Three-Dimensional Structure of a Simian Immunodeficiency Virus Protease/Inhibitor Complex. Implications for the Design of Human Immunodeficiency Virus Type 1 and 2 Protease Inhibitors†

Baoguang Zhao,*,‡ Evon Winborne,‡ Michael D. Minnich,§ Jeffrey S. Culp,§ Christine Debouck,∥ and Sherin S. Abdel-Meguid*,‡

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

Received July 23, 1993; Revised Manuscript Received September 24, 1993®

ABSTRACT: Simian immunodeficiency virus (SIV) proteins have considerable amino acid sequence homology to those from human immunodeficiency virus (HIV); thus monkeys are considered useful models for the preclinical evaluation of acquired immune deficiency syndrome (AIDS) therapeutics. We have crystallized and determined the three-dimensional structure of SIV protease bound to the hydroxyethylene isostere inhibitor SKF107457. Crystals of the complex were grown from 25–32% saturated sodium chloride, by the hanging drop method of vapor diffusion. They belong to the orthorhombic space group *I*222, with a = 46.3 Å, b = 101.5 Å, and c = 118.8 Å. The structure has been determined at 2.5-Å resolution by molecular replacement and refined to a crystallographic discrepancy factor, R ($=\sum||F_o| - |F_c||/\sum|F_o|$), of 0.189. The overall structure of the complex is very similar to previously reported structures of HIV-1 protease bound to inhibitors. The inhibitor is bound in a conformation that is almost identical to that found for the same inhibitor bound to HIV-1 protease, except for an overall translation of the inhibitor, varying along the backbone atoms from about 1.0 Å at the termini to about 0.5 Å around the scissile bond surrogate. The structures of the SIV and HIV-1 proteins vary significantly only in three surface loops composed of amino acids 15–20, 34–45, and 65–70. Superposition of the 1188 protein backbone atoms from the two structures gives an rms deviation of 1.0 Å; this number is reduced to 0.6 Å when atoms from the three surface loops are eliminated from the rms calculation. Given the considerable amino acid sequence homology of SIV and HIV-2 proteases, we predict that their three-dimensional structures will be practically identical. Comparison of the structures of the SIV and HIV-1 protease/SKF107457 complexes suggests that to design inhibitors that are equally effective against HIV-1 and -2, one must maintain a balance in the size of the side chains at P_2, P_1, P_1', and P_2'. In other words, large side chains at one side of the inhibitor (P_1 and P_2') require small ones at the opposite side (P_2 and P_1'). Also, the high similarity in the active-site geometry of the proteases from SIV and HIV confirms the notion that monkeys are useful models for the preclinical evaluation of acquired immune deficiency syndrome (AIDS) therapeutics.

Considerable effort is being devoted to the discovery and development of drugs to combat the human immunodeficiency virus (HIV; Mitsuya et al., 1991), the etiological agent of the acquired immune deficiency syndrome (AIDS). Although much of this effort is focused on HIV type 1 (HIV-1; the most common strain of this virus in the United States and Europe), the emergence of HIV type 2 (HIV-2; the most common strain of this virus in West Africa; Clavel et al., 1986; Guyader et al., 1987) as a growing health threat makes it necessary to develop drugs that combat both types.

Simian immunodeficiency viruses (SIVs) constitute a group of primate retroviruses that are morphologically and antigenically related to HIVs (Henderson et al., 1988; Desrosiers, 1988; Desrosiers & Ringler, 1989). The amino acid sequence of proteins encoded by SIV genomes show greater homology to HIV-2 than to HIV-1 proteins. For example, SIV_{Mac} (that isolated from *Macaca mulatta*) protease differs from HIV-2 (*rod* isolate) protease by only 13 amino acid residues, none of which are in the active site cavity, but it differs by 49 amino acids from HIV-1 protease (Myers et al., 1990; Figure 1). This high degree of sequence homology between SIV and HIV-2 proteases is also reflected in all other proteins constituting these two viruses (Henderson et al., 1988), making SIV an ideal model for HIV-2. Furthermore, the close evolutionary kinship between SIV and HIV suggests that monkeys can be useful animal models to validate many of the strategies being developed for AIDS chemotherapy. One such strategy is inhibition of the virus-encoded protease required for processing the viral *gag* and *gag-pol* polyproteins into enzymes and structural proteins necessary for the formation of infectious virions (Kohl et al., 1988; Meek et al., 1989). The HIV-1, HIV-2, and SIV proteases are each a 99 amino acid residue protein (Figure 1) that functions as a homodimer. The enzyme from HIV-1 was classified as a member of the aspartyl protease family, on the basis of its active-site sequence similarity (Toh et al., 1985; Power et al., 1986) and structural analogy to the well-characterized monomeric enzymes renin, pepsin, rhizopuspepsin, endothiapepsin, and penicillopepsin (Pearl & Taylor, 1987). It is predominantly composed of β-strands; its active site is formed at the interface of the dimer and contains two aspartyl residues, one contributed by each subunit (Lapatto et al., 1989; Navia et al., 1989; Wlodawer et al., 1989). HIV-1 protease undergoes considerable con-

† The refined coordinates for the complex have been deposited in The Protein Data Bank under file name 1SIV.

‡ Department of Macromolecular Sciences.

§ Department of Protein Biochemistry.

∥ Department of Molecular Genetics.

® Abstract published in *Advance ACS Abstracts,* November 15, 1993.

0006-2960/93/0432-13054$04.00/0 © 1993 American Chemical Society

	5								10					15					20				25
HIV-1	—	—	Ile	Thr	—	—	Gln	—	—	Leu	—	—	Ile	Lys	—	Gly	—	—	Leu	Lys	—	Ala	—	—	—
SIV	Pro	Gln	Phe	Ser	Leu	Trp	Arg	Arg	Pro	Val	Val	Thr	Ala	His	Ile	Glu	Gly	Gln	Pro	Val	Glu	Val	Leu	Leu	Asp
HIV-2	—	—	—	—	—	—	Lys	—	—	—	—	—	—	Tyr	—	—	—	—	—	—	—	—	—	—	—

	30								35					40					45				50
HIV-1	—	—	—	—	—	Thr	Val	Leu	Glu	Glu	Met	Ser	—	Pro	Gly	Arg	Trp	Lys	—	—	Met	Ile	—	—	—
SIV	Thr	Gly	Ala	Asp	Asp	Ser	Ile	Val	Thr	Gly	Ile	Glu	Leu	Gly	Pro	His	Tyr	Thr	Pro	Lys	Ile	Val	Gly	Gly	Ile
HIV-2	—	—	—	—	—	—	—	—	Ala	—	—	—	—	—	Asn	Asn	—	Ser	—	—	—	—	—	—	—

	55								60					65					70				75
HIV-1	—	—	—	—	Lys	Val	Arg	Gln	—	Asp	Gln	Ile	Leu	—	—	Ile	Cys	—	His	Lys	Ala	Ile	—	—	Val
SIV	Gly	Gly	Phe	Ile	Asn	Thr	Lys	Glu	Tyr	Lys	Asn	Val	Lys	Ile	Glu	Val	Leu	Gly	Lys	Arg	Ile	Lys	Gly	Thr	Ile
HIV-2	—	—	—	—	—	—	—	—	—	—	—	—	Glu	—	—	—	—	Asn	—	Lys	Val	Arg	Ala	—	—

	80								85							90				95				
HIV-1	Leu	Val	—	Pro	—	—	Val	—	—	Ile	—	—	—	—	—	—	Gln	Ile	—	Cys	Thr	—	—	Phe
SIV	Met	Thr	Gly	Asp	Thr	Pro	Ile	Asn	Ile	Phe	Gly	Arg	Asn	Leu	Leu	Thr	Ala	Leu	Gly	Met	Ser	Leu	Asn	Leu
HIV-2	—	—	—	—	—	—	—	—	—	—	—	—	—	Ile	—	—	—	—	—	—	—	—	—	—

FIGURE 1: Comparison of the amino acid sequences of the proteases from HIV-1 (III_b strain; Muesing et al., 1985; Ratner et al., 1985; Wain-Hobson et al., 1985), SIV_{Mac} (Kornfield et al., 1987), and HIV-2 (*rod* isolate; Guyader et al., 1987).

formational changes upon complexation with inhibitors, particularly in the two "flaps" (flexible β-hairpin structures comprising residues 42–58 of each monomer) which move by as much as 7 Å to tightly embrace the ligands (Miller et al., 1989; Erickson et al., 1990; Fitzgerald et al., 1990; Swain et al., 1990; Bone et al., 1991; Jaskolski et al., 1991; Dreyer et al., 1992, 1993; Murthy et al., 1 992; Thompson et al., 1992; Tomaszek et al., 1992; Abdel-Meguid et al., 1993).

One goal of our research program is to design, as potential AIDS therapeutics, protease inhibitors that are equally effective against HIV-1 and HIV-2. Generally, these inhibitors are peptide analogs designed to mimic a high-energy intermediate involved in peptidolysis (Dreyer et al., 1989; Moore et al., 1989; Tomasselli et al., 1990a; Rich et al., 1990). Various nonhydrolyzable moieties, including the amino acid statine, secondary amines, hydroxyethylenes, hydroxyethylamines, and ketones, have been incorporated in these peptides to replace the scissile peptide bond [for reviews see Greenlee (1990) and Abdel-Meguid (1993)]. Although some of these inhibitors bind with equal affinity to both HIV-1 and HIV-2 proteases, others differ by as much as 2 orders of magnitude (Tomasselli et al., 1990b). These similarities and differences in binding affinities are conserved when SIV protease is substituted for HIV-2 protease. Thus, the high sequence homology and the fact that the processing sites in SIV and HIV-2 *gag* polyproteins are very similar (Henderson et al., 1988) suggest that the proteases from these two viruses will have nearly identical three-dimensional structures. To understand the molecular basis imparting these differences in inhibitor binding affinities, we have determined the crystal structure of SIV protease bound to a hydroxyethylene isostere inhibitor (SKF107457; Figure 2) and compared it to that of HIV-1 protease bound to the same inhibitor. The latter crystal structure has been previously determined in our laboratory (Dreyer et al., 1992; Murthy et al., 1992). This study has shed considerable light on the design of dual inhibitors active against both HIV-1 and HIV-2. Also, it has shown that some HIV-1 protease inhibitors can bind in an identical conformation to the HIV-1 and SIV enzymes. This should confirm the notion that monkeys are ideal animal models to test the effectiveness of HIV-1 protease inhibitors in AIDS chemotherapy.

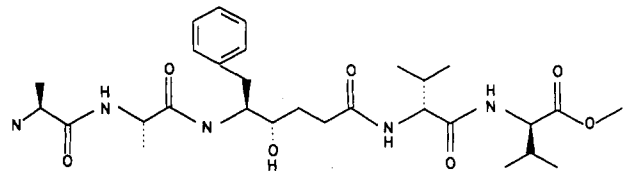

FIGURE 2: Chemical structure of the inhibitor SKF107457. This inhibitor binds to the proteases from SIV and HIV-1 with binding constants (K_i) of 8.4 and 11.7 nM, respectively (Grant et al., 1991).

EXPERIMENTAL PROCEDURES

Recombinant SIV protease, derived from SIV (MM251 isolate), was obtained by expression in *Escherichia coli* (Grant et al., 1991) and was purified to apparent homogeneity as previously described (Grant et al., 1991). The complex of SKF107457 with the protease was prepared by adding a 10 mM aqueous solution of the inhibitor to a 0.2 mg/mL protein solution in 50 mM sodium acetate buffer, pH 5.0, containing 300 mM NaCl, 2 mM EDTA, and 1 mM DTT.[1] The complex was then concentrated by ultrafiltration to a final protein concentration of 4 mg/mL, at room temperature. The synthesis and characterization of the inhibitor have been previously described (Dreyer et al., 1992), while its inhibition constants (K_i) are reported in Figure 2.

Cocrystals were grown from 25–32% saturated NaCl in 50 mM sodium acetate/acetic acid buffer, pH 5.6, using the hanging drop method of vapor diffusion (McPherson, 1976). Typically, 2 μL of the protein solution was mixed with 2 μL of various precipitant concentrations, and the mixtures were vapor-equilibrated at several temperatures against the solutions added to the protein. At every temperature, a three-dimensional matrix was established for each precipitant, in which protein concentration, precipitant concentration, and

[1] Abbreviations: DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; rms, root mean square.

pH were varied. Ammonium sulfate, sodium chloride, and various poly(ethylene glycols) were used as precipitants. Isolated crystals achieve dimensions of about 0.1 mm × 0.1 mm × 0.4 mm within 2 weeks at room temperature. The space group is $I222$ with a = 46.3 Å, b = 101.5 Å, and c = 118.8 Å. Assuming one complete copy of the complex (a protein dimer plus an inhibitor; 23 000 Da)/asymmetric unit, the V_M value is 3.0 Å^3/Da, which is within the range found for most protein crystals (Matthews, 1968). These crystals show diffraction beyond 2.5-Å resolution and are stable in the X-ray beam for several days.

X-ray diffraction data were measured from a single crystal using a Siemens two-dimensional position-sensitive detector. The detector was mounted on a Siemens rotating-anode X-ray generator operated at 40 kV and 100 mA, equipped with a 300-μm focusing cup and producing graphite-monochromated Cu Kα radiation. Diffraction images were recorded, reduced, and visualized using FRAMBO, XENGEN (Howard et al., 1987), and XPREP computer programs, respectively. Approximately 41 170 reflections were measured in 3 days to give 10 018 unique reflections to 2.5-Å resolution, representing 99% of the data. The merging R factor for symmetry-related reflections (R_{sym}) was 0.068 ($R_{sym} = \sum\sum|I_{i(h)} - \bar{I}_{(h)}|/\sum I_{i(h)}$, where $I_{i(h)}$ are the intensities of multiple measurements of reflection h and $\bar{I}_{(h)}$ is their mean).

The structure of the SIV protease/inhibitor complex was determined by molecular replacement using the XPLOR program package (Brunger, 1992). Two starting models were used; they consisted of all atoms of the monomer and the dimer of HIV-1 protease derived from the crystal structure of HIV-1 protease bound to the SB204144 inhibitor (Abdel-Meguid et al., 1993). Rotation functions were calculated using X-ray diffraction data from 10 to 3 Å and a radius of integration of 35 Å. Results from the cross-rotation functions revealed consistently two significant peaks (θ_1, θ_2, θ_3 are 23.1°, 30.0°, 11.1° and 217.2°, 30.0°, 169.2°) related by a 2-fold axis that is consistent with the result from the self-rotation function (φ, ϕ, χ are 30°, 90°, 180°). The heights of the two cross-rotation functions peaks were larger than 15σ, while the self-rotation function peak had a height of 4.5σ. Translation functions were calculated using the space groups $I222$ and $I2_12_12_1$. The correct choice of the space group, $I222$, was determined by graphical examination of crsytal packing of the dimer. The translation peak corresponded to a fractional unit cell of (0.033, 0.092, 0.050); its height was 19σ. The R factor ($R = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively) for the correct molecular replacement solution was 0.48 for data greater than $2\sigma(F_o)$ in the resolution range 8–3 Å. The resulting phases were used to calculate Fourier maps with coefficients $|F_o| - |F_c|$ and $2|F_o| - |F_c|$, into which an atomic model of the SIV protease dimer was built and analyzed using the molecular graphics program FRODO (Jones, 1985) running on an ESV 10 graphic system. Although the electron density of the inhibitor was very clear, no attempt was made to build a model of the inhibitor at this point. The R factor was reduced to 0.24 when data greater than $2\sigma(F_o)$, in the resolution range 6–2.5 Å, were used to refine all atomic positions of the SIV protease dimer. Difference electron density maps calculated using refined phases showed clear, interpretable electron density for the inhibitor. To ensure consistency, the structure of HIV-1 protease bound to SKF107457 (Murthy et al., 1992) was further refined using XPLOR. The model of SKF107457, obtained from the structure of HIV-1 protease bound to SKF107457, was then positioned in the electron density map and the complex was refined using XPLOR. An additional cycle of map fitting, the addition of 34 water molecules, and further refinement resulted in a final crystallographic R factor of 0.189, for the data between 6- and 2.5-Å resolution. The overall rms deviation from ideal geometry for bond lengths was 0.017 Å, while that for bond angles was 3.48°. The refined set of atomic coordinates has been deposited in the Protein Data Bank (PDB; Bernstein et al., 1977).

RESULTS AND DISCUSSION

Description of the Structure. The overall three-dimensional structure of the SIV protease/SKF107457 inhibitor complex is similar to previously published structures of HIV-1 protease/inhibitor complexes (Miller et al., 1989; Erickson et al., 1990; Fitzgerald et al., 1990; Swain et al., 1990; Bone et al., 1991; Jaskolski et al., 1991; Dreyer et al., 1992, 1993; Murthy et al., 1992; Thompson et al., 1992; Tomaszek et al., 1992; Abdel-Meguid et al., 1993), where the protease dimer, having its flaps in a "closed" conformation, embraces the inhibitor and shields it from bulk solvent. The inhibitor binds in the active-site cavity in an extended conformation. It is held by a set of hydrophobic and hydrogen-bonding interactions, many of which are conserved in similar structures (Abdel-Meguid, 1993; Wlodawer & Erickson, 1993). Every heteroatom of the inhibitor is hydrogen-bonded to an atom of the protein, either directly or indirectly through a water molecule (Figure 3). The water molecule centrally located on the enzyme's 2-fold axis near the flaps, found in all ligated structures of HIV-1 protease, is also present in the structure of the SIV protease/SKF107457 inhibitor complex. This water molecule bridges the inhibitor and the protein forming hydrogen bonds with the carbonyl oxygens at P_2 [nomenclature of Schechter and Berger (1967)] and P_1' on one side and with the amide nitrogens of Ile50 and Ile50′ of the flaps on the other side. Two other water molecules interact with the inhibitor and bridge it to the enzyme (Figure 3). One bridges between residue Asp330 and the amide nitrogen at P_3, while the other participates in a bridge between the carbonyl oxygen at P_2' and Thr26, Gly27, and Arg87.

Unlike the case of the HIV-protease/SKF107457 inhibitor complex, where the inhibitor was found in two orientations related to each other by the pseudo-2-fold axis of the enzyme dimer (Dreyer et al., 1992; Murthy et al., 1992), the inhibitor in the SIV protease complex is found in a single orientation. Figure 4 shows clearly defined electron density for the inhibitor, allowing for no ambiguity in this interpretation. The absence of a second orientation of this inhibitor in the SIV protease structure weakens the argument put forth by Murthy et al. (1992) that the presence of two orientations for the inhibitor in the HIV-1 protease/SKF107457 structure is a consequence of having an inhibitor that is small enough to be entirely contained within the active site of the protease dimer, thus making no direct or indirect contribution to the lattice contacts. However, an alternative explanation of why some inhibitors are found in two orientations while others are present in only one must await a careful analysis of the numerous structures of retroviral protease/inhibitor complexes that will soon be available in the PDB (Bernstein et al., 1977).

The structures of the two protein subunits of SIV protease are very similar. Superposition of all 1510 non-hydrogen atoms of the two subunits gave an rms deviation of 1.5 Å, while superposition of the 594 backbone atoms gave an rms deviation of only 0.5 Å. The conformations of almost all the side chains of the two subunits are highly similar, except for those of the

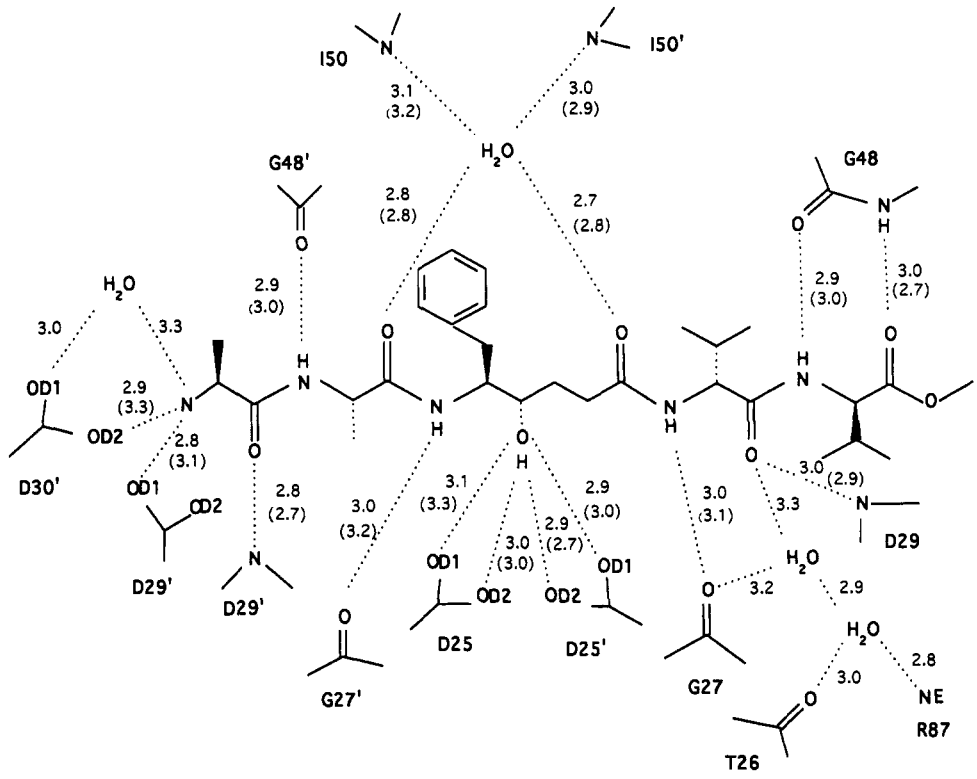

FIGURE 3: Schematic representation of potential hydrogen-bond interactions between the inhibitor and the protein. Distances between pairs of hydrogen-bonded atoms are shown. Those not in parentheses are from the structure of the SIV protease/SKF107457 complex, while those in parentheses are from the structure of the HIV-1 protease/SKF107457 complex.

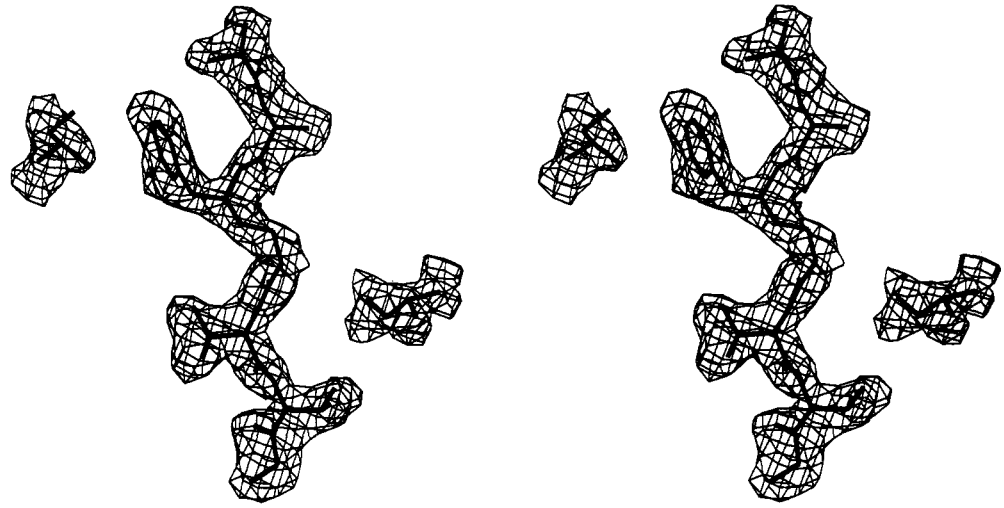

FIGURE 4: Stereoview showing the fit of SKF107457 and SIV protease residues Ile82 and Ile82′ to the final $2|F_o| - |F_c|$ electron density map, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively.

side chains of residues 6, 21, 37, 53, 63, 72, and 82. The latter residue, Ile82, protrudes in the active-site cavity of SIV protease. In the S_1 pocket, the side chain of Ile82 has its δ-carbon atom pointing away from the benzyl group of the inhibitor at P_1, while the same atom of Ile82′ points toward the inhibitor in the cavity created by the absence of a P_1' side chain (Figure 5a).

Comparison with SKF107457 from the Structure of the HIV-1 Protease Complex. The crystal structure of SKF107457 bound to HIV-1 protease was independently determined by Dreyer et al. (1992) and Murthy et al. (1992). The determination of the structure of SKF107457 bound to SIV protease thus allows for a direct comparison of the inhibitor's conformations in the active sites of both enzymes. The inhibitor binds to the active site of both SIV and HIV-1 proteases with similar affinities, having binding constants (K_i) of 8.4 and 11.7 nM, respectively (Grant et al., 1991). Figure 3 shows considerable similarity of potential hydrogen bond distances between the inhibitor and the SIV and HIV-1 proteins. The active-site cavity is formed by residues 8, 23, 25, 27–30, 32, 47–50, 53, 76, 80–82, and 84 (Gustchina & Weber, 1991). All of these residues are identical between SIV and HIV-2 proteases, and only residues 32, 47, 76, and 82 show differences with HIV-1 protease (Figure 1). Given that there are only four amino acid residues that are different in the active-site cavities of the two enzymes (Figure 1), it is reasonable to assume that the conformations of the inhibitor in the active site of both enzymes will be similar. Indeed it is; the two conformations are practically identical except for minor differences at the C-terminus (Figure 5b). Super-

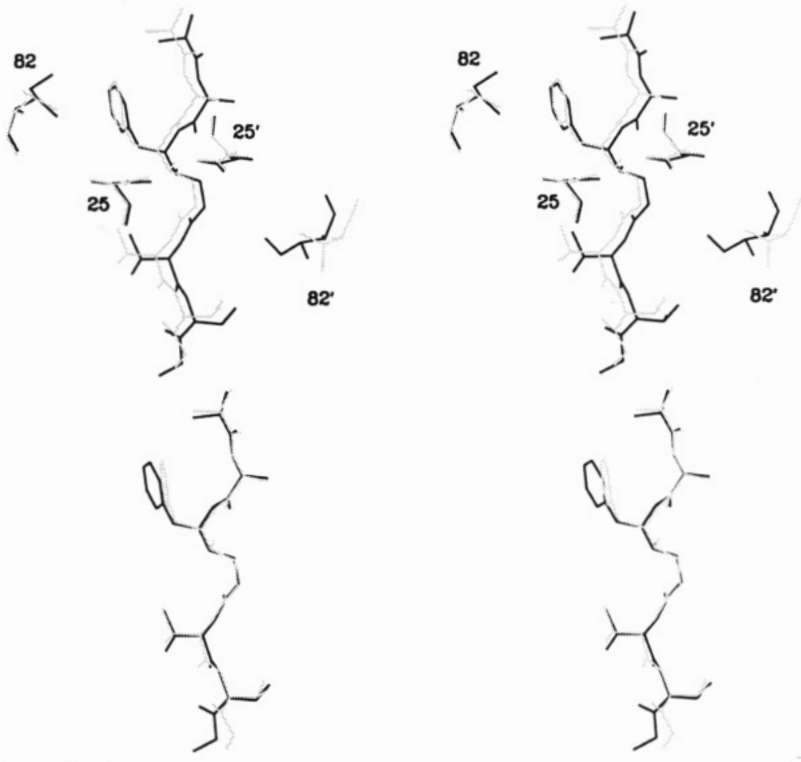

FIGURE 5: Stereoview of the overlay of SKF107457 active-site conformations in the proteases from SIV (solid lines) and HIV-1 (dotted lines). Overlay was achieved by (a, top) superposition of the protein backbone atoms of the two complexes and (b, bottom) superposition of only the 38 backbone atoms of the two inhibitors. The conformations of residues 25, 25′, 82, and 82′ are also shown in (a). Residues 25 and 25′ are the catalytic aspartates in both proteins, while residues 82 and 82′ are isoleucines in SIV protease and valines in HIV-1 protease.

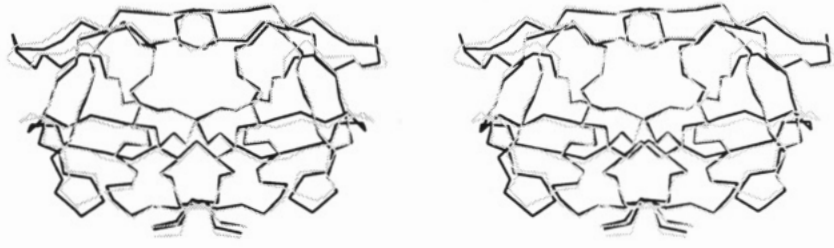

FIGURE 6: Stereoview of the overlay of the α-carbon atoms of the SIV (solid lines) and HIV-1 (dotted lines) protease dimers derived from their crystal structures with SKF107457.

position of all 82 atoms of the inhibitor from the two enzymes gives an rms deviation of 0.53 Å. Similarly, superposition of only the 38 backbone atoms gives an rms deviation of 0.51 Å. The only meaningful difference in binding of SKF107457 to SIV and HIV-1 proteases is an overall translation (Figure 5a) varying along the backbone atoms from about 1.0 Å at the termini to about 0.5 Å around the scissile bond surrogate. This translation is a consequence of amino acid sequence differences in the S_1 and S_2' pockets, where the valine residues at positions 32 and 82 in HIV-1 protease are replaced by the larger isoleucine residues in the SIV enzyme. These two amino acid differences result in a reduction in the size of the S_1/S_1' and S_2/S_2' pockets and consequently a translation of the inhibitor (Figure 5a) to relieve unfavorable van der Waals interactions between the inhibitor's side chains at P_1 and P_2' and the protein.

In all ligated structures of retroviral proteases, the central residues of the inhibitor (P_2–P_2') are buried in the active site and shielded from bulk solvent. The side chains of the inhibitors at P_2 and P_1' point in one direction, while the side chains at P_1 and P_2' point in the opposite direction. The side chains of the SKF107457 inhibitor at P_2, P_1, P_1', and P_2' are methyl, benzyl, hydrogen, and isopropyl, respectively (Figure 2). The two larger side chains, the benzyl and isopropyl, are on one side of the inhibitor while the two smaller ones, hydrogen and methyl, are on the opposite side (Figure 4). Thus, a translation of the inhibitor toward the S_2 and S_1' pockets can be easily accommodated without detrimentally affecting the inhibitor binding affinity. On the other hand, if we add a benzyl group at P_1' of this inhibitor, we should expect lower affinity for binding to SIV protease relative to HIV-1 protease. The inhibitor would contain two large side chains at P_1 and P_1' that must fit in the smaller S_1 and S_1' pockets of the SIV enzyme. Indeed this is the case; such an inhibitor has been reported to bind to SIV protease with a K_i of 1.6 nM (Grant et al., 1991), while it binds to HIV-1 protease with a K_i of 0.6 nM (Dreyer et al., 1992). Since residues comprising the active-sites of SIV and HIV-2 proteases are identical, a similar argument can be made to explain the observation made by Tomasselli et al. (1990b) that HIV-2 is better able to hydrolyze substrates with small amino acids in P_1 and P_1'.

Comparison with HIV-1 Protease. Figure 6 shows the superposition of the SIV and HIV-1 protease structures, the latter derived from the structure with SKF107457 (Murthy et al., 1992). Superposition of the 1188 protein backbone atoms from these two structures gives an rms deviation of 1.0 Å. Although the overall folds of the SIV and HIV-1 proteases are highly similar, there are significant differences in three surface loops composed of amino acid residues 15–20, 34–45, and 65–70, none of which is part of the active-site cavity. The

largest difference in protein bacakbone atoms is in the 34–45 loop, where the α-carbon atoms of residue 40 differ by as much as 4.1 Å. This loop being next to the flap may be the cause of the small, but noticeable, reorientation of the tips of the flaps (Figure 6) and may contribute to the differential binding affinity reported for certain inhibitors with regard to SIV/HIV-2 and HIV-1 proteases (Tomasselli et al., 1990b; Grant et al., 1991). When atoms of these three regions are eliminated from the superposition, the rms deviation is reduced to 0.6 Å for the remaining atoms. Inspection of crystal contacts between symmetry-related dimers of SIV protease and of HIV-1 protease, in their respective crystal lattices, shows that some of these differences may be due to crystal packing forces. In the SIV protease structure, amino acid residues in the loops composed of residues 15–20 and 34–45 are involved in crystal contacts, while that composed of residues 65–70 is not. In the HIV-1 protease/SKF107457 (Dreyer et al., 1992; Murthy et al., 1992), however, only the loop composed of residues 34–45 is involved in crystal contacts. Nevertheless, Mulichak et al. (1993) have reported similar differences in these three loops after they compared the crystal structures of HIV-2 and HIV-1 complexed to inhibitors. The HIV-2 protease/inhibitor molecules in their crystal form differ in their packing from those in our SIV protease structure, confirming the notion that the differences we observe between SIV and HIV-1 protease are genuine.

Inspection of the side-chain conformations of residues having the same amino acid sequence in SIV and HIV-1 proteases shows little conformational change. Although such a lack of differences may be a consequence of the structure determination of the SIV protease structure by molecular replacement using an HIV-1 protease structure as a model, the electron densities for most amino acid residues are well defined, implying a conservation in side-chain conformation between SIV and HIV-1 proteases.

Implication for the Design of Dual HIV-1/HIV-2 Inhibitors. Only 13 amino acid residues are different between the proteases from SIV and HIV-2 (Figure 1), most of which are conservative substitutions on the surface of the protein and none of which are in or near the active-site cavity. Thus for the purpose of this discussion, we will assume that the two structures are identical.

The binding affinity of certain inhibitors with regard to SIV (unpublished results) or HIV-2 (Tomasselli et al., 1990b) proteases was shown to differ by as much as 2 orders of magnitude from their affinity of binding to HIV-1 protease. Inspection of our structure does not directly address this issue, since SKF107457 binds with similar affinities to both SIV and HIV-1 proteases (Figure 2). However, it clearly shows that the S_2, S_1, S_1', and S_2' pockets are shallower than the respective pockets in HIV-1 protease, a consequence of the presence of isoleucines in SIV and HIV-2 proteases versus valines at residues 32 and 82 in the HIV-1 enzyme. The shallowness of the S_1 and S_1' pockets does not necessarily prohibit the accommodation of large side chains at P_1 or P_1', since the δ-carbon atom of Ile82 can assume a conformation that minimizes its protrusion in the active site (Figure 5a). Nevertheless, inhibitors with two large side chains at P_1 and P_1' would be required to assume nonideal conformations in the active site to relieve unfavorable van der Waals contacts. Similarly, the shallowness of S_2 and S_2' pockets dictates a balance in the size of the side chains at P_2 and P_2'. Thus, to design inhibitors that are equally effective against HIV-1 and HIV-2 proteases, it is important to balance the size of the side chains at P_2, P_1, P_1', and P_2' to ensure a proper fit of the inhibitor in a shallower active site. In other words, large side chains at one side of the inhibitor (P_1 and P_2') require small ones at the opposite side (P_2 and P_1').

Further crystallographic studies, particularly with inhibitors that show large differences in binding affinity with regard to HIV-2 (or SIV) and HIV-1 proteases, coupled with structure–activity studies of substrates and inhibitors should supply more direct evidence of the structural requirement for optimal binding to HIV-2 protease.

Monkeys as Viable Animal Models. Although HIV-1 protease was first identified in 1985 (Toh et al., 1985; Debouck et al., 1987) and its three-dimensional structure with and without inhibitor was determined in 1989 (Lapatto et al., 1989; Navia et al., 1989; Wlodawer et al., 1989), therapeutic models have not yet been established. Animal models, such as monkeys, can be invaluable for such a use. The considerable similarity between the structures of SIV and HIV proteases, particularly in the active site, supports the validity of using SIV infection as a model for the preclinical evaluation of HIV protease inhibitors as AIDS therapeutics. This is especially true for the development of inhibitors of HIV-2 infection, a growing health threat in Africa and Asia.

ACKNOWLEDGMENT

We are grateful to Geoffrey Dreyer for useful comments and a gift of the SKF107457 inhibitor, and to Krishna Murthy for assistance with the space group determination.

REFERENCES

Abdel-Meguid, S. S. (1993) *Med. Res. Rev. 13*, 731–778.

Abdel-Meguid, S. S., Zhao, B., Murthy, K. H. M., Winborne, E., Choi, J.-K., DesJarlais, R. L., Minnich, M. D., Culp, J. S., Debouck, C., Tomaszek, T. A., Jr., Meek, T. D., & Dreyer, G. B. (1993) *Biochemistry 32*, 7972-7980.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bone, R., Vacca, J. P., Anderson, P. S., & Holloway, M. K. (1991) *J. Am. Chem. Soc. 113*, 9382–9384.

Brunger, A. T. (1992) *X-PLOR Manual*, Version 3.0, Yale University, New Haven, CT.

Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M.-A., Santos-Ferreira, M. O., Laurent, A. G., Dauguet, C., Katlama, C., Rouzioux, C., Klatzmann, D., Champalimaud, J. L., & Montagnier, L. (1986) *Science 233*, 343–346.

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.

Desrosiers, R. C. (1988) *Annu. Rev. Microbiol. 42*, 607–625.

Desrosiers, R. C., & Ringler, D. J. (1989) *Intervirology 30*, 301–312.

Dreyer, G. B., Metcalf, B. W., Tomaszek, T. A., Jr., Carr, T. J., Chandler, A. C., III, Hyland, L., Fakhoury, S. A., Magaard, V. W., Moore, M. L., Strickler, J. E., Debouck, C., & Meek, T. D. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9752–9756.

Dreyer, G. B., Lambert, D. M., Meek, T. D., Carr, T. J., Tomaszek, T. A., Jr., Fernandez, A. V., Bartus, H., Cacciavillani, E., Hassell, A. M., Minnich, M., Petteway, S. R., Jr., Metcalf, B. W., & Lewis, M. (1992) *Biochemistry 31*, 6646–6659.

Dreyer, G. B., Boehm, J. C., Chenera, B., DesJarlais, R. J., Hassell, A. M., Meek, T. D., Tomaszek, T. A., Jr., & Lewis, M. (1993) *Biochemistry 32*, 937–947.

Erickson, J., Neidhart, D. J., VanDrie, J., Kempf, D. J., Wang, X. C., Norbeck, D. W., Plattner, J. J., Rittenhouse, J. W., Turon, M., Wideburg, N., Kohlbrenner, W. E., Simmer, R., Helfrich, R., Paul, D. A., & Knigge, M. (1990) *Science 249*, 527–533.

Fitzgerald, P. M. D., McKeever, B. M., VanMiddlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C. T., Herber, W. K., Dixon, R. A. F., & Darke, P. L. (1990) *J. Biol. Chem. 265*, 14209–14219.

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.

Greenlee, W. J. (1990) *Med. Res. Rev. 10*, 173–236.

Gustchina, A., & Weber, I. T. (1991) *Proteins 10*, 325–339.

Guyader, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L., & Alizon, M. (1987) *Nature 326*, 662–669.

Henderson, L. E., Benveniste, 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., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383–387.

Jaskolski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Heinrikson, R. L., Schneider, J., Kent, S. B. H., & Wlodawer, A. (1991) *Biochemistry 30*, 1600–1609.

Jones, T. A. (1985) *Methods Enzymol. 115*, 157–171.

Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A. F., Scolnick, E. M., & Sigal, I. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4686–4690.

Kornfield, H., Riedel, N., Viglianti, G. A., Hirsch, V., & Mullins, J. I. (1987) *Nature 326*, 610–613.

Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., Geoghegan, K. F., Hawrylik, S. J., Lee, S. E., Scheld, K. G., & Hobart, P. M. (1989) *Nature 342*, 299–302.

Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491–497.

McPherson, A., Jr. (1976) *Methods Biochem. Anal. 23*, 249–345.

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.

Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H., & Wlodawer, A. (1989) *Science 246*, 1149–1152.

Mitsuya, H., Yarchoan, R., Kageyama, S., & Broder, S. (1991) *FASEB J. 5*, 2369–2381.

Moore, M. L., Bryan, W. M., Fakhoury, S. A., Magaard, V. W., Huffman, W. F., Dayton, B. D., Meek, T. D., Hyland, L., Dreyer, G. B., Metcalf, B. W., Strickler, J. E., Gorniak, J. G., & Debouck, C. (1989) *Biochem. Biophys. Res. Commun. 159*, 420–425.

Muesing, M. A., Smith, D. H., Calredila, C. D., Benton, E. V., Lasky, L. A., & Capon, D. J. (1985) *Nature 313*, 450–458.

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. (1993) *J. Biol. Chem. 268*, 13103–13109.

Murthy, K. H. M., Winborne, E. L., Minnich, M. D., Culp, J. S., & Debouck, C. (1992) *J. Biol. Chem. 267*, 22770–22778.

Myers, G., Rabson, A. B., Berzofsky, J. A., Smith, T. F., & Wong-Staal, F., Eds. (1990) *Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences*, Los Alamos National Laboratory, Los Alamos, NM.

Navia, M. A., Fitzgerald, P. M. D., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., & Springer, J. P. (1989) *Nature 337*, 615–620.

Pearl, L. H., & Taylor, W. R. (1987) *Nature 329*, 351–354.

Power, M. D., Marx, P. A., Bryant, M. L., Gardner, M. B., Barr, P. J., & Luciw, P. A. (1986) *Science 231*, 1567–1572.

Ratner, L., Haseltine, W. A., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. K., Ghrayeb, J., Chang, N. T., Gallo, R. C., & Wong-Staal, F. (1985) *Nature 313*, 277–284.

Rich, D. H., Green, J., Toth, M. V., Marshall, G. R., & Kent, S. B. H. (1990) *J. Med. Chem. 33*, 1285–1288.

Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.

Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Schneider, J., Kent, S. B. H., & Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8805–8809.

Thompson, W. J., Fitzgerald, P. M. D., Holloway, M. K., Emini, E. A., Darke, P. L., McKeever, B. M., Schleif, W. A., Quintero, J. C., Zugay, J. A., Tucker, T. J., Schwering, J. E., Homnick, C. F., Nunberg, J., Springer, J. P., & Huff, J. R. (1992) *J. Med. Chem. 35*, 1685–1701.

Toh, H., Ono, M., Saigo, K., & Miyata, T. (1985) *Nature 315*, 691.

Tomasselli, A. G., Olsen, M. K., Hui, J. O., Staples, D. J., Sawyer, T. K., Heinrikson, R. L., & Tomich, C. S. (1990a) *Biochemistry 29*, 264–269.

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. (1990b) *J. Biol. Chem. 265*, 14675–14683.

Tomaszek, T. A., Jr., Moore, M. L., Strickler, J. E., Sanchez, R. L., Dixon, J. S., Metcalf, B. W., Hassell, A., Dreyer, G. B., Brooks, I., Debouck, C., Meek, T. D., & Lewis, M. (1992) *Biochemistry 31*, 10153–10168.

Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., & Alizon, M. (1985) *Cell 40*, 9–17.

Wlodawer, A., & Erickson, J. W. (1993) *Annu. Rev. Biochem. 62*, 543–585.

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.