Human Immunodeficiency Virus Protease Ligand Specificity Conferred by Residues Outside of the Active Site Cavity[†,‡]

Susan S. Hoog,[§] Eric M. Towler,[||] Baoguang Zhao,[§] Michael L. Doyle,[§] Christine Debouck,[||] and Sherin S. Abdel-Meguid*,[§]

Departments of Macromolecular Sciences and Molecular Genetics, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, Pennsylvania 19406

Received January 24, 1996[⊗]

ABSTRACT: To gain greater understanding of the structural basis of human immunodeficiency virus (HIV) protease ligand specificity, we have crystallized and determined the structures of the HIV-1 protease (Val32Ile, Ile47Val, Val82Ile) triple mutant and simian immunodeficiency virus (SIV) protease in complex with SB203386, a tripeptide analogue inhibitor containing a C-terminal imidazole substituent as an amide bond isostere. SB203386 is a potent inhibitor of HIV-1 protease (K_i = 18 nM) but shows decreased inhibition of the HIV-1 protease (Val32Ile, Ile47Val, Val82Ile) triple mutant (K_i = 112 nM) and SIV protease (K_i = 960 nM). Although SB203386 binds in the active site cavity of the triple mutant in a similar fashion to its binding to the wild-type HIV-1 protease [Abdel-Meguid *et al.* (1994) *Biochemistry 33*, 11671], it binds to SIV protease in an unexpected mode showing two inhibitor molecules each binding to half of the active site. Comparison of these two structures and that of the wild-type HIV-1 protease bound to SB203386 reveals that HIV protease ligand specificity is imparted by residues outside of the catalytic pocket, which causes subtle changes in its shape. Furthermore, this work illustrates the importance of structural studies in order to understand the structure−activity relationship (SAR) between related enzymes.

Human immunodeficiency virus (HIV), a member of the Retroviridae family, is the causative agent of acquired immunodeficiency syndrome (AIDS). Simian immunodeficiency viruses (SIV) constitute a group of primate retroviruses that are morphologically and antigenically related to HIVs (Henderson *et al*., 1988). An essential, late step in the life cycle of HIV, SIV, and other retroviruses is the proteolytic maturation of two polyprotein precursors catalyzed by a virally encoded protease (Debouck *et al*., 1987; Meek *et al*., 1990). The HIV and SIV proteases are classified as members of the aspartyl protease family. They differ from the well-known monomeric enzymes of this class, such as pepsin and renin, in that their structure is achieved by the assembly of two 99 amino acid polypeptides into a functional homodimer (Meek *et al*., 1989; Wlodawer *et al*., 1989). Despite sharing roughly 50% sequence identity to HIV type 1 (HIV-1; the most prevalent strain in the United States and Europe) protease (Figure 1), SIV_{Mac} (that isolated from *Macaca mulatta*) protease displays differences in ligand specificity (Tomasselli *et al*., 1990). SIV_{Mac} protease, however, shares 93% homology to HIV type 2 (HIV-2; the most common strain in West Africa) protease (Figure 1) and displays similar ligand specificity.

As a step toward the development of drugs to combat HIV infection, numerous strategies have led to the design of a wide variety of HIV-1 protease inhibitors. One strategy (Abdel-Meguid *et al*., 1994) led to the design of SB203386 (Figure 2), a potent tripeptide analogue inhibitor containing a C-terminal imidazole substituent as an amide bond isostere. The secondary alcohol was designed to engage both of the catalytic aspartate groups with hydrogen bonds, while bulky P1 and P1′ benzyl groups fill the S1 and S1′ subsites. The N-1 and N-3 atoms act as replacements for amide nitrogen and carbonyl oxygen of the peptidyl backbone and resulted in a substantial improvement in oral bioavailability and pharmacokinetic properties. While a highly potent inhibitor of HIV-1 protease (K_i = 18 nM), SB203386 showed substantially less inhibition of SIV (K_i = 960 nM) and HIV-2 (K_i = 1280 nM) proteases (Abdel-Meguid *et al*., 1994).

Although some inhibitors bind to both HIV-1 and SIV or HIV-2 proteases with equal affinity, others differ by as much as 2 orders of magnitude (Tomasselli *et al*., 1990). Nonetheless, the crystal structures of these three enzymes reveal nearly identical tertiary structures (Zhao *et al*., 1993; Tong *et al*., 1993). The HIV-1 protease active site cavity differs from that of SIV and HIV-2 proteases by only three highly conserved substitutions: Val32Ile, Ile47Val, and Val82Ile in each monomer for a total of six. Given the similarity of the structures of these enzymes, it has been difficult to explain the large difference observed in their ligand binding affinities, particularly in the absence of an understanding of the structural basis for their ligand specificity. Such an understanding is deemed more critical with recent reports of the emergence of drug resistance to HIV-1 grown in tissue culture under HIV-1 protease inhibitor selection pressure (Mellors *et al*., 1995). Thus, if rational design is to enable the design of inhibitors that are therapeutically effective against both HIV-1 and HIV-2 proteases, then there must

[†] This work was supported by NIH Grant 1 RO1 GM50579-01.

[‡] The refined coordinates for the SIV protease and the HIV-1 protease triple mutant have been deposited in the Protein Data Bank under the file names 1TCW and 1TCX, respectively.

* To whom correspondence should be addressed.

[§] Department of Macromolecular Sciences.

[||] Department of Molecular Genetics.

[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1996.

S0006-2960(96)00179-1 CCC: $12.00

```
                    5                  10                  15                  20                  25
HIV-1   -  -  I  T  -  -  Q  -  -  L  -  -  I  K  -  G  -  -  L  K  -  A  -  -  -
SIV     P  Q  F  S  L  W  R  R  P  V  V  T  A  H  I  E  G  Q  P  V  E  V  L  L  D
HIV-2   -  -  -  -  -  -  K  -  -  -  -  -  -  Y  -  -  -  -  -  -  -  -  -  -  -
                   30                  35                  40                  45                  50
HIV-1   -  -  -  -  -  T  V  L  E  E  M  S  -  P  G  R  W  K  -  -  M  I  -  -  -
SIV     T  G  A  D  D  S  I  V  T  G  I  E  L  G  P  H  Y  T  P  K  I  V  G  G  I
HIV-2   -  -  -  -  -  -  -  -  A  -  -  -  -  -  N  N  -  S  -  -  -  -  -  -  -
                   55                  60                  65                  70                  75
HIV-1   -  -  -  -  K  V  R  Q  -  D  Q  I  L  -  -  I  C  -  H  K  A  I  -  -  V
SIV     G  G  F  I  N  T  K  E  Y  K  N  V  K  I  E  V  L  G  K  R  I  K  G  T  I
HIV-2   -  -  -  -  -  -  -  -  -  -  -  -  E  -  -  -  -  N  -  K  V  R  A  -  -
                   80                  85                  90                  95
HIV-1   L  V  -  P  -  -  V  -  -  I  -  -  -  -  -  -  Q  I  -  C  T  -  -  F
SIV     M  T  G  D  T  P  I  N  I  F  G  R  N  L  L  T  A  L  G  M  S  L  N  L
HIV-2   -  -  -  -  -  -  -  -  -  -  -  -  -  I  -  -  -  -  -  -  -  -  -  -
```

FIGURE 1: Comparison of the amino acid sequences of the proteases from HIV-1 (III_b strain), SIV (251 strain), and HIV-2 (*rod* strain).

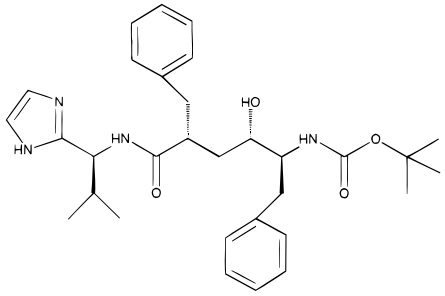

FIGURE 2: Chemical structure of SB203386.

be a structural understanding of the basis for ligand specificity. With this in mind, we have determined the crystal structures of the HIV-1 protease (Val32Ile, Ile47Val, Val82Ile) triple mutant and SIV protease in complex with SB203386 and compared both structures to our previously reported structure of HIV-1 protease bound to SB203386 (Abdel-Meguid *et al*., 1994).

EXPERIMENTAL PROCEDURES

Enzyme Preparation. Recombinant Val32Ile, Ile47Val, Val82Ile HIV-1 protease (triple mutant) was prepared by introduction of amino acid substitutions in the HIV-1 III_b protease coding region by oligonucleotide site-directed mutagenesis and verified by dideoxy sequencing. Triple mutant and SIV proteases were both purified to apparent homogeneity as previously described (Grant *et al*., 1991). The protease/inhibitor complexes were prepared by adding a 2-fold (5-fold for SIV protease) molar excess of inhibitor per protease dimer and concentrating to 3 mg/mL in a buffer containing 50 mM sodium acetate, pH 5.0, 2 mM EDTA, 5 mM DTT, 5% DMSO, and 180 mM NaCl (150 mM NaCl for SIV protease). The synthesis of the inhibitor, SB203386, has been previously described (Abdel-Meguid *et al*., 1994).

Peptidolysis reactions were conducted at 37 °C in a buffer composed of 50 mM 2-(*N*-morpholino)ethanesulfonic acid, pH 6.0, 1 mM EDTA, 1 mM DTT, 0.2 M NaCl, 0.1% Triton X-100, and 10% dimethyl sulfoxide. The reaction mixtures contained 1.25−25 mM concentrations of the substrate Ac-RASQNYPVV-NH_2 and 60−150 ng of HIV-1 or triple mutant protease and were quantified by HPLC. Kinetic constants for substrates were determined by fitting the data to the Michaelis−Menten equation using the nonlinear regression package in Kaleidagraph (Abelbeck Software). K_m values were independently determined for both the HIV-1 protease and the triple mutant protease and found to be 10.2 and 8.9 mM, respectively. The K_i for the HIV-1 protease wild type calculated in this study (18 nM) agreed with the previously determined value reported by Abdel-Meguid *et al.* (1994). Determination of K_i was as previously described by Grant *et al*. (1991).

Microcalorimetry. Microcalorimetric titrations were performed with a Microcal, Inc. (Amherst, MA), MC-2 isothermal titration microcalorimeter. Protease was dialyzed for 2 h at 4 °C against 20 mM sodium acetate, pH 5.0, 150 mM sodium chloride, 2 mM EDTA, and 3% DMSO. Inhibitor solutions were prepared by diluting a 3.0 mg/mL DMSO stock solution (5.6 mM) into buffer to a final concentration of 168 μM inhibitor and 3% DMSO. The pH was adjusted to equal that of the dialyzed protease solution by adding small amounts of 0.1 M NaOH. The samples were degassed at 23 °C with gentle swirling under vacuum just prior to loading into the microcalorimeter. Nonlinear least squares analysis was based on the single-site fitting algorithm from Microcal Origin software (Microcal, Inc., 1993). All data were well described individually by a binding mode with a single equilibrium constant and binding enthalpy change as adjustable parameters. Because the ionization enthalpy change of acetate is so small (−0.08 kcal/mol), the binding enthalpy change reported herein need not be corrected for proton uptake or release by the buffer and is regarded as an intrinsic enthalpy change (Stoll & Blanchard, 1990).

Crystallization. The triple mutant protease crystals were grown out of 20% saturated ammonium sulfate in 0.2 M sodium acetate, pH 5, while SIV protease crystals were grown out of 25−30% NaCl in 100 mM sodium acetate, pH 4.2, by vapor diffusion in hanging drops. Crystals of both complexes grew at room temperature to full size within 1 week.

Data Collection. Diffraction data were collected on a Siemens four-circle area detector and reduced using XENGEN (Howard *et al*., 1987). The space group of the triple mutant was determined to be $P2_12_12$ with cell dimensions $a = 59.6$ Å, $b = 87.5$ Å, and $c = 47.0$ Å and one protease dimer per asymmetric unit. Data extending to 2.3 Å were 85% complete with $R_{sym} = 0.060$ ($R_{sym} = \Sigma\Sigma|I_{i(h)} - \bar{I}_{(h)}|/\Sigma I_{i(h)}$, where $I_{i(h)}$ are the intensities of multiple measurements of reflection h and $\bar{I}_{(h)}$ is their mean). The space group of the SIV protease was $P2_1$ with cell dimensions $a = 56.1$ Å, $b = 41.5$ Å, $c = 53.1$ Å, and $\beta = 99.6°$ with one dimer per asymmetric unit. Data were 93% complete to 2.4 Å with $R_{sym} = 0.074$.

Structure Determination and Refinement. The HIV-1 protease triple mutant/SB203386 complex was determined

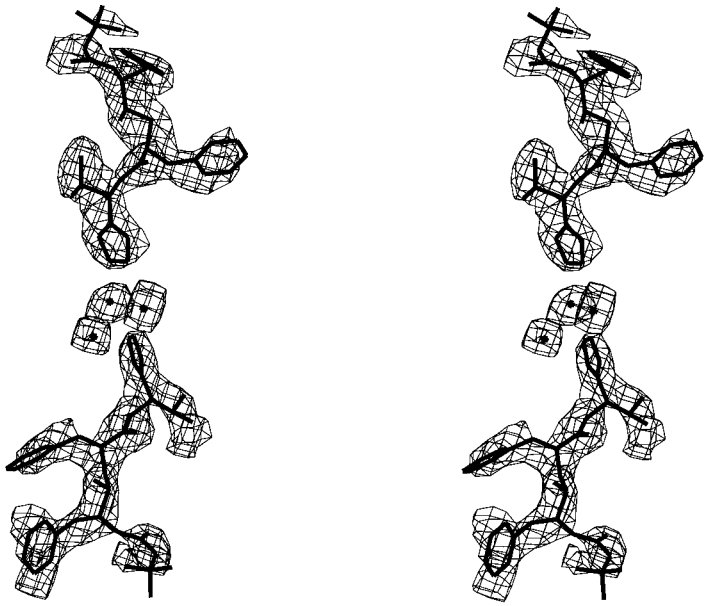

FIGURE 3: $F_o - F_c$ omit map contoured at 2σ showing electron density of SB203386 and water molecules in the active site of the SIV protease/SB203386 complex. The two inhibitors were omitted from the structure factor calculation.

using phases from the isomorphous structure 1HTG (Jhoti *et al*., 1994) to an initial *R*-factor of 0.29 ($R = \Sigma||F_o| - |F_c||/\Sigma |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively). Following several rounds of model building and X-PLOR (Brunger, 1992) refinement using data greater than $2\sigma(F_o)$ in the resolution range 8.0−2.3 Å, and aided by calculation of annealed omit maps, the *R*-factor dropped to 0.180. The final rms deviations from ideal values for bond lengths and bond angles are 0.016 Å and 3.1°, respectively. The mean *B*-factors for the protein and inhibitor were 19.7 and 17.9 Å^2, respectively.

The structure of the SIV protease/inhibitor complex was determined by molecular replacement using X-PLOR program package (Brunger, 1992). Rotation functions were calculated using diffraction data from 10 to 4 Å with a radius of integration of 25 Å for the self-rotation function and from 6 to 3 Å with a vector length of 35 Å for the cross-rotation function. Results from the latter revealed two significant peaks (θ_1, θ_2, θ_3 are 56.5°, 63.0°, 56.5° and 84.9°, 45.0°, 196.9°) related by a 2-fold axis that are consistent with the result from the self-rotation function. The translation function using data from 10 to 3 Å resolution for reflections greater than 2σ gave an *R*-factor of 0.40. The translation peak corresponded to fractional coordinates of 0.441, 0.000, and 0.121, with a height of 19σ. Using data from 7 to 2.4 Å resolution with $|F_o| > 3\sigma|F_o|$, the *R*-factor was reduced to 0.26 following a single round of simulated annealing. Difference electron density maps calculated using refined phases showed clear interpretable density for the protease; however, density in the active site cavity was broken and ambiguous.

Inspection of annealed omit maps at $\pm 2\sigma$ and 3σ revealed density for several water molecules in the center and continuous density of length consistent with that of the inhibitor at both sides of the binding site extending outward from the P2 and P2′ sites (Figure 3). At this point we assumed that SIV protease binds two inhibitors per protease dimer or that it binds one, but owing to the symmetric nature of the enzyme we see the inhibitor twice, each with one-half occupancy. The latter scenario was less likely due to

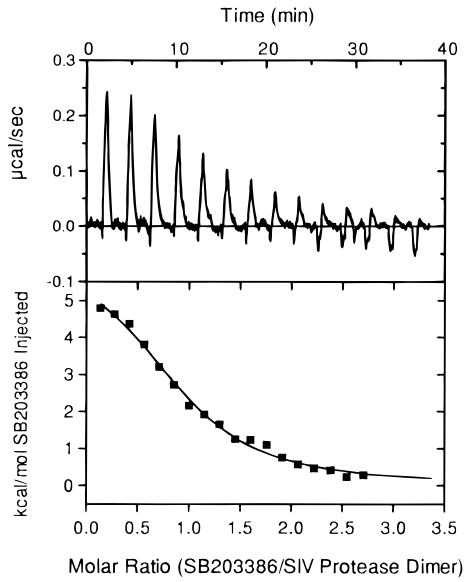

FIGURE 4: (Top portion) Power versus time plot of the raw data resulting from titration of 10.6 μM SIV protease with several 12 μL injections of 168 μM SB203386. The amount of heat absorbed upon each injection of SB203386 is proportional to the area under each "spike". (Bottom portion) Observed heat normalized to moles of SB203386 injected versus molar ratio of SB203386 added per SIV protease dimer. Nonlinear regression analysis gives a $K_d = 3$ μM, $\Delta H = 6.4$ kcal/mol of SB203386, and an apparent stoichiometry of 0.93 inhibitor molecule per protease dimer. Conditions were 27 °C, pH 5.0, 20 mM sodium acetate, 150 mM NaCl, 2 mM EDTA, and 3% DMSO.

the nonsymmetric nature of the density on both sides of the active site. Alternatively, SIV protease could bind one inhibitor molecule on one side of the active site with the other being occupied by a peptide of host cell origin as that reported by Jhoti *et al.* (1994). To resolve these issues, titration calorimetry experiments were performed in order to gain a stoichiometric ratio of the inhibitor to the enzyme for this complex. Figure 4 shows calorimetry data for the titration of SIV protease with SB203386. Analysis of the data by nonlinear regression gave a K_d of 3 μM, which is in good agreement with the K_i value of 1 μM. The apparent

Table 1: Inhibitor Binding Equilibrium Parameters of SIV and Mutant Protease Determined by Titration Microcalorimetry

protease	K_d (nM)	ΔG (kcal/mol inhbitor)	ΔH (kcal/mol inhibitor)	$-T\Delta S$ (kcal/mol inhibitor)	apparent stoichiometry (inhibitor/dimer)
SIV	3000	-7.5 ± 0.3	$+6.4 \pm 0.6$	-14.0 ± 0.7	0.93 ± 0.10
triple mutant	150	-9.3 ± 0.3	$+2.5 \pm 0.3$	-11.8 ± 0.4	1.05 ± 0.10

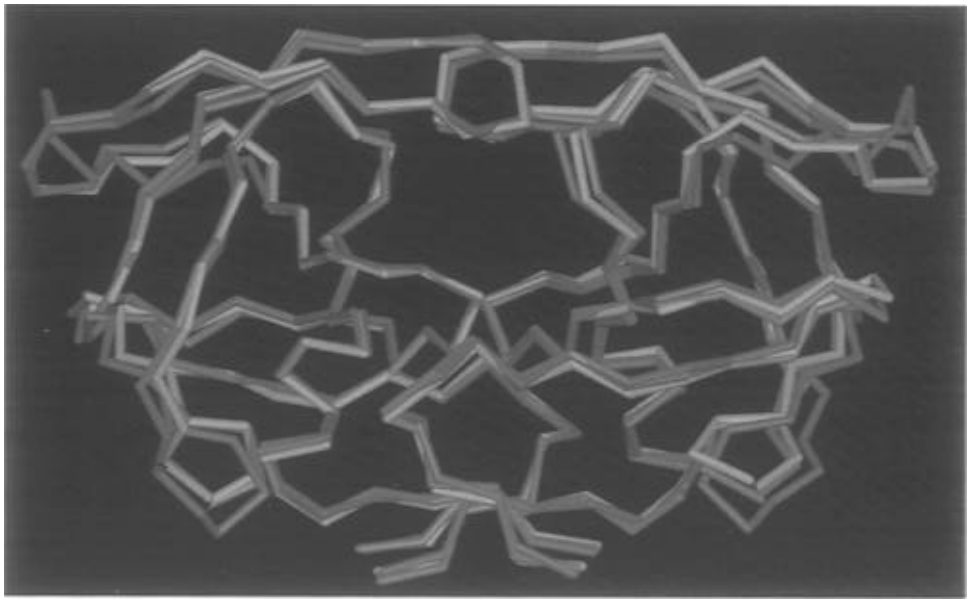

FIGURE 5: Superposition of the α-carbon atoms of HIV-1 protease (yellow), SIV protease (purple), and HIV-1 protease triple mutant (green) structures, each derived from its structure with SB203386. Despite the absence of an inhibitor in the center of the binding site of SIV protease, the flaps are in the closed conformation.

stoichiometry of the reaction was determined to be 0.93 inhibitor per protease dimer, a value which is supportive of a 1:1 binding stoichiometry (Table 1). However, the calorimetry data do not disprove the possibility of having two inhibitor molecules unequally bound with different K_d's, one with the measured K_d and one with a much weaker binding constant, since the calorimetry data can only detect K_d's tighter than about 30 μM. To explore the possibility that a peptide of host cell origin is bound to half of the active site, a Ser-Leu-Ile-Gln peptide was fit into one side of the binding site, while the inhibitor was fit on the other side. The Ser-Leu-Ile-Gln peptide appeared to fit well in the electron density, satisfied several hydrogen bonds, and made several favorable contacts. However, in order to confirm or better identify this fragment, N-terminal sequencing on dissolved crystals of the complex was performed but resulted in no sequence other than that of the protease. Thus, despite the fortuitous fit to the density, the origin of such a peptide could not be explained. It is not part of a known substrate of SIV (or HIV) protease and is not found in its sequence. Additionally, its origin as an artifact of purification is unlikely because SIV protease was purified under identical procedures as both the HIV-1 and triple mutant proteases, in which no similar peptide was found.

The most likely explanation of these data is that there are two inhibitors bound simultaneously to the SIV protease dimer, each occupying one-half of the active site starting at P2 and P2′, but the inhibitors bind much weaker to one side than the other. This is evidenced by the lack of symmetry in the electron density on both sides of the active site and by differences in the temperature factor of each inhibitor (the mean *B*-factors for the 39 atoms of each of the two inhibitors are 22.4 and 27.0 Å^2, respectively). Cycles of positional refinement together with grouped *B*-factor refinement were performed, resulting in a final *R*-factor of 0.197 with root mean square deviation in bond length of 0.018 Å and 3.7°.

RESULTS AND DISCUSSION

Structure of Val32Ile, Ile47Val, Val82Ile Triple Mutant/SB203386 Complex. The triple mutant contains the binding pocket of SIV (or HIV-2) while retaining all other structural features of HIV-1 protease. The crystal structure of HIV-1 protease/SB203386 has been previously determined by Abdel-Meguid *et al.* (1994). As seen in Figure 5, the overall backbone structures of the native and mutant enzyme complexes are basically identical, with an rms deviation of 0.57 Å for all 198 α-carbon pairs, where the largest deviations are in external loops in which changes in conformation are most likely the result of crystal contacts.

Figure 6a shows a superposition of the active sites of the two enzymes. The conformation of the inhibitors in both structures is essentially identical with small changes in the positions of the P1 and P1′ phenyl groups. All hydrogen-bonding interactions seen in the HIV-1 structure are maintained in the mutant structure with very minor differences in hydrogen bond length. Specifically, key interactions of the imidazole allowing it to mimic the amide bond between P3′ and P2′ remain the same in the triple mutant structure as in the HIV-1 structure (Abdel-Meguid *et al.*, 1994). However, the triple mutant shows 6-fold reduction in binding affinity (K_i = 112 nM) for SB203386 when compared to the HIV-1 enzyme (K_i = 18 nM). Figure 6b shows the S1/S1′ subsite interactions with the P1/P1′ benzyl groups of the inhibitor. In the wild-type enzyme, both Cγ methyl groups of Val82′ make van der Waals contact with the benzyl group

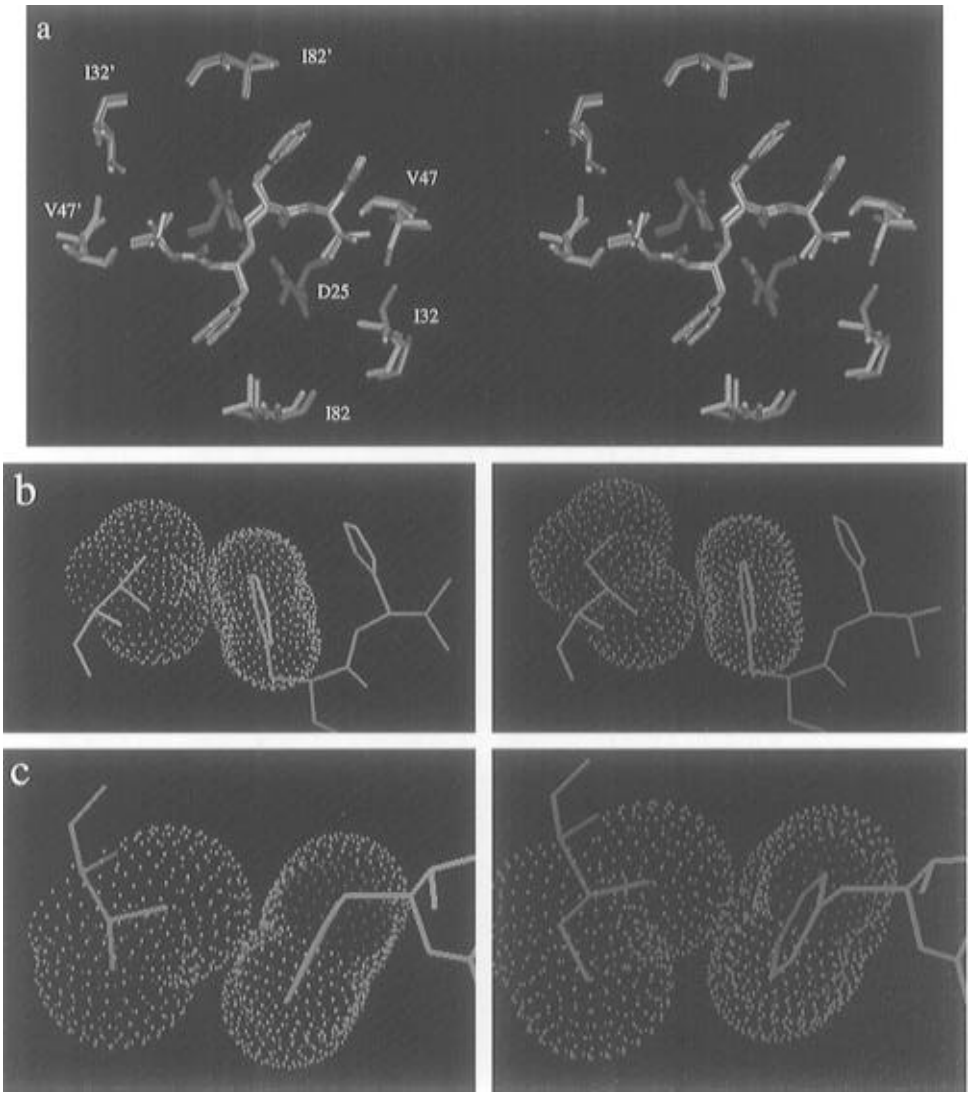

FIGURE 6: (a) Stereoview of residues comprising the active site cavity of the triple mutant (green) and native HIV-1 protease (yellow) showing the bound SB203386 inhibitor. The two catalytic aspartates (Asp25 and Asp25′) are in red. (b) Close-up look at the van der Waals contacts between the inhibitor and residue 82′. In the native structure (yellow) the Cγ1 and Cγ2 methyl groups are in van der Waals contact with the P1′ benzyl group. The mutant structure (green) shows that only the Cγ2 methyl of Ile82′ is in contact with the P1 benzyl, while the Cγ1 and Cδ1 methyls are pointing away from the ring. (c) Close-up look at the van der Waals contacts between the inhibitor and residue 82. Compared to the wild type (yellow), the mutant enzyme (green) maintains van der Waals contact with the inhibitor through reorientation of the Cγ methyls of Ile82 and a small movement of the P1 benzyl.

of the inhibitor. This contact is reduced to a single Cγ2 methyl group in the mutant structure because of the steric constraint of the Ile side chain. Figure 6c shows interactions of the P1 benzyl group with Ile82, where the P1 side chain in the mutant enzyme compensates for the presence of the larger Ile by moving about 1 Å relative to that in the wild type to maintain a single van der Waals contact with the I82Cγ methyl group. Compensatory changes in a HIV-1 protease mutant have been seen before in the Val82Ala/A-77003 complex structure (Baldwin *et al*., 1995), where unexpected shifts at Ala82, combined with a slight reorientation of the P3 substituent, resulted in interaction surfaces that were similar in appearance to those observed in the wild-type enzyme. The remaining mutations (Val32Ile, Ile47Val, Val32′Ile, Ile47′Val) in the triple mutant resulted in little change in the binding of the valinyl or *t*-Boc groups of the inhibitor. The loss of a single methyl group in the Val82Ala/A-77003 structure has been postulated to be the basis of a 4-fold reduction in inhibitor affinity with the Val82Ala mutant binding to A-77003 (Baldwin *et al*., 1995). Likewise, the 6-fold reduction in SB203386 binding affinity must be due to the cumulative small changes seen around positions 82 and 82′.

Structure of SIV Protease/SB203386 Complex. The overall three-dimensional structure of the SIV protease/SB203386 inhibitor complex is similar to previously published structures of inhibitor-bound SIV protease (Zhao *et al*., 1993; Rose *et al*., 1993). Superposition of α-carbon atoms of the present structure to that of Zhao *et al*. (1993) resulted in a rms deviation of 0.54 Å.

As discussed under Experimental Procedures, the binding mode of SB203386 seen here was unexpected despite the relatively poor binding of this inhibitor to SIV protease (K_i 960 nM). The conformation of the inhibitor on either side of the enzyme is not completely identical. An overlap of the two conformations reveals the greatest deviation at the C-terminal end beyond the second phenyl group, where for one side the electron density is largely disordered. On both sides, the inhibitor appears to be held in the active site by the same three hydrogen bonds; two are to the amide nitrogen

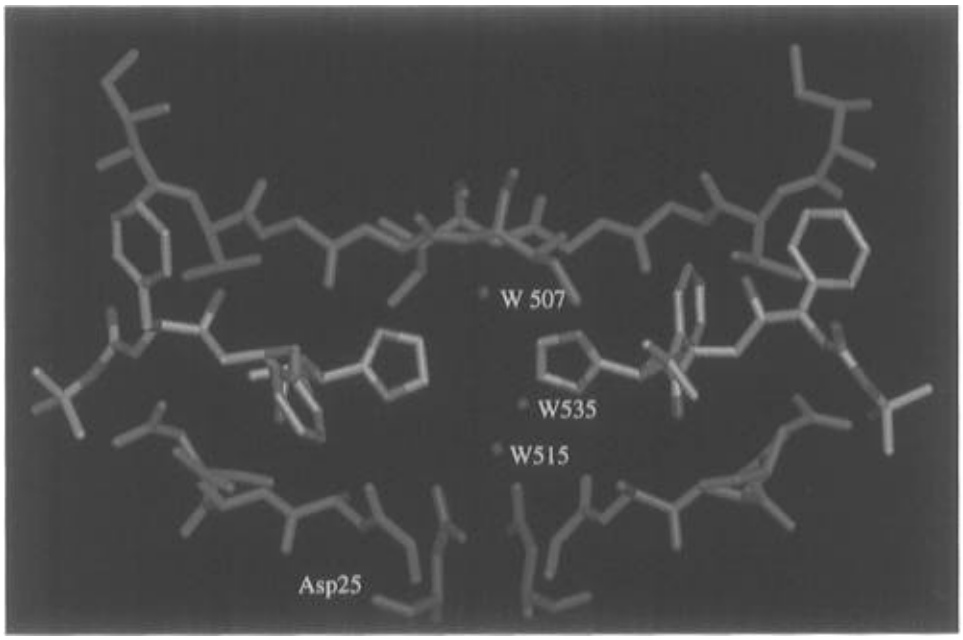

FIGURE 7: View of the binding cleft of the SIV protease/SB203386 complex containing the two inhibitors. The residues surrounding the active site are shown in purple, except for Asp25 and Asp25′, which are shown in red. Solvent molecules are red spheres.

and carbonyl oxygen of Gly48 (and Gly48′) and the third is to the amide nitrogen of Asp29 (and Asp29′). The *t*-Boc group on both sides extends outside the binding site into the solvent and does not form any interactions. There are no interactions between the inhibitor and the catalytic aspartate residues (25 and 25′) or the central water between Ile50 and Ile50′.

The calorimetry data show that the binding enthalpy change is +6.4 kcal/mol inhibitor, a value which is considerably unfavorable (Table 1). Thus the binding reaction must be driven by a favorable change in entropy. Although this may seem counterintuitive, entropy-driven binding reactions of small hydrophobic compounds with proteins have been seen before (Weber *et al.*, 1994). Table 1 shows a less positive ΔH and a more favorable reaction overall (ΔG) for the triple mutant as compared to SIV protease. This is a likely indication of the larger number of hydrogen bonds in the mutant/SB203386 complex as compared to the SIV protease inhibitor complex.

The present structure shows the flaps closed over the active site despite an absence of an inhibitor in the center of the binding site. There are, however, three central water molecules. Figure 7 shows W507 bridging the flaps closed down over the active site, maintaining hydrogen bonds of 2.8 and 2.7 Å to Ile50′ and Ile50, respectively. W515 is positioned between the catalytic aspartates as if in an apo structure, maintaining hydrogen bonds to the four carboxylate oxygens (3.1 and 3.3 Å to Asp25, 3.3 and 3.7 Å to Asp25′). Nucleophilic attack on the carbonyl carbon of the scissile bond of a bound substrate by this water molecule leads to the tetrahedral transition state intermediate mimicked by the hydroxyl of many tight binding hydroxyethylene isostere inhibitors. A water molecule between the catalytic aspartates in a closed apo structure has been previously seen by Wilderspin and Sugrue (1994; PDB code 1SIP). In addition, there is a third water molecule, W535, in the binding site that hydrogen bonds to W515, the carbonyl oxygen of G27′, and an imidazole nitrogen of the inhibitor.

Comparison with HIV-1 Protease/SB203386 Complex. Pairwise structural comparison of complexes of HIV-1 and SIV proteases (Zhao *et al.*, 1993) and HIV-1 and HIV-2 proteases (Mulichak *et al.*, 1993; Tong *et al.*, 1993; Priestle *et al.*, 1995) containing the same inhibitor has been previously reported. However, in most of these reports the inhibitor binds with affinities that differ by less than an order of magnitude to both members of the pair, whereas SB203386 shows a 54-fold difference in binding between HIV-1 and SIV proteases (Abdel-Meguid *et al.*, 1994), with little conformational difference in the enzyme structures (Figure 5). The rms deviation between all α-carbon atoms of the two enzymes is 1.10 Å. Aside from the kinetic data, there is little in the overall tertiary structure to allow prediction of the present binding mode. The largest differences between these two structures are in residues 33−44 and 61−75. These residues are part of loops on the surface of the protease and are distant from the active site cavities. It is difficult to conclude that these structural differences are responsible for the different binding modes of the inhibitor.

Implications. HIV-1 and SIV (or HIV-2) proteases differ in their substrate specificity (Tomasselli *et al.*, 1990). It has been proposed that HIV-2 (or SIV) protease may have a preference for ligands containing smaller hydrophobic residues at P1 and P1′ positions, as a consequence of the presence of isoleucine in HIV-2 and SIV proteases versus valine at residue 82 in the HIV-1 enzyme (Tomasselli *et al.*, 1990). This preference was postulated to be the reason for the 6-fold reduction in binding of CGP53820 to HIV-2 protease (Priestle *et al.*, 1995). Our triple mutant structure further supports this conclusion. As we have shown above, the 6-fold loss of SB203386 binding affinity between the wild-type and triple mutant is a consequence of conformational changes around positions 82 and 82′. The above reasoning, however, cannot adequately explain the mode of binding of SB203386 to SIV protease. We, as well as Zhao *et al.* (1993), observed that the S1 and S1′ pockets do not necessarily preclude the accommodation of large side chains

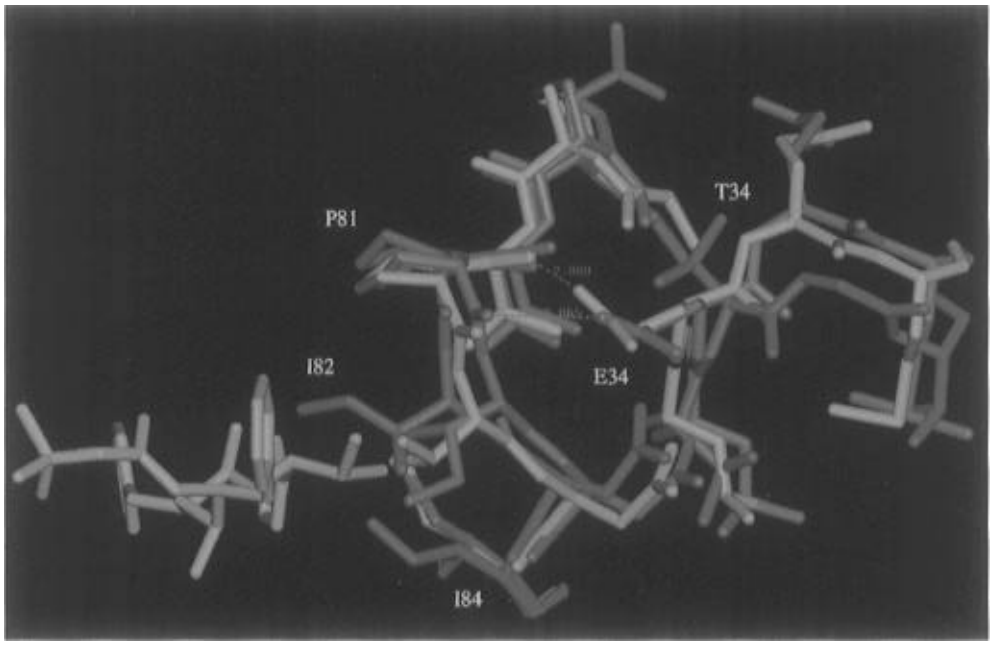

FIGURE 8: View of aligned residues 78−84 and 33−36 in the S1 subsite in all three structures. The inhibitor of the HIV-1 complex structure is shown to illustrate the smaller binding cavity in the SIV structure around Ile82. Color code: HIV-1 complex, yellow; mutant complex, green; SIV complex, purple.

at P1 and P1′ since the δ-carbon atom of Ile82 assumes a conformation that minimizes its protrusion in the active site. Thus, an explanation for the observed mode of binding lies elsewhere.

The structures presented herein suggest that differences in ligand binding specificity between HIV-1 and SIV proteases are conferred by residues outside of the binding site. As seen from the structure of the triple mutant, the changes in the active site cavity account for only 6-fold reduction in affinity of SB203386. Thus, the additional 9-fold reduction seen between the triple mutant and SIV protease must be due to changes resulting from amino acid residues distal to the active site cavity that preclude proper positioning of the inhibitor in the binding site. Given that there is a total of 49 amino acids that are different between SIV and HIV-1 protease, any one or a combination of these residues could be responsible for the additional loss of binding affinity.

Main chain flexibility of residues 80−84 has been implicated as responsible for differences in ligand binding between HIV-1 and HIV-2 proteases (Priestle *et al*., 1995; Mulichak *et al*., 1993). Analysis of the average temperature factor of the wild-type, mutant, and SIV protease structures containing SB203386 in that region reveals no significant differences. The average *B*-factor of backbone atoms for residues 80−84 and 80′−84′ as a fraction of the overall mean *B*-factor is 0.75, 0.66, and 0.51 for the HIV-1, triple mutant, and SIV protein structures, respectively, indicating that this region is relatively stable in all three structures. However, analysis of these three structures shows that in that region the backbone of the SIV structure extends further toward the active site cavity by more than 1 Å, causing a significant change in the shape of the active site. The distance between the β-carbon atoms of 82 and 82′ is 16.8 Å in the SIV structure, while this same distance is 19.1 and 18.7 Å in the triple mutant and the wild-type HIV-1 protease structures, respectively. This is a consequence of a hydrogen bond that is present in the wild-type structure between the side chain carboxyl of Glu34 and the backbone carbonyl of Thr80 (2.9 Å), which is not possible in SIV where residue 34 is a threonine and is therefore unable to maintain that bond. A similar hydrogen bond exists in the mutant structure between the carboxylate of Glu34 and the backbone carbonyl of Pro81 (3.1 Å), as can be seen in Figure 8.

Our data clearly show that ligand specificity is not only imparted by residues comprising the active site cavity but also influenced by global perturbations of the overall structure that causes subtle changes in the shape (and probably the size) of the active site cavity. These perturbations are a consequence of structural changes caused by amino acids distal to the catalytic pocket. Furthermore, this work illustrates the importance of structural studies in order to understand the structure−activity relationship (SAR) between related enzymes. Inspection of SAR values alone cannot possibly predict differences in modes of binding, particularly when differences in binding affinities are significant.

ACKNOWLEDGMENT

The authors thank Jeff Stebbins, Evon Winborne, Dean McNulty, Seth Fisher, and Kannen Ramachandren for technical assistance. They also thank the two anonymous reviewers for their perceptive comments.

REFERENCES

Abdel-Meguid, S. S., Metcalf, B. W., Carr, T. J., Demarsh, P., DesJarlais, R. L., Fisher, S., Green, D. W., Ivanoff, L., Lambert, D. M., Murthy, K. H. M., Petteway, S. R., Pitts, W. J., Tomaszek, T. A., Winborne, E., Zhao, B., Dreyer, G. B., & Meek, T. D. (1994) *Biochemistry 33*, 11671−11677.

Baldwin, E. T., Bhat, T. N., Liu, B., Pattabiraman, N., & Erickson, J. W. (1995) *Nat. Struct. Biol. 2*, 244−249.

Brünger, A. T. (1992) *X-PLOR Manual, Version 3.1*, Yale University Press, New Haven, CT.

Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W., & Rosenberg M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8903−8906.

Grant, S. K., Deckman, I. C., Minnich, M. D., Culp, J., Franklin, S., Dreyer, G. B., Tomaszek, T. A., Jr., Debouck, C., & Meek, T. D. (1991) *Biochemistry 30*, 8424−8434.

Henderson, L. E., Beneviste, R. E., Sowder, R., Copeland, T., D., Schultz, A. M., & Oroszlan, S. (1988) *J. Virol. 62*, 2587−2595.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Puolos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383−387.

Jhoti, H., Singh, O. M. P., Weir, M. P., Cooke, R., Murray-Rust, P., & Wonacott, A. (1994) *Biochemistry 33*, 8417−8427.

Meek, T. D., Dayton, B. D., Metcalf, B. W., Dreyer, G. B., Strickler, J. E., Gorniak, J. G., Rosenberg, M., Moore, M. L., Magaard, V. W., & Debouck, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1841−1845.

Meek, T. D., Lambert, D. M., Dreyer, G. B., Carr, T. J., Tomaszek, T. A., Jr., Moore, M. L., Strickler, J. E., Debouck, C., Hyland, L. J., Matthews, T. J., Metcalf, B. W., & Petteway, S. R. (1990) *Nature 343*, 90−92.

Mellors, J. W., Larder, B. A., & Schinazi, R. F. (1995) *Int. Antiviral News 3*, 8−13.

Microcal, Inc. (1993) *ITC Data Analysis in Origin, Version 2.9*, Northampton, MA.

Mulichak, A. M., Hui, J. O., Tomasselli, A. G., Heinrikson, R. L., Curry, K. A., Tomich, C.-S., Thaisrivongs, S., Sawyer, T. K., & Watenpaugh, K. D. (1993) *J. Biol. Chem. 268*, 13103−13109.

Priestle, J. P., Fässler, A., Rösel, J., Tintelnot-Blomley, M., Strop, P., & Grütter, M. G. (1995) *Structure 3*, 381−389.

Rose, R. B., Rosé, J. R., Salto, R., Craik, C. S., & Stroud, R. M. (1993) *Biochemistry 32*, 12498−12507.

Stoll, V. S., & Blanchard, J. S. (1990) *Methods Enzymol. 182*, 24−38.

Tomasselli, A. G., Hui, J. O., Sawyer, T. K., Staples, D. J., Bannow, C., Reardon, I. M., Howe, W. J., DeCamp, D. L., Craik, C. S., & Heinrikson, R. L. (1990) *J. Biol. Chem. 265*, 14675−14683.

Tong, L., Pav, S., Pargellis, C., Do, F., Lamarre, D., & Anderson, P. C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8387−8391.

Weber, P. C., Pantoliano, M. W., Simmons, D. M., & Salemme, F. R. (1994) *J. Am. Chem. Soc. 116*, 2717−2724.

Wilderspin, A. F., & Sugrue, R. J. (1994) *J. Mol. Biol. 239*, 97−103.

Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., & Kent, S. B. H. (1989) *Science 245*, 616−621.

Zhao, B., Winborne, E., Minnich, M. D., Culp, J. S., Debouck, C., & Abdel-Meguid, S. S. (1993) *Biochemistry 32*, 13054−13060.

BI960179J