# Structural Role of the 30's Loop in Determining the Ligand Specificity of the Human Immunodeficiency Virus Protease[†,‡]

Manal A. Swairjo,[§] Eric M. Towler,[||,⊥] Christine Debouck,[||] and Sherin S. Abdel-Meguid*[,§]

*Departments of Structural Biology and Molecular Biology, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, Pennsylvania 19406*

*Received April 8, 1998; Revised Manuscript Received June 9, 1998*

ABSTRACT: The structural basis of ligand specificity in human immunodeficiency virus (HIV) protease has been investigated by determining the crystal structures of three chimeric HIV proteases complexed with SB203386, a tripeptide analogue inhibitor. The chimeras are constructed by substituting amino acid residues in the HIV type 1 (HIV-1) protease sequence with the corresponding residues from HIV type 2 (HIV-2) in the region spanning residues 31−37 and in the active site cavity. SB203386 is a potent inhibitor of HIV-1 protease ($K_i$ = 18 nM) but has a decreased affinity for HIV-2 protease ($K_i$ = 1280 nM). Crystallographic analysis reveals that substitution of residues 31−37 (30's loop) with those of HIV-2 protease renders the chimera similar to HIV-2 protease in both the inhibitor binding affinity and mode of binding (two inhibitor molecules per protease dimer). However, further substitution of active site residues 47 and 82 has a compensatory effect which restores the HIV-1-like inhibitor binding mode (one inhibitor molecule in the center of the protease active site) and partially restores the affinity. Comparison of the three chimeric protease structures with those of HIV-1 and SIV proteases complexed with the same inhibitor reveals structural changes in the flap regions and the 80's loops, as well as changes in the dimensions of the active site cavity. The study provides structural evidence of the role of the 30's loop in conferring inhibitor specificity in HIV proteases.

Human immunodeficiency virus (HIV)[1] is the etiological agent of acquired immune deficiency syndrome (AIDS). An essential late step in the life cycle of HIV, as well as the related simian immunodeficiency virus (SIV), is the post-translational processing of the polyprotein precursors $Pr^{GAG}$ and $Pr^{GAG-POL}$ into mature structural proteins by a virally encoded protease (*1*−*3*). The HIV and SIV retroviral proteases belong to the ubiquitous aspartyl protease family. However, their structures differ from those of monomeric aspartyl proteases in that they are formed by the assembly of two 99-amino acid polypeptides into a functional homodimer (*4, 5*). Their essential role in the viral life cycle and their unique functional and structural properties that distinguish them from their cellular counterparts have made them an attractive target for the treatment of AIDS. Efforts to develop inhibitors of HIV protease have yielded a class of compounds that suppress HIV replication in vitro and in vivo. Several such inhibitors have been approved by the U. S. Food and Drug Administration and have been successfully used in human patients as anti-AIDS therapies (for reviews, see refs *6* and *7*).

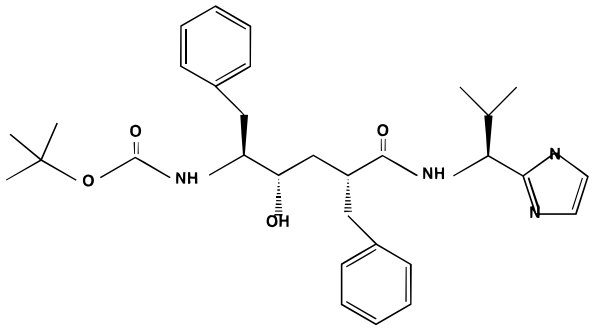


FIGURE 1: Chemical structure of SB203386.

Numerous structure-based design strategies have been implemented to develop inhibitors of HIV proteases (for reviews, see refs *8*−*11*). One such strategy has led to the design of SB203386 (Figure 1), a potent tripeptide analogue inhibitor containing a C-terminal imidazole substituent as an amide bond isostere (*12*). The secondary alcohol was designed to engage both 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 of the imidazole were shown to act as replacements for the amide nitrogen and carbonyl oxygen of the peptidyl backbone which resulted in a substantial improvement in oral bioavailability and pharmacokinetic properties. Although a potent competitive inhibitor of HIV-1 protease ($K_i$ = 18 nM), SB203386

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

[‡] The coordinates for the three structures reported in this study have been deposited in the Brookhaven Protein Data Bank [1BDR for HIV-1 (2: 31, 33−37), 1BDL for HIV-1 (2: 31−37), and 1BDQ for HIV-1 (2: 31−37, 47, 82)].

* To whom correspondence should be addressed. E-mail: sherin_s_abdel-meguid@sbphrd.com.

[§] Department of Structural Biology.

[||] Department of Molecular Biology.

[⊥] Present address: Structural Biochemistry Program, SAIC Frederick, Frederick, MD 21701.

[1] Abbreviations: HIV, human immunodeficiency virus; AIDS, acquired immune deficiency syndrome; SIV, simian immunodeficiency virus; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; PEG, polyethylene glycol.

S0006-2960(98)00784-3 CCC: $15.00 
Published on Web 07/17/1998

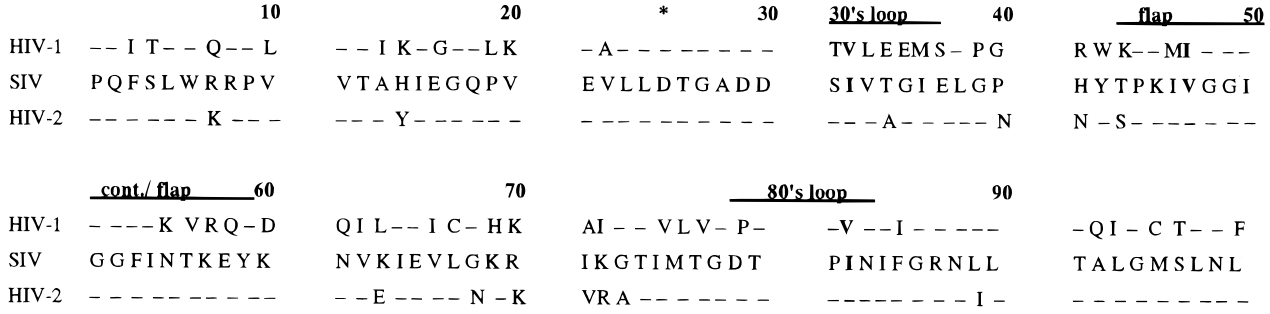


FIGURE 2: Comparison of the 99-amino acid sequences of HIV-1 ($III_b$ strain), HIV-2 (ROD strain), and SIV (251 strain) proteases. Residues constituting the 30's loop, the flaps, and the 80's loop are indicated. Active site residues which differ between HIV-1 and HIV-2/SIV are bold. Active site aspartate is indicated with an asterisk.

Table 1: Inhibition of Wild-Type and Chimeric HIV Proteases by SB203386

| protein | $K_i$[a] ± SD | ratio[b] | inhibitor mode of binding[c] |
|---|---|---|---|
| HIV-1 | 18 ± 0.00 | 1 | one inhibitor molecule per active site |
| HIV-1 (2: 32, 47, 82)[d] | 110 ± 50 | 6 | one inhibitor molecule per active site |
| HIV-1 (2: 31, 33−37) | 210 ± 20 | 12 | one inhibitor molecule per active site |
| HIV-1 (2: 31−37) | 1410 ± 180 | 78 | two inhibitor molecules, each occupying half of active site |
| HIV-1 (2: 31−37, 47, 82) | 460 ± 20 | 26 | one inhibitor molecule per active site |
| HIV-2 | 1280 ± 50 | 71 | not determined, but expected to be similar to that for SIV |
| SIV | 960 ± 50 | 53 | two inhibitor molecules, each occupying half of active site |

[a] $K_i$ values in nanomolar. All values except SIV values previously reported in Towler et al. (*18*). SIV values from Hoog et al. (*13*). [b] Ratio of $K_i$(enzyme studied)/$K_i$(HIV-1). [c] Inhibitor mode of binding as seen in the crystal structures of complexes with SB203386: wild-type HIV-1 protease from Abdel-Meguid et al. (*12*), HIV-1 (2: 32, 47, 82) and SIV proteases from Hoog et al. (*13*), and the remainder from this study. [d] Numbers in parentheses indicate the substituted residues in that chimera; e.g., HIV-1 (2: 32, 47, 82) is HIV-1 protease with HIV-2 protease residues at positions 32, 47, and 82.

shows a substantially decreased inhibition of HIV-2 ($K_i$ = 1280 nM) and SIV ($K_i$ = 960 nM) proteases (*13*). This is consistent with the fact that HIV-2 and SIV proteases share 89% of their amino acid sequences (including residues of the active site cavity), whereas HIV-1 and HIV-2 are only 45% identical in sequence (Figure 2).

Comparative analysis of the crystal structures of HIV-1 and SIV proteases complexed with SB203386 reveals nearly identical protease tertiary structures (*13*), but a different mode of binding of the inhibitor. In HIV-1 protease, one inhibitor molecule is found bound to the protease active site whereas two inhibitor molecules, each occupying one half of the active site, bind with different affinities to SIV protease (*13*). It was suggested that this different mode of binding was due to the differences in active site residues of HIV-1 and HIV-2/SIV proteases. In the active site cavity, HIV-1 and HIV-2/SIV proteases differ at residues 32, 47, and 82 in each monomer (Figure 2). However, mutating these three residues in HIV-1 protease to their counterparts from HIV-2 and SIV does not reduce the SB203386 binding affinity to that of HIV-2 and SIV, nor does it invoke the SIV protease-like inhibitor mode of binding (*13*). This finding has led to the proposal by Hoog et al. (*13*) that inhibitor specificity and mode of binding are conferred, at least in part, by residues outside the active site cavity.

Recently, Towler et al. (*18*) further analyzed residues of HIV-1 and HIV-2 proteases that are responsible for differences in binding affinity for SB203386 and other inhibitors. Using a series of chimeric proteins in which domains of the respective proteases were exchanged, they demonstrated that SB203386 affinity is conferred by a combination of the active site residues (32, 47, and 82) along with a loop comprised of residues 31−37 ("the 30's loop"), which mostly lie outside the active site cavity. To better understand the structural basis of ligand specificity in HIV proteases, we sought to study these chimeric proteins crystallographically. In this study, we report the crystal structures of the three chimeric proteases: HIV-1 (2: 31, 33−37), HIV-1 (2: 31−37), and HIV-1 (2: 31−37, 47, 82). Each of these chimeras is an HIV-1 protease in which the residues indicated in parentheses are substituted by those of HIV-2 protease. For example, HIV-1 (2: 31, 33−37) is an HIV-1 protease in which residues 31 and 33−37 have been substituted by those of HIV-2 protease. These chimeras have been previously characterized by their affinity for SB203386 (Table 1), and their structures are compared to the previously determined crystal structures of wild-type HIV-1 (*12*), triple-mutant HIV-1 (2: 32, 47, 82), and SIV (*13*) proteases complexed with SB203386. The crystal structures give new insight into the role of residues 31−37 (referred to here as the 30's loop) and residues 81−85 (part of the "80's loop") in determining the inhibitor mode of binding and specificity. Understanding the structural basis of ligand specificity is becoming increasingly important in light of recent reports of the emergence of drug resistance in patients treated with HIV protease inhibitors (*14*, *15*).

## MATERIALS AND METHODS

*Preparation of Enzyme−Inhibitor Complexes*. Three chimeric proteases were studied: HIV-1 (2: 31, 33−37), HIV-1 (2: 31−37), and HIV-1 (2: 31−37, 47, 82). The chimeras were constructed, purified, assayed, and characterized as previously described (*13*, *16*−*18*). Their substrate binding properties were shown to be comparable to those of the wild-type enzyme (*18*). The synthesis of SB203386 and its inhibition kinetics were also reported previously (*12*, *18*). Complexes of chimeric proteases with SB203386 were prepared for crystallization by adding DMSO-solubilized

Table 2: Data Acquisition and Structure Refinement Statistics for HIV-1 (2: 31, 33−37)−, HIV-1 (2: 31−37)−, and HIV-1 (2: 31−37, 47, 82)−SB203386 Complex Structures

| parameter | HIV-1 (2: 31, 33−37) | HIV-1 (2: 31−37) | HIV-1 (2: 31−37, 47, 82) |
|---|---|---|---|
| space group | $P6_1$ | $P6_1$ | $P6_1$ |
| unit cell dimensions (Å) | $a = b = 62.68$ $c = 83.49$ | $a = b = 63.14$ $c = 83.55$ | $a = b = 62.91$ $c = 83.56$ |
| resolution (Å) | 2.5 | 2.8 | 2.5 |
| no. of reflections | | | |
| total | 83 770 | 53 648 | 43 218 |
| unique | 6058 | 4054 | 5858 |
| redundancy | | | |
| total | 7.3 | 2.9 | 3.1 |
| highest-resolution shell | 7.3 | 2.9 | 3.0 |
| $R_{sym}$ (%)[a] | 12 | 13 | 14 |
| completeness (%) | | | |
| total | 99.8 | 95.6 | 99.0 |
| $I > 2\sigma I$ | 72.6 | 71.8 | 76.2 |
| highest shell with $I > 2\sigma I$ | 67.1 (2.57−2.5 Å) | 74.2 (2.9−2.8 Å) | 53.7 (2.59−2.5 Å) |
| no. of reflections used in refinement[b] | 6058 | 4054 | 5858 |
| refinement resolution range (Å) | 20−2.5 | 20−2.8 | 20−2.5 |
| no. of non-H atoms | | | |
| total | 1595 | 1584 | 1594 |
| protein | 1500 | 1502 | 1502 |
| water | 56 | 4 | 53 |
| SB203386 | 39 | 78 | 39 |
| crystallographic *R*-factor (%)[c], $F > 0$ | 17.4 | 20.1 | 19.1 |
| free *R*-factor (%)[d], $F > 0$ | 24.8 | 25.5 | 26.1 |
| mean *B*-factor (Å$^2$)[e] | | | |
| protein overall | 17.5 | 19.6 | 19.1 |
| main chain | 14.3 | − | 17.6 |
| side chain | 18.6 | − | 20.3 |
| inhibitor molecule | 10.9 | 19.6, 23.0 | 19.3 |
| inhibitor occupancy factor | 0.9 | 0.9[g] | 1.0 |
| rms deviations from standard values | | | |
| bond lengths (Å) | 0.014 | 0.019 | 0.014 |
| bond angles (deg) | 3.4 | 3.4 | 3.3 |
| dihedral angles (deg) | 27.8 | 29.2 | 28.5 |
| improper angles (deg) | 1.4 | 1.8 | 1.5 |
| Ramachandran plot (%)[f] | | | |
| favored | 94.2 | 86.5 | 92.3 |
| allowed | 5.8 | 2.8 | 7.7 |
| generous | 0.0 | 0.6 | 0.0 |

[a] $R_{sym} = 100[\sum\sum|\langle I(h)\rangle - I(h)_i|]/\sum I(h)_i$, where $I(h)_i$ is the $i$th measurement of reflection $h$ and $\langle I(h)\rangle$ is the mean intensity of symmetry-related measurements. [b] Reflections where $F > 0$ were used in refinement. [c] Crystallographic *R*-factor $= 100[\sum||F_{obs}(h)| - |F_{calc}(h)||]/\sum|F_{obs}(h)|$, where $F_{obs}(h)$ and $F_{calc}(h)$ are the observed structure factor amplitude and the structure factor amplitude calculated from the model, respectively. [d] The free *R*-factor was calculated by using 90% of the data in the last round of refinement. [e] In HIV-1 (2: 31, 33−37)− and HIV-1 (2: 31−37, 47, 82)−SB203386 structures, overall *B*-factor refinement was followed by restrained group *B*-factor refinement for the protein and combined *B*-factor and occupancy refinement for the inhibitor, taking the entire inhibitor as one group of atoms. In HIV-1 (2: 31−37)−SB203386, only the overall *B*-factor was refined. [f] The Ramachandran plot was generated in PROCHECK (*24*). [g] Occupancy factor for each of the two inhibitors in this structure.

inhibitor to 3.5 mg/mL protease solutions [containing 200 mM NaCl, 5 mM DTT, 2 mM EDTA, and 50 mM NaAc (pH 5.0)] at a 5:1 inhibitor:protease dimer molar ratio and incubating the mixture for 24 h at 4 °C.

*Crystallization of Enzyme−Inhibitor Complexes*. Crystals of the HIV-1 (2: 31, 33−37)−SB203386 complex were grown in 28% saturated ammonium sulfate and 0.2 M NaAc (pH 5.2). Those of the HIV-1 (2: 31−37)−SB203386 complex were grown in 10% PEG 1000, 0.2 M ammonium sulfate, and 0.1 M MES (pH 6.0) while those of the HIV-1 (2: 31−37, 47, 82)−SB203386 complex in 35% saturated ammonium sulfate and 0.2 M NaAc (pH 5.6). All crystals were grown using the vapor diffusion method in hanging drops at 21 °C. Needle crystals appeared in 1 week and grew to full size (0.2 mm along the shortest edge) in several months.

*X-ray Data Acquisition*. Diffraction data were measured at room temperature from crystals of all three complexes on an 18 cm MAR detector mounted on a Rigaku generator operated at 50 kV and 100 mA with monocromatized CuKα radiation in 1° oscillation frames. The data were processed using the HKL program (*19*). Crystals of all three protease−inhibitor complexes belong to space group $P6_1$ with similar unit cell dimensions (Table 2) and one protease dimer in the asymmetric unit. Data collection parameters are summarized in Table 2.

*Structure Determination and Refinement*. The crystal structures of all three protease−inhibitor complexes were determined by difference Fourier methods using X-PLOR (*20*) and phases obtained from the previously determined, isomorphous crystal structure of the wild-type HIV-1−SB203386 complex (*8*; PDB entry 1SBG). The initial model contained 93 of the 99 amino acid residues of the protease monomer, with the 30's loop (residues 31−37) excluded and active site residues 47 and 82 mutated to alanines. Inhibitor and solvent molecules were excluded from the initial model.

The structures were refined using the simulated annealing protocols in X-PLOR in iterative rounds of refinement and manual model rebuilding. Electron density visualization and model building were carried out in O (*21*). Structure refinement parameters and statistics are summarized in Table 2.

## RESULTS AND DISCUSSION

*Overall Structures and Comparison to Structures of Wild-Type HIV-1, Triple-Mutant HIV-1 (2: 32, 47, 82), and SIV Protease Complexes with SB203386.* All three chimeras studied herein crystallize in the space group $P6_1$, with unit cell parameters differing by less than 1% (Table 2). This allowed for direct comparison between them, unencumbered by concerns over differences due to different crystal packing. The overall backbone structures of the three chimeras, (2: 31, 33−37), (2: 31−37), and (2: 31−37, 47, 82), are nearly identical, with pairwise rms deviations of backbone atoms not exceeding 0.6 Å. Furthermore, the protein backbone in these three structures is nearly identical to that of the wild-type HIV-1 protease, the HIV-1 protease triple mutant (2: 32, 47, 82), and the SIV protease, each in complex with SB203386 (*12*, *13*). Electron density maps, calculated with no inhibitor in the model, clearly showed inhibitor density in the active site. This density could be accounted for by one inhibitor molecule in HIV-1 (2: 31, 33−37)− and HIV-1 (2: 31−37, 47, 82)−SB203386 structures (Figure 3a). The inhibitor conformation in these two complex structures is nearly identical to that found in the wild-type enzyme (Figure 3b), keeping the same interactions with the protein. However, the electron density in the active site of HIV-1 (2: 31−37) cannot be accounted for by one inhibitor molecule. Instead, the inhibitor is found to adopt a binding mode similar to that found in the crystal structure of the SIV−SB203386 complex (*13*), with two inhibitor molecules per protease dimer (Figure 3c). Refinement of *B*-factors and occupancy factors of the inhibitors resulted in almost full occupancy of both inhibitor sites but a higher *B*-factor for one inhibitor molecule relative to the other (Table 2). Similar observations have been made with the SIV−SB203386 complex (*13*). Each of the two bound inhibitor molecules occupies one half of the active site, with weaker binding to one half than the other.

Inhibition and structural data for all six protease−SB203386 complexes are summarized in Table 1. The proteins are listed in order of increasing number of HIV-2 residues, with wild-type HIV-1 protease at one end and HIV-2/SIV protease at the other. A trend of decreasing binding affinity of SB203386 for the protein is observed as the number of HIV-2 protease residues is increased, except for HIV-1 (2: 31−37, 47, 82) which reverts to a more HIV-1 protease-like binding. Similarly, the inhibitor mode of binding shifts toward the SIV/HIV-2 protease-like mode of binding with an increasing number of HIV-2 protease residues, except for HIV-1 (2: 31−37, 47, 82) which reverts to the HIV-1 protease-like mode of binding. In this work, we will assume that SB203386 has a mode of binding to HIV-2 protease similar to its mode of binding to SIV protease (*13*). This assumption is reasonable given their high sequence homology (Figure 2) and is supported by the data presented in this work. Thus, substitution of only seven amino acid residues (31−37) of HIV-1 protease with those of HIV-2 protease renders the former similar to the latter, in both the inhibitor binding affinity and mode of binding. On the other hand, this is reverted toward a more HIV-1 protease-like structure by the additional substitution of residues 47 and 82, suggesting that these two residues exert a compensatory effect on the active site cavity that abates the effect exerted by changing residues 31−37. These observations led to a closer examination of structural differences appearing in three regions in the protease: residues 31−37 (30's loops), residues 43−59 (the flaps), and residues 80−83 (part of the 80's loop).

*Structural Role of the 30's Loops and Their Interactions with the Flaps.* Residues 31 and 33−37 in all three chimeras contain the amino acid sequence of HIV-2 protease. Structural comparison of the chimeras with wild-type HIV-1 protease in the region of residues 34−37 of both protease monomers reveals an altered backbone conformation due to the introduction of the conformationally flexible glycine residue at position 35 in place of the wild-type glutamate. Further, in at least one of the monomers of wild-type HIV-1 protease, Glu35 forms a salt bridge with Arg57 which is located close to the N-terminal end of one of the two $\beta$-strands forming the flap. The absence of this salt bridge in the chimeras results in an apparent displacement of the flap (residues 43−59) away from the 30's loop (Figure 4a). The displacement in each flap occurs in a direction nearly parallel to the plane of the $\beta$-sheet forming the flap starting at residue 43. The average displacements of the $\alpha$-carbons in flap residues 43−59, relative to their positions in wild-type HIV-1 protease, as seen in the chimeric proteases as well as in the triple-mutant HIV-1 (2: 32, 47, 82) and SIV protease are summarized in Table 3. Unlike that in the triple mutant, in the chimeric proteins this displacement exceeds 1.0 Å in at least one of the two monomers in each of the three chimeras. This displacement is large when compared with rms deviations from wild-type HIV-1 in the rest of the protease structure (Table 3). Notably, the tips of the flaps, formed by residues 49−52 which assume a "hairpin loop" structure, have also separated from each other, relative to their positions in wild-type HIV-1 (Figure 4b). The displacement of the flap tips, however small, occurs in the plane perpendicular to the plane of the hairpin loop. Thus, the polypeptide main chain atoms in the flap tips are no longer within hydrogen-bonding distance. A hydrogen bond between the peptide backbone amide hydrogen of Gly51 in the flap tip of one monomer and the carbonyl oxygen of Ile50 in the other monomer is thought to play a role in stabilizing inhibitor binding and inducing closure of the flaps in HIV proteases (*22*). In all three protease chimeras studied herein, no hydrogen bonds between the flap tips are seen in the crystal structures. The distance between the amide nitrogen atom of Gly51 from one monomer and the carbonyl oxygen atom of Ile50 from the other monomer is $\geq$5.0 Å for all three chimeras (Table 3).

The role of the flaps in ligand binding to HIV proteases is not well understood. In all postulated mechanisms of catalysis and inhibition, the initial step is thought to be ligand interaction with an "open" form of the enzyme (*23*). Proper closure of the flaps then stabilizes the bound inhibitor. The flap shifts observed in the chimeric protease structures may suggest a decreased stability in the closed form of the enzyme

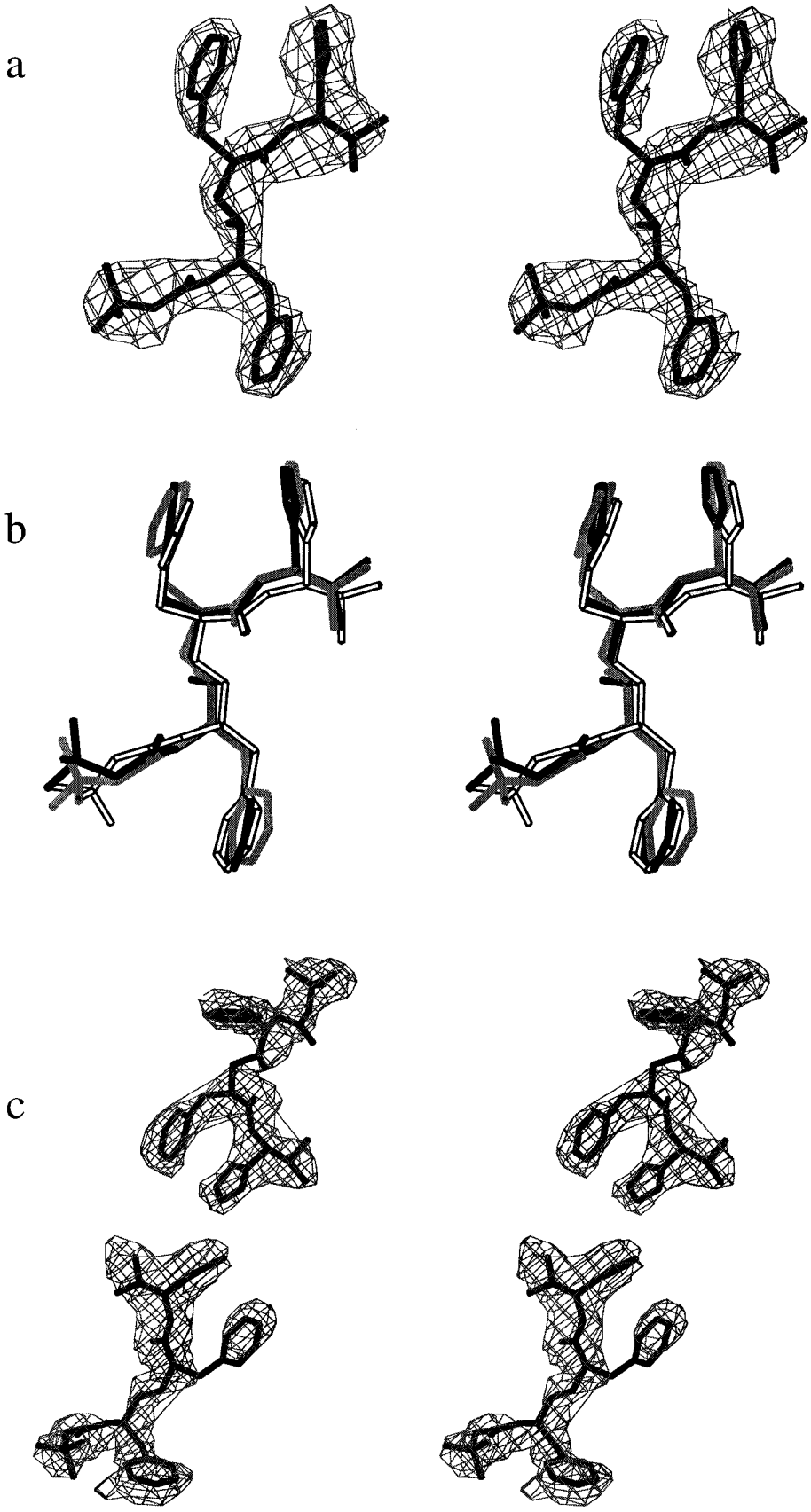


FIGURE 3: (a) Simulated annealed (2.5 Å, temperature of 2000 K) omit map, contoured at 2.5$\sigma$, of the center of the active site cavity of the HIV-2 (2: 31, 33−37)−SB203386 complex, showing the electron density for the inhibitor. The inhibitor is omitted from the phase calculation. The model of the bound inhibitor is also shown. (b) Superposition of the inhibitor model from the crystal structures of wild-type HIV-1 (solid black), HIV-1 (2: 31, 33−37) (gray), and HIV-1 (2: 31−37, 47, 82) (hollow black) complexes with SB203386. (c) Simulated annealed (2.8 Å, temperature of 2000 K) omit map, contoured at 2.5$\sigma$, showing the electron density for the two SB203386 inhibitor molecules bound in the active site of HIV-1 (2: 31−37). The inhibitor models are omitted from the map calculation.

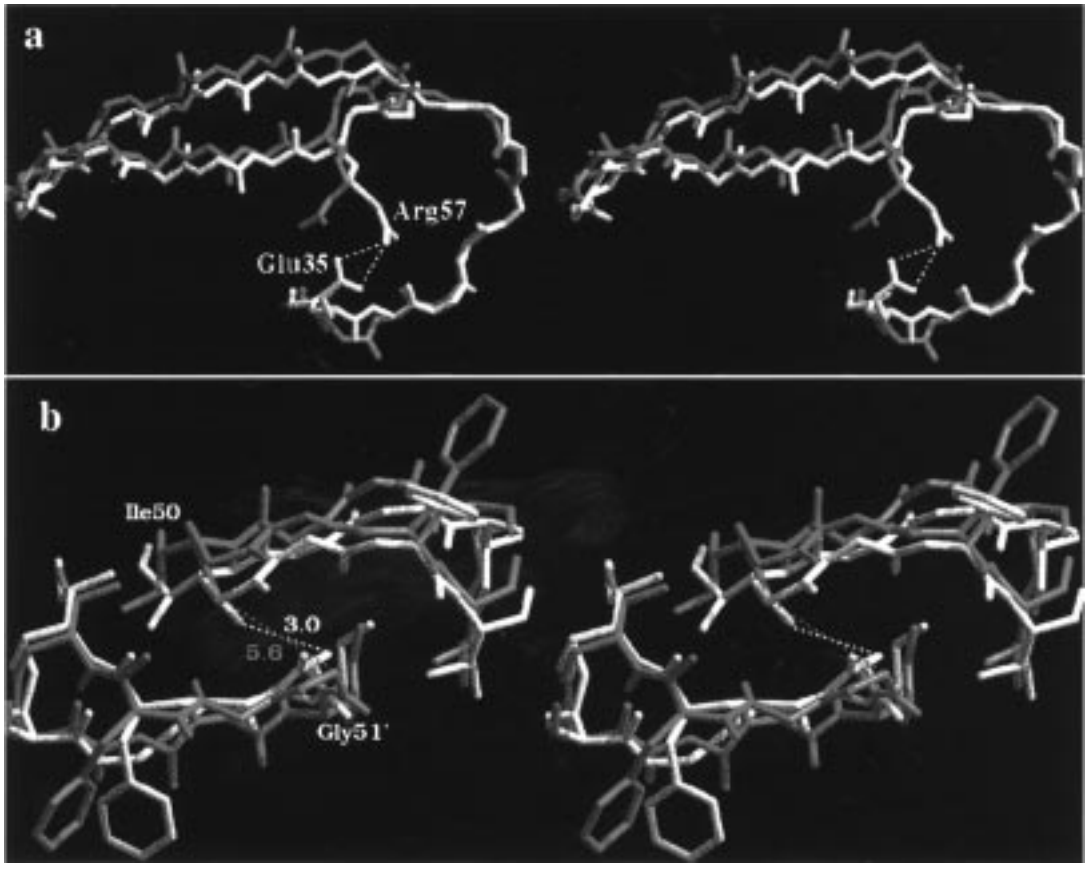


FIGURE 4: Stereoview of the superposition of the HIV-1 (2: 31, 33−37) (cyan) and wild-type HIV-1 (yellow) structures in one flap region. (a) The backbone atoms of residues 34−59 from one monomer are shown. The side chains of residues Glu35 and Arg57 forming a salt bridge in the wild-type enzyme are shown. (b) Flap tips from both monomers, in a stereoview almost down the dimer 2-fold axis. The distance (angstroms) between the Ile50 carbonyl oxygen from one monomer and the Gly51 backbone nitrogen from the other is indicated for each structure.

Table 3: Changes in Atomic Positions (in Angstroms) between Mutant and Wild-Type HIV Proteases

| protein | $C_\alpha$ rmsd values of protease dimer excluding flaps[a] | mean displacement of $C_\alpha$ atoms of flap residues 43−59 (monomer A/monomer B) | intermonomer distance between Ile50 carbonyl O and Gly51 amide N | intermonomer distance between $C_\beta$ atoms of residues 82 | intermonomer distance between $C_\beta$ atoms of residues 47 | intramonomer distance between $C_\beta$ atoms of residues 47 and 82 (monomer A/monomer B) |
|---|---|---|---|---|---|---|
| HIV-1[c] | – | – | 3.0 | 18.7 | 17.9 | 14.3/14.3 |
| HIV-1 (2: 32, 47, 82)[b] | 0.57 | 0.4/0.6 | 2.9 | 19.1 | 18.2 | 14.1/14.5 |
| HIV-1 (2: 31, 33−37) | 0.37 | 0.8/1.1 | 5.7 | 18.9 | 18.1 | 14.7/14.4 |
| HIV-1 (2: 31−37) | 0.49 | 1.0/1.3 | 5.4 | 17.6 | 18.6 | 14.4/14.6 |
| HIV-1 2: 31−37, 47, 82) | 0.42 | 0.9/1.2 | 5.0 | 17.6 | 20.2 | 15.1/15.4 |
| SIV[c] | 0.54 | 1.1/0.6 | 2.9 | 16.8 | 18.4 | 14.3/13.3 |

[a] The structures were three-dimensionally aligned with wild-type HIV-1 by least-squares fitting of $C_\alpha$ atoms of all protease residues except for flap residues 43−59 in both monomers. [b] Values obtained from Brookhaven Data Bank entries 1TCX and 1TCW for HIV-1 (2: 32, 47, 82) and SIV proteases, respectively. [c] Values obtained from Brookhaven Data Band entry 1SBG.

in solution. This may explain the overall decrease in binding affinity of SB203386 for the chimeras relative to that of the wild-type enzyme.

Our results point to the role of the 30's loop residues, particularly Glu35, in maintaining the long-range interactions within the polypeptide chain, necessary for inhibitor binding and stabilization. Although the overall tertiary structure of the protease is unaffected by the HIV-1 to HIV-2 protease Glu to Gly change at position 35, the subtle changes seen in the flap positions may be directly related to the disruption of the Glu35−Arg57 salt bridge.

*Shift of the 80's Loops in HIV-1 (2: 31−37) and HIV-1 (2: 31−37, 47, 82).* HIV-1 (2: 31−37) contains the HIV-2 protease Ile32 in the active site cavity instead of the HIV-1 protease Val32. The backbone conformation of Ile32 in HIV-1 (2: 31−37) is similar to that seen in the wild-type enzyme. However, a small movement toward the active site cavity in the polypeptide main chain at residues 79−83 (the 80's loop) in both monomers, relative to its position in the wild-type enzyme, is notable (Figure 5). This movement is also seen in HIV-1 (2: 31−37, 47, 82), and previously in SIV protease (*13*), both of which also contain an isoleucine at position 32. As in the SIV protease−SB203386 complex, this further extension of residues 79−83 toward the active site cavity results in a significant change in its shape. The distance between the $\beta$-carbon atom at position 82 in

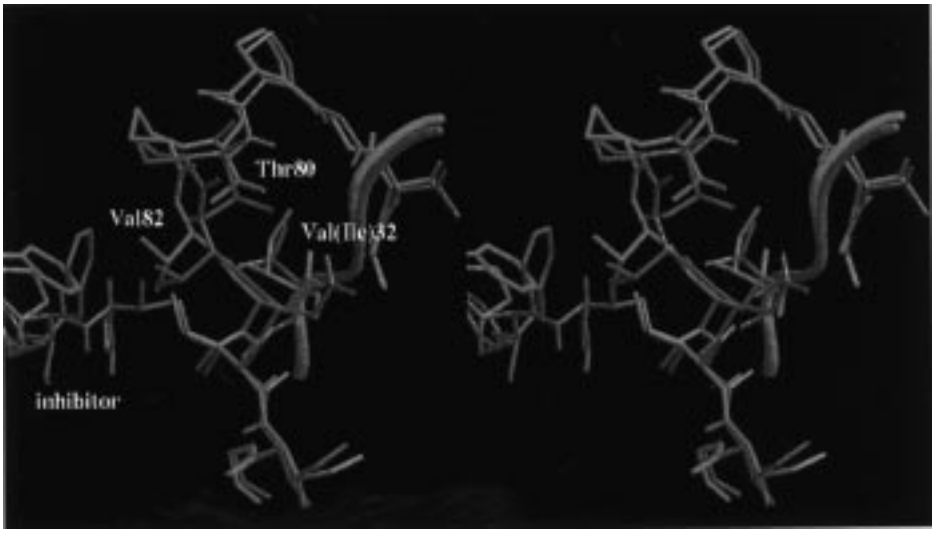


FIGURE 5: Stereoview of the superposition of HIV-1 (2: 31−37) (red) and HIV-1 (2: 31, 33−37) (cyan) in the 80's loop region of one monomer showing the shift in the positions of residues 79−83 toward the active site cavity. Shown in the foreground is the $C_\alpha$ trace for residues 31−35 and the side chain of residue 32. The inhibitor is partially shown to indicate the location of the active site. Thr80 and Val82 are also indicated.

monomer A and its symmetry-related atom in monomer B is more than 1 Å shorter in HIV-1 (2: 31−37), HIV-1 (2: 31−37, 47, 82), and SIV than in HIV-1 (31, 33−37) and wild-type HIV-1 (Table 3). This apparent decrease in the size of the active site cavity seen in HIV-1 (2: 31−37) and SIV protease may explain the lack of a bound inhibitor molecule in the center of the active site and the binding of two inhibitor molecules, each to one half of the active site, instead. The shift in the 80's loop may be directly related to the changes in the amino acid sequence of the 30's loop, particularly the size of the amino acid side chain at position 32. In the wild-type enzyme, the Val32 side chain makes a van der Waals contact with $C_\gamma$ of Thr80. Although a similar contact is present between Ile32 and Thr80 in HIV-1 (2: 31−37) and HIV-1 (2: 31−37, 47, 82), steric effects from the slightly bigger isoleucine side chain may cause the residues adjacent to Thr80 to shift (Figure 5).

The above conclusions are confirmed by the fact that the mutant protein containing a valine at position 32 [HIV-1 (2: 31, 33−37)] does not exhibit the 80's loop shift in its crystal structure, whereas the proteins that contain an isoleucine at position 32 [HIV-1 (2: 31−37), HIV-1 (2: 31−37, 47, 82), and SIV] do. However, the triple-mutant HIV-1 (2: 32, 37, 82) does not exhibit a shift in its 80's loop (*13*), indicating that although the size of residue 32 is necessary for the 80's loop shift it is not by itself sufficient; other sequence changes in the 30's loop are needed. When the $K_i$ values and inhibitor mode of binding of wild-type HIV-1 protease, HIV-1 (2: 31, 33−37), HIV-1 (2: 31−37), and SIV protease (Table 1) are compared, it appears that the effect of the HIV-2 sequence at positions 31 and 33−37 and the effect of the Val32Ile mutation are separable and additive. Introducing the HIV-2 sequence at residues 31 and 33−37 in wild-type HIV-1 protease [resulting in HIV-1 (2: 31, 33−37)] increases the $K_i$ value 12-fold but maintains the HIV-1 protease-like inhibitor mode of binding. Adding to this change, the HIV-2 isoleucine at position 32 [as in HIV-1 (2: 31−37)] increases the $K_i$ to a value comparable to that of SIV/HIV-2 proteases (53−78-fold higher than that of HIV-1 protease) and changes the inhibitor mode of binding to that of SIV protease (Table 1). This correlation indicates that the 12-fold increase in $K_i$ seen for HIV-1 (2: 31, 33−37) protease is a consequence of the 30's loop sequence and its indirect effects on the flap structure, whereas the novel inhibitor mode of binding seen in the HIV-1 (2: 31−37) and SIV protease structures is largely due to structural changes in the vicinity of residue 32, such as in the 80's loop.

*Active Site Cavity Changes Resulting from Ile47Val and Val82Ile Mutations.* HIV-1 (2: 31−37, 47, 82) contains not only the HIV-2 protease 30's loop but also its active site cavity. Interestingly, this chimera has an SB203386 binding affinity that is much closer to that of HIV-1 protease than it is to that of SIV or HIV-2 proteases (Table 1). Furthermore, the crystal structure of the HIV-1 (2: 31−37, 47, 82)−SB203386 complex reveals an inhibitor mode of binding similar to that of wild-type HIV-1 (Figure 3b). Given the trend in Table 1 and going from HIV-1 (2: 31−37) to HIV-1 (2: 31−37, 47, 82), one would have expected that the latter mutant would have maintained an HIV-2 protease-like binding. The reversion toward an HIV-1 protease-like binding mode and affinity can only indicate a compensatory effect in the size and shape of the HIV-1 (2: 31−37, 47, 82) active site cavity resulting from the introduction of additional mutations at 47 and 82. Analysis of the distances between active site cavity residues (Table 3) clearly reveals differences between HIV-1 (2: 31−37, 47, 82) and the other enzymes. Although the intermonomer distance between the $C_\alpha$ atoms of residues 82 in HIV-1 (2: 31−37, 47, 82) is comparable to that in HIV-1 (2: 31−37) (Table 3), the intermonomer distance between the $C_\alpha$ atoms of residues 47 is longer in HIV-1 (31−37, 47, 82) than in the other chimeras as well as wild-type HIV-1 and SIV proteases (Table 3). Also, the intramonomer distance between the $C_\beta$ atoms of residues 47 and 82 in each of the two monomers is longer in HIV-1 (2: 31−37, 47, 82) than in HIV-1 (2: 31−37) and in wild-type HIV-1 (Table 3). This amounts to a narrowing of the active site cavity by 1.1 Å along the dimension spanning the distance between the Ile82 residues of both monomers and its enlargement by 2.3 Å along a

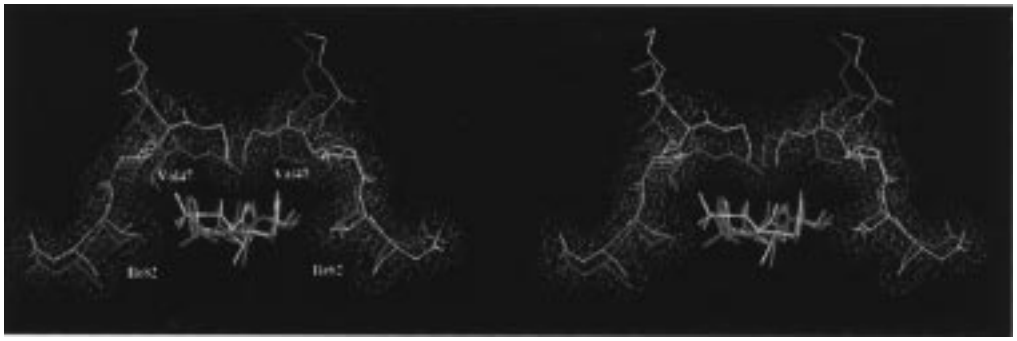

FIGURE 6: Stereoview showing the deformation of the active site cavity in HIV-1 (2: 31−37, 47, 82) (magenta), relative to wild-type HIV-1 (yellow). A van der Waals surface representation of atoms in the polypeptide segments containing residues 46−48 and 79−83 is shown for each structure. The view is along the normal to the dimer 2-fold axis. Active site residues 47 and 82 are labeled as in the HIV-1 (2: 31−37, 47, 82) sequence, monomer A on the left and monomer B on the right. The bound inhibitor molecule is shown.

dimension nearly perpendicular to the first as compared with the wild-type enzyme (Figure 6). The average *B*-factor of the 80's loop atoms (residues 79−83 from both monomers) increased from 10.8 Å$^2$ in the wild-type enzyme to 16.9 Å$^2$ in HIV-1 (2: 31−37, 47, 82). The average *B*-factor value of highly ordered residues 23−29 is 10.4 and 13.0 Å$^2$ in the wild-type and HIV-1 (2: 31−37, 47, 82) enzymes, respectively. This apparent increase in *B*-factors is consistent with an increase in structural flexibility in the 80's loop region. Similarly, the average *B*-factor of residues 45−49 from both monomers increased from 13.9 Å$^2$ in the wild-type enzyme to 30.4 Å$^2$ in HIV-1 (2: 31−37, 47, 82), indicative of even higher structural flexibility in that region. It is likely that this apparent structural flexibility combined with the compensatory enlargement of the active site cavity along the intermonomer distance between residues 47, observed in HIV-1 (2: 31−37, 47, 82) relative to HIV-1 (2: 31−37), is responsible for the restoration of the wild-type mode of inhibitor binding in HIV-1 (2: 31−37, 47, 82). It is possible that similar compensatory effects can be achieved by mutations at residues other than 47 and 82. We have focused in our study on residues 47 and 82 because they are the only active site cavity residues that are different between HIV-1 and HIV-2/SIV proteases. A similar compensatory effect of residues 47 and 82 has been previously observed in the comparative inhibition of the mutants HIV-1 (2: 32) and HIV-1 (2: 32, 47, 82), which exhibit 16- and 6-fold increases in their $K_i$ values by SB203386 relative to that of the wild-type enzyme, respectively (*18*).

The study reported herein enhances our understanding of ligand specificity toward HIV-1 and HIV-2 proteases. The results from this study confirm the importance of long-range interactions in imparting enzyme ligand specificity. They reveal a high degree of structural flexibility within the enzyme active site, manifested in changing active site dimensions in response to sequence mutations outside the active site itself. This "malleability" of the active site appears to be a significant determining factor in the enzyme's inhibition properties in solution. This may provide the virus with a significant opportunity to develop resistance to protease inhibitors while keeping an intact protease active site. This consideration is becoming increasingly important in light of the appearance of drug-resistant HIV variants in patients treated with HIV protease inhibitors (*14*, *15*) and the fact that proteases from many drug-resistant HIV variants contain mutations at residues outside the active site cavity.

## REFERENCES


1. Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W., and Rosenberg, M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8903.
2. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M., and Sigal, I. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4686.
3. Oroszlan, S., and Luftig, R. B. (1990) *Curr. Top. Microbiol. Immunol. 157*, 153.
4. 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., and Debouck, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1841.
5. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989) *Science 245*, 616.
6. McDonald, C. K., and Kuritzkes, D. R. (1997) *Arch. Intern. Med. 157*, 951.
7. Korant, B. D., and Rizzo, C. J. (1997) *Adv. Exp. Med. Biol. 421*, 279.
8. Abdel-Meguid, S. S. (1993) *Med. Res. Rev. 13*, 731.
9. Wlodawer, A., and Erickson, J. W. (1993) *Annu. Rev. Biochem. 62*, 543.
10. Reich, S. H., Melnick, M., Davies, J. F., Appelt, K., Lewis, K. K., Fuhry, M. A., Pino, M., Trippe, A. J., Nguyen, D., Dawson, H., Wu, B.-W., Musick, L., Kosa, M., Kahil, D., Webber, S., Gehlhaar, D. K., Andrada, D., and Shetty, B. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3298.
11. von-der-Helm, K. (1996) *Biol. Chem. 377*, 765.
12. Abdel-Meguid, S. S., Metcalf, B. W., Carr, T. J., Demarsh, P., DesJarlais, R. L., Fisher, S., Green, D., Ivanoff, L., Lambert, D. M., Murthy, K. H. M., Peteway, S. R. J., Pitts, W. J., Tomaszek, T. A. J., Winborne, E., Zhao, B., Dreyer, B. B., and Meek, T. D. (1994) *Biochemistry 33*, 11671.
13. Hoog, S. S., Towler, E. M., Zhao, B., Doyle, M., Debouck, C., and Abdel-Meguid, S. S. (1996) *Biochemistry 35*, 10279.
14. Condra, J. H., Holder, D. J., Schleif, W. A., Blahy, O. M., Danovich, R. M., Gabryelski, L. J., Graham, D. J., Laird, D., Quintero, J. C., Rhodes, A., Robbins, H. L., Roth, E., Shivaprakash, M., Yang, T., Chodakewitz, J. A., Deutsch, P. J., Leavitt, R. Y., Massari, F. E., Mellors, J. W., Squires, K. E., Steigbigel, R. T., Teppler, H., and Emini, E. A. (1996) *J. Virol. 70*, 8270.
15. Schock, H. B., Garsky, V. M., and Kuo, L. C. (1996) *J. Biol. Chem. 271*, 31957.
16. Grant, S. K., Deckman, I. C., Minnich, M. D., Culp, J., Franklin, S., Dreyer, G. B., Tomaszek, T. A., Jr., Debouck, C., and Meek, T. D. (1991) *Biochemistry 30*, 8424.
17. Towler, E. M., Stebbins, J. W., and Debouck, C. (1996) *Gene 183*, 259.

18. Towler, E. M., Thompson, S. K., Tomaszek, T., and Debouck, C. (1997) *Biochemistry 36*, 5128.
19. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307.
20. Brünger, A. T. (1992) *X-PLOR 3.1 Manual*, Yale University Press, New Haven, CT.
21. Jones, T. A., and Kjelgaard, M. (1992) *O-The Manual*, Uppsala University, Uppsala, Sweden.
22. Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B. H., and Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8805.
23. Meek, T. D., Rodriguez, E. J., and Angeles, T. S. (1994) *Methods Enzymol. 241*, 127.
24. Howard, A. J., et al. (1987) *J. Appl. Crystallogr. 20*, 383.

BI980784H