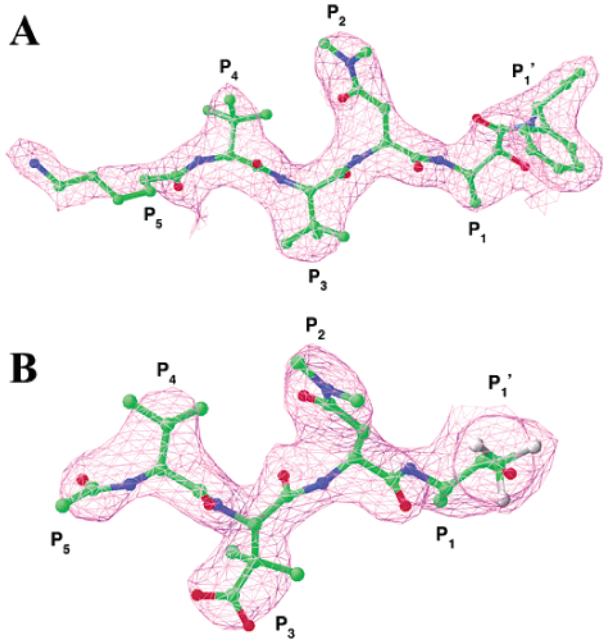

FIGURE 2: Observed electron density for inhibitors. Final $2F_o - F_c$ electron density map for (A) inhibitor 2 in the orthorhombic crystal form at 2.7 Å resolution and (B) inhibitor 3 at 2.5 Å resolution. The contour level is at 1σ. Produced with Ribbons (*26*).

inhibitor 2 (Figure 1), which results in a 2-fold increase in the inhibitory activity (*11*). The structural analysis shows that this ethyl group is in van der Waals contact with residues Cys161, Ser162, and Ile231 of the protease (Figure 3C), producing about 70 Å^2 of buried surface area and therefore explaining the beneficial effect of this group. The structure also clearly indicates that the ethyl substituent is stereospecific. The other stereoisomer at this position would not only point the ethyl group into the solvent, but it will also cause steric hindrance between the ethyl group and the oxygen atom of the P_1' amide group (Figure 3C). SAR studies showed that introduction of the other stereoisomer at this position caused an 18-fold loss in the activity of the inhibitor (*11*).

About 660 Å^2 of the surface area of inhibitor 2 is buried at the interface with the protease. The majority of the surface burial is accounted for by the interactions at the P_3, P_1, and the P_1' positions with the protease (Figure 3D), corroborating the importance of these residues for inhibitor binding. The P_4 and P_2 residues are exposed to the solvent and make only small contributions to this burial.

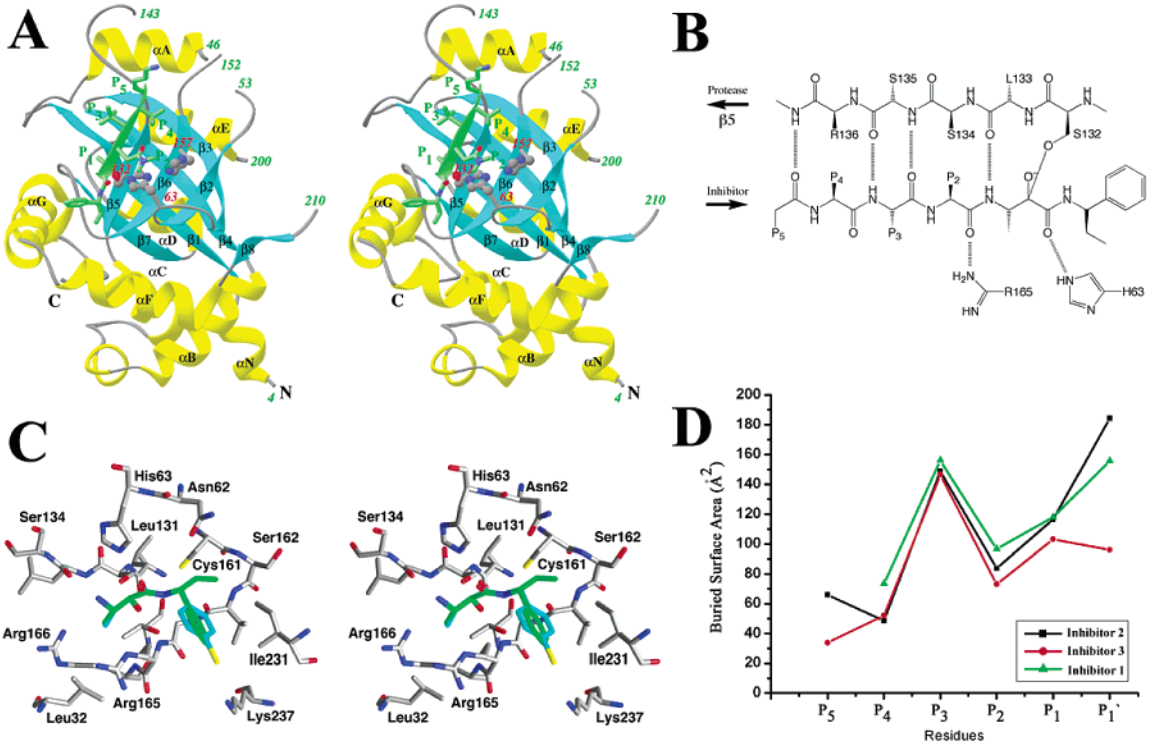

FIGURE 3: Crystal structure of the inhibitor 2 complex in the orthorhombic crystal form. (A) Schematic drawing of the structure of HCMV protease in complex with inhibitor 2. The inhibitor is shown in green. Produced with Ribbons (*26*). (B) Schematic drawing of the hydrogen-bonding interactions between inhibitor 2 and HCMV protease. (C) Stereo diagram showing the interactions of the P_1' residue of inhibitor 2 (in green) and the protease. The P_1' group of inhibitor 1 is shown (in cyan) for comparison. Produced with Grasp (*27*). (D) Contribution of the P_5 to P_1' groups of the inhibitors to the buried surface area at the interface with the protease.

Binding Mode of Inhibitor 3. As compared to inhibitor 1, inhibitor 3 contains an acetyl group as a mimic for the P_5 residue, a valine as the P_4 residue, a dimethyl aspartic acid residue at the P_3 position, and a trifluoromethyl ketone (TFMK) as the activated carbonyl group (Figure 1). Crystals of HCMV protease in complex with this inhibitor are also in the orthorhombic form, isomorphous to that of the inhibitor 1 and inhibitor 2 complexes described earlier. The overall conformation of the protease and the inhibitors are similar among these different complexes. The binding mode of the inhibitor is clearly defined by the electron density map (Figure 2B). The acetyl group forms a hydrogen bond with the main chain amide of Arg137, equivalent to that from the P_5 residue of inhibitor 2 (Figure 3B).

The TFMK group is on the surface of the protease, with one of its fluorine atoms making a hydrogen bond with His63 (Figure 4C), the second member of the catalytic triad, similar to that from the carbonyl oxygen of the P_1' group in inhibitor 2 and inhibitor 1 (Figure 3B). Inhibitors containing the TFMK group are about 5-fold weaker as compared to those with the α-ketoamide group as in inhibitor 2 (Figure 1) (*11*). Part of this loss in potency is likely because of the absence of interactions observed for the phenyl ring in the α-ketoamide P_1' group (Figure 3C). In addition, the TFMK exists mostly as a hemiketal in solution, which is inactive as it is not susceptible to the attack by the catalytic Ser132 residue (*11*).

However, the introduction of the dimethyl aspartic acid residue at the P_3 position caused a significant conformational change in the S_3 binding pocket (Figure 4A). The side chain of this residue is bound such that the three carbon atoms are pointed toward the bottom of the pocket (Figure 4B), in a conformation that is essentially the same as that of the *tert*-butyl side chains of inhibitor 2 and inhibitor 1 (Figure 4A). The carboxylic oxygens of the side chain are pointed toward the rim of the S_3 pocket, formed by ion-pair interactions among Arg165, Glu31, and Arg137 (Figure 4B) (*12*). The carboxylic oxygens of the P_3 side chain in inhibitor 3 cause a conformational change in the side chain of Glu31, because of steric and charge repulsions, such that it now only interacts with that of Arg165. Possibly as a consequence, the side chain of Arg137 becomes disordered in the inhibitor 3 complex. (The main chain remains ordered, as it is stabilized by the hydrogen bond to the P_5 residue). This creates an opening in the side of the S_3 pocket in the inhibitor 3 complex as compared to the inhibitor 2 complex, and the P_3 side chain becomes accessible to solvent (Figure 4A,B). Surprisingly, this compound has only about a 14-fold loss in potency as compared to an inhibitor with *tert*-butylglycine at the P_3 position (see discussion below) (*11*). In addition, residues 139−143 are disordered in this complex, whereas they are partly ordered in the inhibitor 1 and inhibitor 2 complexes.

Two Water Molecules in the Active Site of the Protease. The X-ray diffraction data for the inhibitor 3 complex extends to 2.5 Å resolution, allowing us for the first time to locate ordered solvent molecules in such complexes. The structure reveals the presence of two water molecules in the active site (Figure 4C), at positions similar to those observed earlier in the free enzyme structures (*6*), the di-isopropyl phosphate (DIP) complex of HSV-2 protease (*24*), and the phenylmethylsulfonate (PMSF) complex of HCMV protease (*22*). These two waters are part of the oxyanion hole of the

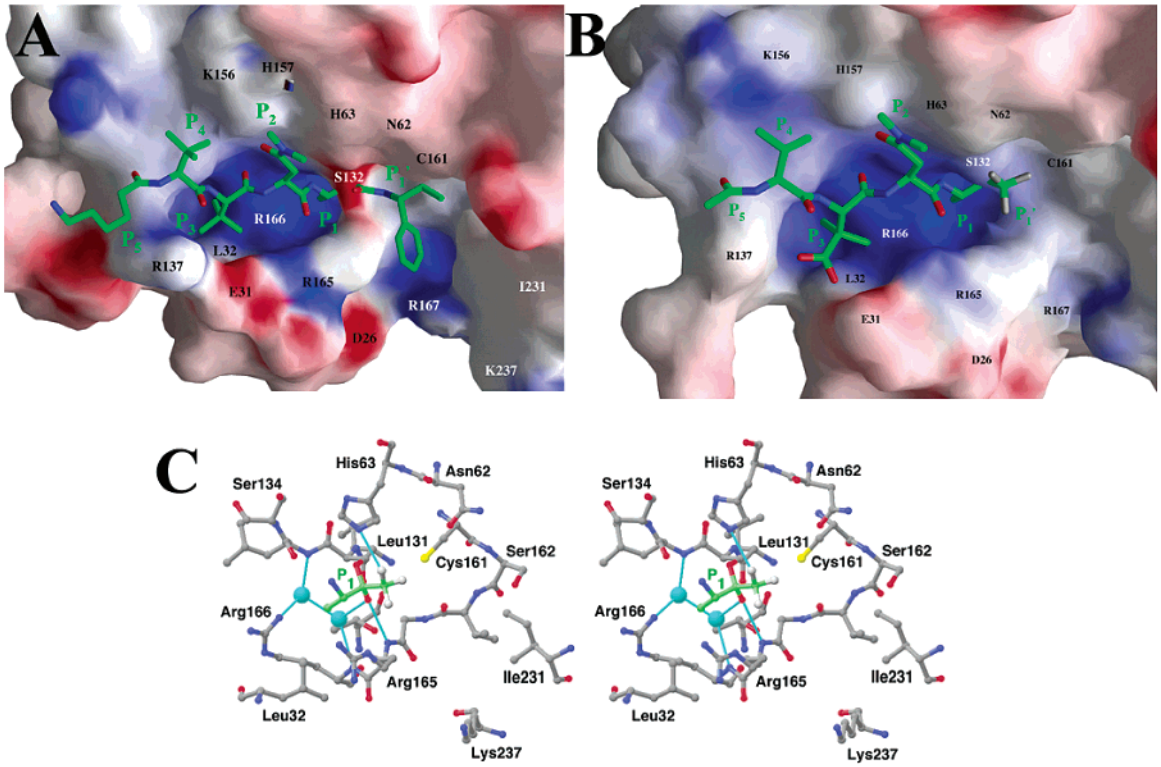

FIGURE 4: Crystal structure of the inhibitor 3 complex. (A) Molecular surface of HCMV protease in complex with inhibitor 2, colored according to electrostatic potential. The inhibitor is shown in the stick model. (B) Molecular surface of HCMV protease in complex with inhibitor 3. (C) Stereo diagram showing the binding of the P_1 (Ala) residue and the TFMK group of inhibitor 3 to the protease. Also shown are the two ordered water molecules in the active site (in cyan) at the bottom of the S_1 pocket. Panels A and B were produced with Grasp (*27*), and panel C was produced with Ribbons (*26*).

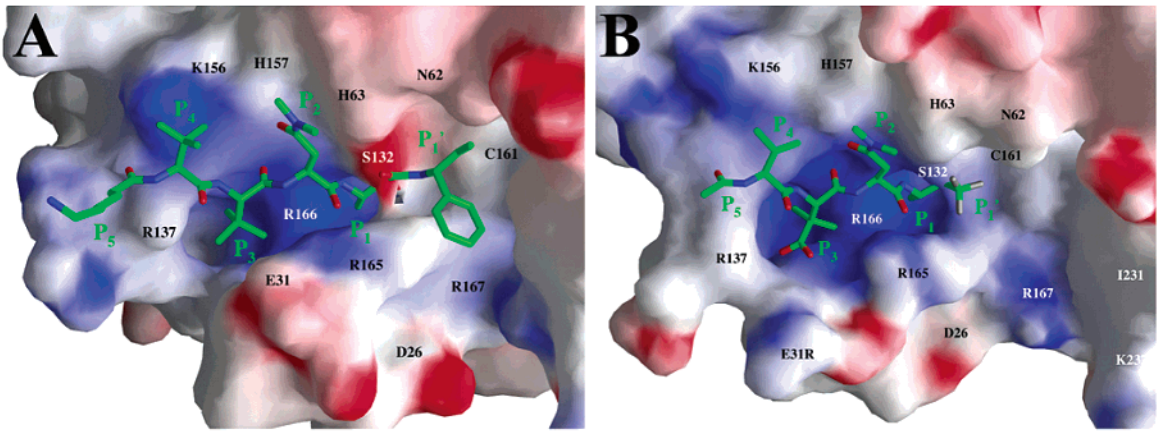

FIGURE 5: Flexibility of the S_3 binding pocket. (A) Molecular surface of HCMV protease in complex with inhibitor 2, in the tetragonal crystal form, colored according to electrostatic potential. The inhibitor is shown in the stick model. (B) Molecular surface of the E31R mutant of HCMV protease in complex with inhibitor 3, in the orthorhombic crystal form. Produced with Grasp (*27*).

protease, which stabilizes the developing negative charge on the scissile amide carbonyl during the nucleophilic attack. The oxyanion is stabilized by a direct hydrogen bond to the main chain amide of Arg165 and an indirect hydrogen bond to the side chain guanidinium group of Arg166 through the two water molecules (Figure 4C). Arg166 is strictly conserved among the herpesvirus proteases, and the R166A mutant has 1500-fold loss in the k_{cat} of the enzyme (*25*). These two water molecules are located at the bottom of the S_1 pocket. Inhibitors that are capable of replacing them may have significantly improved potency against the protease, as was observed when an ordered water molecule in the active site of HIV protease is replaced by the inhibitor (*10*).

Binding Mode of Inhibitor 2 in the Tetragonal Crystal Form. We have also obtained crystals of HCMV protease in complex with inhibitor 2 in the tetragonal crystal form that is isomorphous to that of the free enzyme crystals reported by us earlier (*6*). The inhibitor is bound to only one of the two protease molecules in the asymmetric unit. Interestingly, in this binding mode, the side chain of Arg137 is also disordered, similar to that observed in the inhibitor 3 complex. As a consequence, there is an opening in the side of the S_3 pocket (Figure 5A), even though the P_3 side chain is only a *tert*-butyl group in this inhibitor. Structural studies in this crystal form, and on the inhibitor 3 complex, therefore suggest that the S_3 pocket may have a significant degree of

Table 2: Summary of Kinetic Parameters

enzyme	k_{cat} (s^{-1})[a]	K_m (μM)[a]	k_{cat}/K_m ($s^{-1}\,M^{-1}$)[a]
wild-type (with A143Q)	0.033	10.1	3300
E31S	0.015 (2)	21.6 (0.5)	685 (5)
E31R	0.0028 (12)	37.6 (0.3)	74 (44)
R137E	0.0053 (6)	10.1 (1)	524 (6)

[a] The parameters for the wild-type enzyme are from ref *22*. The numbers in parentheses are the ratios between the wild-type and mutant values. See Materials and Methods for details on the kinetic assay for HCMV protease.

flexibility and plasticity. This is consistent with SAR results, which showed that side chains as large as the adamantyl group can be tolerated at the P_3 position with only a 3-fold loss in the potency of the inhibitor (*11*).

Inhibitor SAR as well as substrate studies consistently show that a hydrophobic side chain is preferred at the P_3 position (*3*, *11*). However, the crystal structure shows that the S_3 pocket is rather hydrophilic in nature (Figure 4A). The observation of Arg166 at the bottom of the S_3 pocket prompted the synthesis of compounds with Asp, Glu, and homo-Glu residues at the P_3 position, in an attempt to establish ionic interactions with this Arg residue. Surprisingly, all of these compounds showed significant reductions in inhibitory activity. Inhibitor 3 is one of the few active compounds that contain an acidic group at the P_3 position. Our structural analysis showed that the carboxylate group is pointed away from the Arg166 side chain, and the binding mode more closely mimics that of the *tert*-butyl side chain (Figure 4B). The molecular basis for the requirement of a hydrophobic residue at the P_3 position is currently not understood.

Mutagenesis Studies of Residues Forming the S_3 Pocket. Of the residues that form the S_3 pocket, only Arg166 is strictly conserved among the herpesvirus proteases, and mutation of Arg166 to Ala leads to a 1500-fold loss in the k_{cat} of the enzyme (*25*). Arg165 is conserved in most herpesvirus proteases, and the structures show that its guanidinium group is hydrogen bonded to the P_2 carbonyl oxygen (Figure 3B). Surprisingly, the R165A mutation has only a minor impact on the activity of the protease (*25*). The other residues in this pocket are not conserved. For example, although the Glu31 residue is generally maintained among the α- and β-subfamilies, it shows variation to Val, Cys, and other residues in the γ-subfamily. Similarly, the Arg137 residue is poorly conserved in the γ-subfamily, where changes to His and Cys are observed. Interestingly, murine CMV, a β-herpesvirus, contains an Arg at residue 31 and a Ser at residue 137.

The structural analysis showed that the Glu31 and Arg137 residues help form the S_3 binding pocket (Figure 4A). To characterize the functional importance of these two residues in the catalysis by HCMV protease, we have created single-site mutants at these positions and determined their kinetic parameters (Table 2). These mutants (E31R, E31S, and R137E) were made in an attempt to disrupt the ion-pair interactions that form the side of the S_3 pocket (Figure 4A), through charge repulsion or reduction in the bulk of the side chain.

Our kinetic data showed that these mutations have only minor effects on the catalysis by the protease, roughly a 5-fold loss in the k_{cat}/K_m of the enzyme for the E31S and R137E mutants (Table 2). The small effect of the R137E mutation is consistent with our structural observations showing that this side chain can be disordered in the binding of inhibitor 3 (Figure 4B) and inhibitor 2 (Figure 5A). As compared to the E31S mutation, the E31R mutation has a somewhat larger effect on the activity of the enzyme, with a 45-fold loss in the k_{cat}/K_m (Table 2). Overall, mutations of the Glu31 and Arg137 residues are generally tolerated by the enzyme, which is consistent with the sequence variability and structural flexibility of the S_3 pocket.

To understand the molecular basis for the effect of the E31R mutation, we have determined the crystal structure of this mutant in complex with inhibitor 3 at 2.6 Å resolution (Table 1). The crystal is in the orthorhombic form, isomorphous to that of the complex with wild-type protease. The mutation did not introduce any significant structural change in the protease, and the interactions of the inhibitor with the mutant enzyme is similar to that of the wild-type complex. Specifically, the side chain of the Arg137 residue is disordered in all four molecules (Figure 5B). The newly introduced Arg31 side chain is ordered in only two of the four molecules, where they have little interaction with the inhibitor (Figure 5B). Overall, the structural observations are in agreement with the small effects of the mutations in the S_3 pocket on the catalysis by the enzyme.

ACKNOWLEDGMENT

We thank Gerwald Jogl, Xiao Tao, and Zhiru Yang for help with data collection at synchrotron radiation sources and Lisette Lagacé and Marie-Josée Massariol for providing the A143Q expression construct of HCMV protease.

REFERENCES

1. Knipe, D. M., and Howley, P. M. (2001) *Fields Virology*, Vol. 2, Lippincott Williams & Wilkins, Philadelphia.
2. Gibson, W., Welch, A. R., and Hall, M. R. T. (1995) *Perspect. Drug Discovery 2*, 413−426.
3. Waxman, L., and Darke, P. L. (2000) *Antivir. Chem. Chemother. 11*, 1−22.
4. Batra, R., Khayat, R., and Tong, L. (2001) *Nature Struct. Biol. 8*, 810−817.
5. Pray, T. R., Reiling, K. K., Demirjian, B. G., and Craik, C. S. (2002) *Biochemistry 41*, 1474−1482.
6. Tong, L., Qian, C., Massariol, M.-J., Bonneau, P. R., Cordingley, M. G., and Lagace, L. (1996) *Nature 383*, 272−275.
7. Qiu, X., Culp, J. S., DiLella, A. G., Hellmig, B., Hoog, S. S., Janson, C. A., Smith, W. W., and Abdel-Meguid, S. S. (1996) *Nature 383*, 275−279.
8. Shieh, H.-S., Kurumbail, R. G., Stevens, A. M., Stegeman, R. A., Sturman, E. J., Pak, J. Y., Wittwer, A. J., Palmier, M. O., Wiegand, R. C., Holwerda, B. C., and Stallings, W. C. (1996) *Nature 383*, 279−282.
9. Chen, P., Tsuge, H., Almassy, R. J., Gribskov, C. L., Katoh, S., Vanderpool, D. L., Margosiak, S. A., Pinko, C., Matthews, D. A., and Kan, C.-C. (1996) *Cell 86*, 835−843.
10. Batra, R., Khayat, R., and Tong, L. (2001) *Protein Pept. Lett. 8*, 333−342.
11. Ogilvie, W., Bailey, M., Poupart, M.-A., Abraham, A., Bhavsar, A., Bonneau, P., Bordeleau, J., Bousquet, Y., Chabot, C., Duceppe, J.-S., Fazal, G., Goulet, S., Grand-Maitre, C., Guse, I., Halmos, T., Lavallee, P., Leach, M., Malenfant, E., O'Meara, J., Plante, R., Plouffe, C., Poirier, M., Soucy, F., Yoakim, C., and Deziel, R. (1997) *J. Med. Chem. 40*, 4113−4135.
12. Tong, L., Qian, C., Massariol, M.-J., Deziel, R., Yoakim, C., and Lagace, L. (1998) *Nature Struct. Biol. 5*, 819−826.

13. Bonneau, P. R., Grant-Maitre, C., Greenwood, D. J., Lagace, L., LaPlante, S. R., Massariol, M.-J., Ogilvie, W. W., O'Meara, J. O., and Kawai, S. H. (1997) *Biochemistry 36*, 12644−12652.
14. LaPlante, S. R., Bonneau, P. R., Aubry, N., Cameron, D. R., Deziel, R., Grand-Maitre, C., Plouffe, C., Tong, L., and Kawai, S. H. (1999) *J. Am. Chem. Soc. 121*, 2974−2986.
15. Pinko, C., Margosiak, S. A., Vanderpool, D., Gutowski, J. C., Condon, B., and Kan, C.-C. (1995) *J. Biol. Chem. 270*, 23634−23640.
16. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
17. Jogl, G., Tao, X., Xu, Y., and Tong, L. (2001) *Acta Crystallogr. D57*, 1127−1134.
18. CCP4. (1994) *Acta Crystallogr. D50*, 760−763.
19. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
20. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
21. Holskin, B. P., Bukhtiyarova, M., Dunn, B. M., Baur, P., DeChastonay, J., and Pennington, M. W. (1995) *Anal. Biochem. 227*, 148−155.
22. Khayat, R., Batra, R., Massariol, M.-J., Lagace, L., and Tong, L. (2001) *Biochemistry 40*, 6344−6351.
23. Lu, W., Qasim, M. A., Laskowski, M., Jr., and Kent, S. B. H. (1997) *Biochemistry 36*, 673−679.
24. Hoog, S. S., Smith, W. W., Qiu, X., Janson, C. A., Hellmig, B., McQueney, M. S., O'Donnell, K., O'Shannessy, D., DiLella, A. G., Debouck, C., and Abdel-Meguid, S. S. (1997) *Biochemistry 36*, 14023−14029.
25. Liang, P.-H., Brun, K. A., Feild, J. A., O'Donnell, K., Doyle, M. L., Green, S. M., Baker, A. E., Blackburn, M. N., and Abdel-Meguid, S. S. (1998) *Biochemistry 37*, 5923−5929.
26. Carson, M. (1987) *J. Mol. Graphics 5*, 103−106.
27. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.

BI027045S