A Symmetric Inhibitor Binds HIV-1 Protease Asymmetrically[†,‡]

Geoffrey B. Dreyer,*[,§] Jeffrey C. Boehm,[§] Balan Chenera,[§] Renée L. DesJarlais,[∥] Anne M. Hassell,[⊥,ℱ] Thomas D. Meek,[§] and Thaddeus A. Tomaszek, Jr.[§]

Departments of Medicinal Chemistry, Physical and Structural Chemistry, and Macromolecular Sciences, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, Pennsylvania 19406

Mitchell Lewis*

Johnson Research Foundation, Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received October 21, 1992; Revised Manuscript Received November 12, 1992

ABSTRACT: Potential advantages of C_2-symmetric inhibitors designed for the symmetric HIV-1 protease include high selectivity, potency, stability, and bioavailability. Pseudo-C_2-symmetric monools and C_2-symmetric diols, containing central hydroxymethylene and (*R*,*R*)-dihydroxyethylene moieties flanked by a variety of hydrophobic P1/P1′ side chains, were studied as HIV-1 protease inhibitors. The monools and diols were synthesized in 8–10 steps from D-(+)-arabitol and D-(+)-mannitol, respectively. Monools with ethyl or isobutyl P1/P1′ side chains were weak inhibitors of recombinant HIV-1 protease (K_i > 10 μM), while benzyl P1/P1′ side chains afforded a moderately potent inhibitor (apparent K_i = 230 nM). Diols were 100–10 000× more potent than analogous monools, and a wider range of P1/P1′ side chains led to potent inhibition. Both classes of compounds exhibited lower apparent K_i values under high-salt conditions. Surprisingly, monool and diol HIV-1 protease inhibitors were potent inhibitors of porcine pepsin, a prototypical asymmetric monomeric aspartic protease. These results were evaluated in the context of the pseudosymmetric structure of monomeric aspartic proteases and their evolutionary kinship with the retroviral proteases. The X-ray crystal structure of HIV-1 protease complexed with a symmetric diol was determined at 2.6 Å. Contrary to expectations, the diol binds the protease asymmetrically and exhibits 2-fold disorder in the electron density map. Molecular dynamics simulations were conducted beginning with asymmetric and symmetric HIV-1 protease/inhibitor model complexes. A more stable trajectory resulted from the asymmetric complex, in agreement with the observed asymmetric binding mode. A simple four-point model was used to argue more generally that van der Waals and electrostatic force fields can commonly lead to an asymmetric association between symmetric molecules.

The aspartic protease of human immunodeficiency virus-1 (HIV-1)[1] cleaves the viral *gag* and *gag–pol* polyproteins to produce functional components of the infectious virion. Inhibitors of HIV-1 protease effectively block viral maturation and infectivity in vitro and represent a promising class of potential AIDS therapeutics [Meek et al. (1990) and Lambert et al. (1992); for a review see Huff (1991)]. One strategy for inhibitor design employs dipeptide surrogates, such as hydroxyethylene isosteres (Figure 1), that are thought to resemble a high-energy enzyme-bound intermediate involved in peptidolysis (Dreyer et al., 1989, 1992; Huff, 1991). Although this approach has led to a variety of potent HIV-1 protease inhibitors, difficulties in formulation and delivery of these agents have slowed efforts to apply them therapeutically. Derived from peptide substrates, such compounds bear a resemblance to peptides including amide links and an overall amino-to-carboxyl directionality, and this vestigial peptide structure may contribute to their poor pharmacokinetic properties (Plattner & Norbeck, 1990).

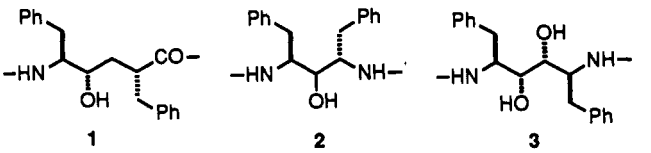


FIGURE 1: Subunits of HIV-1 protease inhibitors: **1**, hydroxyethylene isostere analogue of the dipeptide Phe-Phe; **2**, pseudosymmetric monool; **3**, symmetric diol.

A greater divergence from peptide structure could result by designing binding ligands directly from the three-dimensional structure of the enzyme. In addition to potential pharmacological benefits from a reduction of peptide structure, such inhibitors could be highly selective toward the enzyme for which they are designed. A starting point for such efforts is provided by X-ray structures of HIV-1 protease complexed with substrate analogue inhibitors (Miller et al., 1989; Fitzgerald et al., 1990; Swain et al., 1990; Jaskólski et al., 1991; Dreyer et al., 1992; Thompson et al., 1992; Thanki et al., 1992). HIV-1 protease has an unusual homodimeric, C_2-symmetric structure in which the symmetry axis bisects the substrate binding groove. Accordingly, potent inhibitors have been devised that incorporate binding elements of substrate analogues within pseudosymmetric or symmetric structures (Erickson et al., 1990; Kempf et al., 1990; Chenera et al., 1991; Bone et al., 1991; Babine et al., 1992). Two examples are the diamino alcohol **2** and diamino diol **3** shown in Figure

† Supported in part by NIH grant GM-39526.

‡ The crystallographic coordinates have been deposited in the Brookhaven Protein Data Bank.

§ Department of Medicinal Chemistry, SmithKline Beecham.

∥ Department of Physical and Structural Chemistry, SmithKline Beecham.

⊥ Department of Macromolecular Sciences, SmithKline Beecham.

ℱ Present address: Glaxo Inc., 5 Moore Drive, Research Triangle Park, NC 27709.

[1] Abbreviations: Cbz, benzyloxycarbonyl; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; HIV-1, human immunodeficiency virus type 1; rms, root mean square; THF, tetrahydrofuran.

0006-2960/93/0432-937$04.00/0

1, which can be thought of as dimeric versions of the amino-terminal half of the hydroxyethylene isostere **1**.

One might expect that a C_2-symmetric inhibitor designed to bind the symmetric protease would exhibit a highly symmetric binding mode. In the crystal structure of HIV-1 protease complexed with A74704, an inhibitor based on structure **2**, the bound conformation of the inhibitor and the enzyme dimer is approximately symmetric but still exhibits significant deviations from symmetry (Erickson et al., 1990). It is not clear whether this is due to crystal packing forces or to the asymmetry of the inhibitor, since the central nonstereogenic hydroxyl group in structure **2** breaks the C_2 symmetry. On the other hand, even asymmetric substrate analogues exhibit a quasi-symmetric binding mode with relatively minor asymmetries in the enzyme dimer. As a consequence, in the crystal structures of HIV-1 protease complexed with a series of hydroxyethylene isosteres we found that the inhibitors bind in two orientations related by the pseudo 2-fold symmetry axis of the protein (Dreyer et al., 1992). We therefore anticipated that a C_2-symmetric inhibitor, e.g., based on structure **3**, should bind symmetrically to the protease, removing the positional ambiguity associated with 2-fold disorder.

With the intention of examining symmetry as a design element for HIV-1 protease inhibitors, we have developed a synthesis of structures **2** and **3** in which the hydrophobic side chains can be easily varied (Chenera et al., 1991). We have compared these pseudosymmetric and symmetric compounds as HIV-1 protease inhibitors, examined the effects of changes in the P1/P1′ side chains[2] and flanking amino acyl residues, and investigated their activity against a monomeric, unsymmetrical aspartic protease, porcine pepsin. We have also determined the X-ray crystallographic structure of the complex of HIV-1 protease with a symmetric inhibitor.

MATERIALS AND METHODS

Synthesis of Inhibitors. (1,5-Cyclooctadiene)(pyridine)-(tricyclohexylphosphine)iridium(I) hexafluorophosphate [*Ir*(cod)(py)(Cy_3P)PF_6] was prepared as described (Stork & Kahne, 1983). 1,2:5,6-Dianhydro-3,4-*O*-isopropylidene-D-mannitol (**9**) was prepared by the procedure of Le Merrer et al. (1987). THF was distilled from sodium/benzophenone ketyl. Ether refers to commercial anhydrous diethyl ether. Pyridine, dimethylformamide, and DMSO were anhydrous grade (Aldrich). All other reagents and solvents were purchased and used without further purification. Flash chromatography was carried out with Kieselgel 60 silica (E. Merck). Complete spectral and analytical data for all compounds are available as supplementary material.

Monool Inhibitors. Illustrated for Compound I. (*2R,4R*)-*1,2,:4,5-Dianhydro-3-(benzyloxy)arabitol* (**4**). Toluenesulfonyl chloride (38.8 g, 204 mmol) was added in portions to a stirring solution of D-(+)-arabitol (15.2 g, 100 mmol) in dry pyridine (350 mL) cooled in an ice bath. After 3 h the solution was allowed to warm to room temperature and then was poured into a mixture of 700 mL of cold 6 N HCl and 500 mL of ether. The layers were separated, and the aqueous layer was extracted with 800 mL of ether. The combined organic layers were washed with 3% $NaHCO_3$ (400 mL), dried over $MgSO_4$, and concentrated in vacuo to give the crude primary ditosylate (34.2 g, 74%), which was then dissolved in 200 mL of THF and added to a stirred suspension of NaH (230 mmol; 10 g of a 55% dispersion in oil, rinsed with hexane to remove oil) in 300 mL of THF at 0 °C. After 1 h, benzyl bromide (12 mL, 100 mmol) in THF (10 mL) was added dropwise. After 1 h further, the ice bath was removed and the mixture was stirred overnight. The mixture was recooled to 0 °C and excess NaH was decomposed by slow dropwise addition of 50 mL of water (*caution*: evolution of hydrogen!). The mixture was diluted with 750 mL of ether and washed with water. The organic layer was dried over $MgSO_4$ and concentrated. The residue was purified by flash chromatography eluting with hexane followed by 1:10 ethyl acetate/hexanes to yield **4** (6.10 g, 30% overall from arabitol) as an oil. Anal. ($C_{12}H_{14}O_3$), C, H.

(3R,5R)-3,5-Dihydroxy-4-(benzyloxy)heptane (**5**, R = Me). To a suspension of CuI (2.87 g, 15 mmol) in 40 mL of ether at −35 °C was added methyllithium (20 mL, 1.5 M in ether; 30 mmol). The resulting colorless solution was stirred at −30 °C for 30 min and then cooled to −78 °C. A solution of the diepoxide **4** (930 mg, 4.50 mmol) in 10 mL of ether was added. The reaction was allowed to warm to 25 °C over 4 h and then saturated aqueous NH_4Cl and concentrated aqueous NH_3 were added. Extraction of the organic layer and trituration of the residue with ethyl acetate/hexanes provided **5** (R = Me) (1.01 g, 96%), mp 70–72 °C. Anal. ($C_{14}H_{22}O_3$) C, H.

(3R,5R)-3,5-Di(methylsulfonyloxy)-4-(benzyloxy)heptane (**6**, *R = Me*). Methanesulfonyl chloride (1.0 mL, 13 mmol) was added dropwise to diol **5** (430 mg, 1.84 mmol) in pyridine (5 mL) at 0 °C. The mixture was allowed to warm to 25 °C and, after 12 h, was diluted with cold 6 N HCl (10 mL) and extracted with CH_2Cl_2. The organic extract was washed with 3% $NaHCO_3$, dried over $MgSO_4$, and concentrated. The residue was purified by flash chromatography (1:4 ethyl acetate/hexanes) to provide **6** (R = Me) (588 mg, 81% yield). Anal. ($C_{16}H_{26}O_7S_2$) C, H.

(3S,5S)-3,5-Diazido-4-(benzyloxy)heptane (**7**, *R = Me*). A mixture of dimesylate **6** (R = Me) (588 mg, 1.49 mmol) and sodium azide (1.40 g, 21.0 mmol) in 5 mL of dimethylformamide was heated to 70 °C for 12 h. After cooling, ethyl acetate (20 mL) was added and the mixture was filtered and concentrated in vacuo. Flash chromatography (1:20 ethyl acetate/hexanes) of the residue provided **7** (R = Me) (353 mg, 83% yield).

(3S,5S)-3,5-Diamino-4-(benzyloxy)heptane. $LiAlH_4$ (200 mg, 5.3 mmol) was added to a solution of diazide **7** (R = Me) (328 mg, 1.14 mmol) in 5 mL of THF at 0 °C. The mixture was allowed to warm to 25 °C and was stirred for 5 h. The mixture was stirred with 0.5 mL of 15% NaOH for 15 min, diluted with 150 mL of ether, and filtered. Concentration of the filtrate provided the diamine (278 mg, 100%).

(3S,5S)-3,5-Diamino-4-hydroxyheptane Dihydrochloride (**8**, *R = Me*). To a solution of (3*S*,5*S*)-3,5-diamino-4-(benzyloxy)heptane (52 mg, 0.22 mmol) in methanol (4 mL) were added 20% $Pd(OH)_2$ on carbon (50 mg) and 37% aqueous HCl (two drops). The mixture was stirred under 1 atm H_2 for 16 h, filtered, and concentrated to provide **8** (R = Me) (53 mg, 100%).

(3S,5S)-3,5-Bis-[[(carbobenzyloxy)valyl]amino]-4-hydroxyheptane (I). To a stirring solution of Cbz-Val (133 mg, 0.500 mmol) in 2 mL of THF at −40 °C were added

[2] In discussing the monool and diol inhibitors, we use a loose adaptation of the convention of Schechter and Burger (1967) in designating inhibitor side chains P1, P2′, etc., and the corresponding enzyme specificity pockets S1, S2′, etc. For a C_2-symmetric structure, P*i* and P*i*′ (and S*i* and S*i*′) are in principle equivalent; hence we refer to the *i*th pair of side chains and binding pockets as P*i*/P*i*′ and S*i*/S*i*′, respectively. In discussing the complex with compound VII, we again refer to distinct P*i* and P*i*′ side chains since the asymmetric binding mode makes them nonequivalent.

N-methylmorpholine (65 μL, 0.50 mmol) and isobutyl chloroformate (65 μL, 0.50 mmol). After 10 min, a solution of diamine hydrochloride **8** (R = Me) (25 mg, 0.11 mmol) and *N*-methylmorpholine (50 μL) in dimethylformamide (1 mL) was added. The cold bath was removed, and after 12 h the mixture was diluted with ethyl acetate and washed with 5% HCl, 5% $NaHCO_3$, and brine. Concentration of the organic layer and flash chromatography of the residue (ethyl acetate/hexanes) provided I (18 mg). Anal. ($C_{33}H_{48}N_4O_7$) C, H, N.

Synthesis of II–IV proceeded much as described for I. Reaction of **4** with Ph_2CuLi or $(2\text{-propenyl})_2CuLi$ yielded **5** (R = Ph, 82%; R = 2-propenyl, 96%). Quantitative hydrogenation of 2-propenyl to *i*-Pr in **5** was carried out under 1 atm H_2 using $Ir(cod)(py)(Cy_3P)PF_6$ as the catalyst as described for **10** (R = 2-propenyl) below. When R = Ph or *i*-Pr, the yield of **7** from dimesylate **6** was lowered to 45–50% due to competing elimination reactions; purification was readily achieved following $LiAlH_4$ reduction to the diamine (before removal of the benzyl protecting group) by column chromatography on Florisil.

Diol Inhibitors. Illustrated for Compound VIII. (4R,5R,6R,7R)-4,7-Dihydroxy-5,6-(isopropylidenedioxy)-2,9-dimethyldeca-2,9-diene (**10**, *R = 2-Propenyl*). Lithium wire pieces (700 mg, 100 mmol; 1% Na) were stirred with 2-bromopropene (7.0 mL, 75 mmol) in ether (75 mL) at 0 °C. Reaction was complete after 1 h to afford a 0.50 M solution of 2-propenyllithium as determined by titration using diphenylacetic acid. A 32-mL (16 mmol) portion of this solution was added to a stirred suspension of CuI (1.5 g, 7.8 mmol) in ether (45 mL) at −78 °C, and the resultant mixture was warmed to −45 °C over 30 min. The dark mixture was recooled to −78 °C, and a solution of diepoxide **9** (2.50 g, 13.5 mmol) in 50 mL of ether was added. After 1 h the reaction mixture was allowed to warm to 25 °C and then saturated aqueous NH_4Cl and ammonium hydroxide were added. Extractive workup and flash chromatography of the residue (1:8 ethyl acetate/hexanes) gave **10** (R = 2-propenyl) (3.20 g, 88%). Anal. ($C_{15}H_{26}O_4$) C, H.

(4R,5R,6R,7R)-4,7-Dihydroxy-5,6-(isopropylidenedioxy)-2,9-dimethyldecane (**10**, *R = i-Pr*). Diol **10** (R = 2-propenyl) (2.10 g, 7.75 mmol) was stirred with $Ir(cod)(py)(Cy_3P)PF_6$ (0.50 g, 0.62 mmol) in CH_2Cl_2 under 1 atm H_2 for 24 h. Ether was added, and the mixture was filtered through silica gel. Removal of solvents gave **10** (R = i-Pr) (1.95 g, 93%).

(4R,5R,6R,7R)-4,7-Di(methylsulfonyloxy)-5,6-(isopropylidenedioxy)-2,9-dimethyldecane (Dimesylate Derivative of **10**). Methanesulfonyl chloride (1.5 mL, 19 mmol) was added to a stirring solution of diol **10** (R = i-Pr) (1.18 g, 4.31 mmol) in pyridine (10 mL) at 0 °C. After 2 h, the mixture was warmed to 25 °C, stirred an additional 12 h, and then poured into 50 mL of ice-cold 3 N HCl. Extractive workup and flash chromatography (1:4 ethyl acetate/hexanes) provided the dimesylate of **10** (1.78 g, 96%).

(4S,5R,6R,7S)-4,7-Diamino-5,6-(isopropylidenedioxy)-2,9-dimethyldecane (**11**, *R = i-Pr*). A solution of the dimesylate derivative of **10** (R = i-Pr) (1.19 g, 2.66 mmol) and NaN_3 (1.1 g, 17 mmol) in 10 mL of DMSO was heated to 100 °C for 10 h. The mixture was cooled, diluted with ether, and filtered. The filtrate was washed with 50 mL of water, dried over Na_2SO_4, and concentrated to give the diazide (810 mg). A 790-mg portion of the crude diazide was stirred for 12 h with $Pd(OH)_2$ on carbon (400 mg) in ethyl acetate (15 mL) under 1 atm H_2. The mixture was filtered and concentrated. The residue was purified by chromatography on 10 g of Florisil, eluting with hexanes, then 1:4 ethyl acetate/hexanes, and finally methanol to provide diamine **11** (R = i-Pr) (620 mg, 86% overall). Anal. ($C_{15}H_{32}N_2O_2$) C, H, N.

(4S,5R,6R,7S)-4,7-Bis[[(benzyloxycarbonyl)amino]valyl]-amino-5,6-(isopropylidenedioxy)-2,9-dimethyldecane (Acetonide of IX). Isobutyl chloroformate (106 μL, 0.800 mmol) was added to a stirring solution of Cbz-Val (200 mg, 0.800 mmol) and *N*-methylmorpholine (130 μL, 1.20 mmol) in 5 mL of THF at −40 °C. After 30 min, a solution of diamine **11** (R = i-Pr) (75 mg, 0.28 mmol) in 5 mL of THF was added. The reaction mixture was stirred at −40 °C for 2 h and at 25 °C for 16 h, diluted with ethyl acetate (100 mL), and washed successively with 5% HCl and 5% $NaHCO_3$. The organic layer was dried over $MgSO_4$ and concentrated. The residue was triturated with hexanes and ether to provide acetonide-protected IX (190 mg, 93%). Anal. ($C_{41}H_{62}N_4O_8$) C, H, N.

(4S,5R,6R,7S)-4,7-Bis[[[(benzyloxycarbonyl)amino]valyl]amino]-5,6-dihydroxy-2,9-dimethyldecane (IX). The acetonide of IX (45 mg, 0.061 mmol) was suspended in 5 mL of 70% acetic acid and heated to 85 °C for 10 h. The mixture was cooled and concentrated in vacuo. The residue was triturated with ether to give pure IX (40 mg, 93%). Anal. ($C_{38}H_{58}N_4O_8 \cdot (1/2H_2O)$) C, H, N.

Synthesis of V–VIII and X–XIII proceeded generally as described for IX but with the following modifications. Reaction of **9** with benzyl alcohol/NaH in THF afforded **10** with R = OBn (75%). Reaction of **9** with Ph_2CuLi or CuI/allylmagnesium bromide afforded **10** with R = Ph (87%) or R = allyl (100%). Reaction of **9** with lithium phenylacetylide/$BF_3 \cdot Et_2O$ at −78 °C in THF provided **10** with R = PhC≡C (100%). Hydrogenation of **10** where R is allyl or PhC≡C was effected quantitatively with 1 atm H_2 and 10% Pd/C in methanol to afford **10** with R = *n*-Pr or $Ph(CH_2)_2$, respectively. Azide displacement of the bis(mesylate) of **10** was accompanied by varying degrees of elimination (depending on R) to provide, after reduction with $H_2/Pd(OH)_2$ or $LiAlH_4$ and Florisil chromatography, **11** where R = Ph, BnO, *n*-Pr, and $Ph(CH_2)_2$ in respective overall yields of 57%, 59%, 76%, and 53%. Conversion of **8** and **11** to the final products proceeded as described for I and IX.

Enzymes and Substrates. Recombinant HIV-1 protease, corresponding to the BH10 clone of HIV-1, was obtained using the PRO4 expression vector in *Escherichia coli* strain AR58 and purified to >95% homogeneity as previously described (Strickler et al., 1989). Porcine pepsin A was obtained as a lyophilized powder from Boehringer Mannheim and used without further purification. The substrates Ac-ArgAlaSerGlnAsnTyrProValVal-NH_2 (Moore et al., 1989) and PheGlyHis[4-nitro-Phe]PheAlaPhe-OMe were obtained from Bachem Bioscience, Inc. and were of >95% purity.

Enzyme Inhibition Assays. For HIV-1 protease, 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, 200 mM NaCl, 1 mM dithiothreitol, 0.1% (v/v) Triton X-100, and 10% (v/v) DMSO (MENDT buffer) in reaction mixtures containing 1–10 mM concentrations of the substrate Ac–ArgAlaSerGlnAsnTyrProValVal-NH_2, variable concentrations of the inhibitor, and 20 nM HIV-1 protease. Alternatively, a buffer containing 100 mM sodium acetate (pH 4.7), 1 mM EDTA, 1 mM dithiothreitol, 1.0 M NaCl, 1 mg/mL bovine serum albumin, and 10% DMSO (AEDBN buffer) was used at 30 °C. Reactions at pH 3.5 and 4.5 were conducted in a buffer composed of MENDT containing additional 50 mM concentrations each of glycine, sodium acetate, and Tris (GAMNT-EDT buffer), adjusted to pH 3.5 or 4.5 and constant ionic strength as shown by conductivity

measurements. HPLC quantification of reaction products and determination of inhibition constants (Dixon analysis) were performed as previously (Moore et al., 1989; Dreyer et al., 1989). For some inhibitors, competitive inhibition was verified from double-reciprocal plots of initial velocity vs substrate concentration at several fixed concentrations of inhibitor (Lineweaver–Burk analysis). For compounds which were inhibitory at concentrations equivalent to those of the enzyme, apparent inhibition constants were obtained from plots of remaining initial rate vs inhibitor concentration at several fixed enzyme concentrations, as described (Dreyer et al., 1992). Inhibition studies of porcine pepsin were performed by spectrophotometric assay (Medzihradszky, et al., 1970) at 37 °C in a reaction mixture composed of 40 mM sodium formate (pH 3.5 or 4.5), 1% DMSO, 10 nM pepsin, ≤0.1 mM of the substrate PheGlyHis[4-nitro-Phe]PheAlaPhe-OMe, and variable inhibitor concentration, using Dixon analysis. Inhibition of pepsin was assumed to be linear competitive.

Crystallography. Large needle-shaped cocrystals of HIV-1 protease complexed with the diol VII (SKF108361) were grown as previously described for a series of hydroxyethylene isostere complexes (Dreyer et al., 1992). Precession photography demonstrated that the space group of the crystals was either $P6_1$ with strong noncrystallographic 22 symmetry or $P6_122$, with unit cell dimensions $a = b = 63.4$ Å and $c = 84.0$ Å. The crystals appeared to be isomorphous with those of the hydroxyethylene isostere complexes and with those reported by Erickson et al. (1990) for the complex of HIV-1 protease with the inhibitor A74704. To be consistent with these previous reports, the structure was solved by assuming that the additional symmetry was noncrystallographic and the space group was $P6_1$.

The three-dimensional diffraction data were recorded using a Siemens multiwire X-ray area detector that was mounted on a rotating anode X-ray source, operated at 40 kV and 50 mA using a 200 μm focal cup with a graphite monochromator. Data were recorded from a single crystal, by collecting a series of 10′ oscillation frames (400 s per frame). A total of 15 346 observations recorded to a nominal resolution of 2.3 Å were reduced to 4201 unique intensity measurement with a scaling residual of 0.059. Ninety-one percent of the theoretically possible data to 2.6 Å was observed as judged by $I/\sigma(I) > 2$. Data beyond 2.6 Å were not reliable and were discarded.

A preliminary set of phases was derived from the model of the HIV-1 protease complex with the inhibitor A74704 (Erickson et al., 1990). The model for the protein of the symmetric diol complex was refined with the restrained least-squares program PROLSQ (Hendrickson, 1985). Since the two halves of the enzyme are chemically identical, constraints were imposed on the dimer to maintain 2-fold symmetry in the initial cycles of refinement. After several cycles of refinement, the R value dropped to 30% for data between 10 and 3 Å. Electron density maps were calculated with coefficients $2|F_o| - |F_c|$ and $3|F_o| - 2|F_c|$ using the calculated phases. The higher resolution data with d spacings of >2.6 Å were added, and further cycles of refinement lowered the residual to less than 26%. A further improvement of the model was obtained by using omit maps.

The symmetric diol was built in an extended conformation in the active site in a manner similar to the hydroxyethylene isosteres (Dreyer et al, 1992). Surprisingly, the best fit of the diol inhibitor to the electron density did not position the inhibitor 2-fold axis coincident with the pseudo 2-fold axis of the protein. As was observed with the hydroxyethylene isosteres, the electron density in the vicinity of the inhibitor binding site was diffuse and the diol could be built in two orientations that were related by the pseudocrystallographic 2-fold axis. Additional cycles of refinement were performed using the program XPLOR (Brünger et al., 1987). The noncrystallographic constraints imposed on the dimer were replaced with restraints, and the residual dropped to 22.6%. The restraints were gradually weakened to allow the structure to deviate from 2-fold symmetry. A final residual of 18.3% was calculated for all data to 2.6 Å using 1516 protein atoms, 72 inhibitor atoms, and 1 water molecule. The rms deviation of the bond lengths was 0.018 Å and the rms deviation of bond angle distances was 0.037 Å.

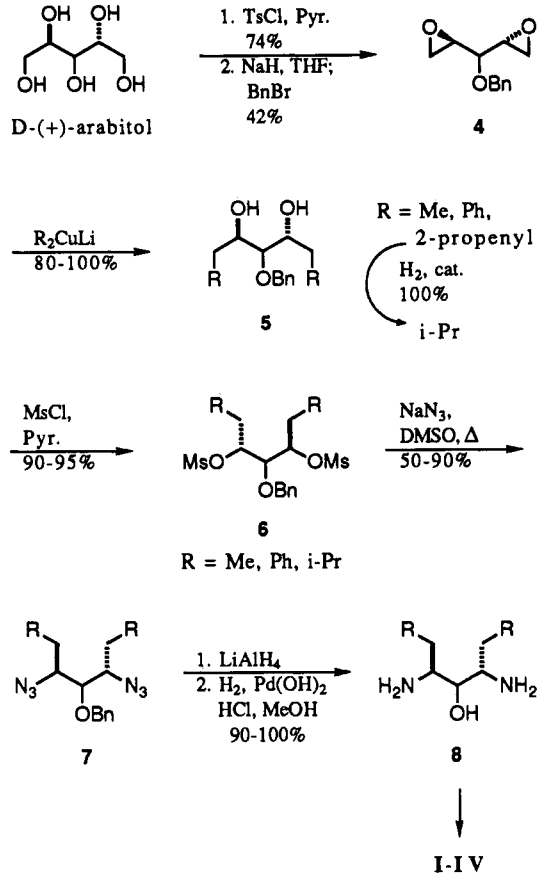
Molecular Dynamics Simulations. Molecular dynamics calculations were performed on a model solvated complex of HIV-1 protease with the diol inhibitor VII using the program AMBER Revision 3A and the parameter set parm89a (Weiner et al., 1986; Seibel et al., 1989). Calculations were performed beginning with two different initial conformations of the inhibitor: one C_2 symmetric and the other asymmetric. The initial models were built from the X-ray crystallographic structures of the complexes of HIV-1 protease with the inhibitors MVT101 [Miller et al. (1989); Brookhaven Protein Data Bank (PDB) file name 4hvp] and JG365 [Swain et al. (1990); PDB file name 7hvp]. The protein structure for the models was taken from the coordinates of the MVT101 complex. The symmetric inhibitor was built so that the conformation of its AlaAla residues approximated the conformations of the P3–P2 portions of the inhibitors MVT101 and JG365 and was positioned in the active site of HIV-1 protease with its C_2-symmetry axis coinciding with the protein pseudo-C_2 axis. The asymmetric inhibitor was constructed similarly, but with one of its central hydroxyl groups coinciding with the position of the hydroxyl group of JG365. Consistent with pH studies on the binding of substrates and inhibitors [e.g., see Hyland et al. (1991)], one carboxyl group of the Asp25, Asp25′ pair was protonated and the other was unprotonated. Charges for the neutral aspartate were calculated using MOPAC 5.0 modified to fit partial atomic charges to the electrostatic potential (Bessler et al., 1990). Counterions were placed near charged amino acids to make an electrostatically neutral system, and any water molecules identified in the MVT101 complex were included. The system was solvated with a cap of water defined by spheres of radius 22 Å centered on the atoms of residues Asp25 and Asp25′. In order to relieve any unfavorable interactions, the waters and counterions were minimized and subjected to 1 ps of room-temperature (298 K) molecular dynamics using a distance-dependent dielectric. The system was then subjected to 400 ps of room-temperature molecular dynamics, during which the inhibitor, all protein residues within 8 Å of the inhibitor, and all counterions and water molecules within 22 Å of residues Asp25 and Asp25′ were allowed to move. Coordinate sets and energy were saved every picosecond.

RESULTS

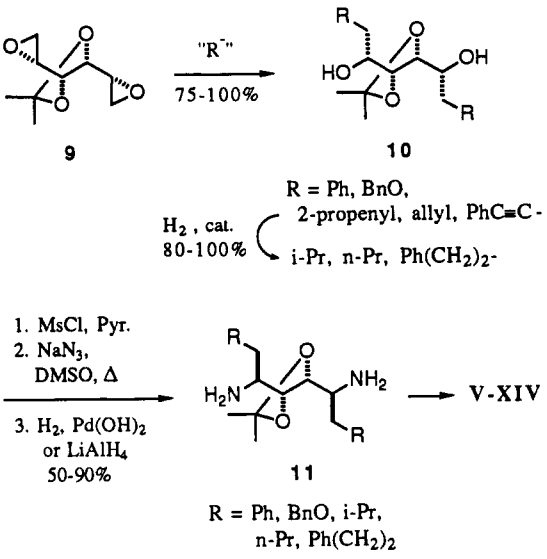
Inhibitor Synthesis. Efficient access to the two structural classes represented by **2** and **3** is afforded by the two-directional chain synthesis strategy outlined in Schemes I and II (Chenera et al., 1991). Methods similar to those shown in Scheme II have also been reported by Ghosh et al. (1991) and Jadhav and Woerner (1992). Alternative syntheses of structures **2** and **3** have been described by Kempf et al. (1992).

The syntheses of the monools I–IV (Table I) and diols V–XIV (Table II) followed analogous pathways proceeding through the known diepoxides **4** and **10**, readily available from

Scheme I: Synthesis of Pseudosymmetric Monool Inhibitors

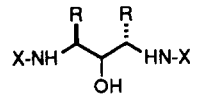


Scheme II: Synthesis of C_2-Symmetric Diol Inhibitors



D-(+)-arabitol and D-(+)-mannitol, respectively (Le Merrer et al., 1987; Nakatsuka et al., 1990; Chenera et al., 1991). Reaction of **4** and **10** with organocuprate reagents or other nucleophiles (e.g., benzyl alkoxide) effected introduction of the side chains needed to fill the enzyme S1/S1′ pockets. Unsaturated side chains (R = 2-propenyl, allyl, phenylpropynyl) were subjected to catalytic hydrogenation at this stage. For **5** and **10**, where R = 2-propenyl, hydrogenation of the disubstituted homoallylic olefins using palladium on carbon was extremely slow and accompanied by double-bond migration. Diimide could be used to effect the hydrogenation on a 1-mmol scale, but on a larger scale, low yields and incomplete reactions were realized. Best results were achieved with the homogeneous catalyst $Ir(cod)(py)(Cy_3P)PF_6$, which led to rapid and quantitative hydrogenation (Stork & Kahne, 1983). Furthermore, with use of the iridium catalyst the benzyl protecting group within **5**, needed for subsequent steps, remained intact.

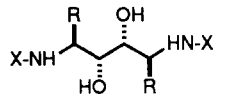
Table I: Inhibition of HIV-1 Protease by Pseudosymmetric Monools[a]



cmpd	X	R	K_i (nM) MENDT	AEDBN
I	Cbz-Val	Et	50 000	nd
II	Cbz-Val	*i*-Bu	80 000	2000
III	AlaAla	*i*-Bu	80 000	nd
IV[b]	Cbz-Val	Bn	230	13

[a] Assays were conducted either in MENDT buffer (pH 6.0, 0.2 M NaCl, 37 °C) or in AEDBN buffer (pH 4.7, 1.0 M NaCl, 30 °C), using the substrate AcArgAlaSerGlnAsnTyrProValVal-NH_2. Inhibition constants (nd, not done) were determined to within SEM ≤ 15% by Dixon analysis or (for IV in AEDBN buffer) by methods for tight-binding inhibitors (Dreyer et al., 1992). [b] Identical to the inhibitor A74704 of Erickson et al. (1990).

Table II: Inhibition of HIV-1 Protease by Symmetric Diols[a]



cmpd	X	R	K_i (nM) MENDT	AEDBN
V	Cbz-AlaAla	Et	5800	1400
VI	Cbz-AlaAla	*i*-Bu	580	30
VII (SKF108361)	AlaAla	*i*-Bu	800	nd
VIII	Cbz-Ala	*i*-Bu	2400	1600
IX	Cbz-Val	*i*-Bu	11	2.9
X	Cbz-Val	*n*-Bu	4.5	1.1
XI[b]	Cbz-Val	Bn	2.9	0.73
XII	Cbz-Val	$Ph(CH_2)_3$	8.6	nd
XIII	Cbz	$Ph(CH_2)_3$	14300	nd
XIV	Cbz-Val	$PhCH_2OCH_2$	180	nd

[a] Assays were conducted either in MENDT buffer (pH 6.0, 0.2 M NaCl, 37 °C) or in AEDBN buffer (pH 4.7, 1.0 M NaCl, 30 °C), as described in Table I (nd, not done). Inhibition constants were determined with SEM ≤ 15% by Dixon analysis or, for those compounds with K_i ≤ 11 nM, by methods for tight-binding inhibitors. [b] Identical to compound **24** in Kempf et al. (1990).

Introduction of the amino groups within **8** and **11** was accomplished by S_N2 displacement of the dimesylate derivatives of **5** and **10** with sodium azide in dimethylformamide or DMSO at elevated temperature, followed by reduction of the azido groups. The yield of the azide displacement reaction was dependent on the identity of the group R, ranging from nearly quantitative when R = Me to 50–60% when R = Ph due to a competing E_2 process (Chenera et al., 1991). In cases where elimination byproducts were formed, they were most conveniently removed following azido group reduction but prior to removal of the benzyl (in **7**) or acetonide (in **11**) protecting groups. Reduction of the azido groups could be carried out either by catalytic hydrogenation or with $LiAlH_4$, although the latter procedure was of course necessary in order to retain the benzyl protecting group within **7** and the P1/P1′ side chains when R = OBn.

HIV-1 Protease Inhibition. Several examples of the monool and diol structures were prepared with varying P1/P1′ side chains[2] and flanking aminoacyl substituents. These compounds were examined in our standard assay for inhibition of

HIV-1 protease (MENDT buffer, pH 6.0, 0.2 M NaCl, 37 °C) (Meek et al., 1989; Dreyer et al., 1989, 1992; Hyland et al., 1991). Monools possessing P1/P1′ ethyl or isobutyl side chains all exhibit weak competitive inhibition with K_i values of 50–80 μM (Table I). Incorporation of benzyl side chains affords the moderately potent inhibitor IV, identical to the inhibitor A74704 described by Erickson et al. (1990). This narrow tolerance for side-chain variation indicates that the three-carbon separation between P1 and P1′ may be suboptimal for spanning the S1 and S1′ pockets. Models built from the X-ray crystallographic structure of the HIV-1 protease complex with A74704 (Erickson et al., 1990) suggest that ethyl and isobutyl side chains (I–III) inadequately fill the S1/S1′ specificity pockets. In addition, the isobutyl groups in II and III appear to encounter unacceptably close contacts with Leu23 and Ile84 in the S1 pocket, and Leu23′ and Ile84′ in S1′.

The symmetric diol inhibitors are 2–4 orders of magnitude more potent than the corresponding monools under identical assay conditions using MENDT buffer (Table II). The greater potency of symmetric diols relative to analogous monools was previously noted by Kempf et al. (1990). A greater range of P1/P1′ groups is acceptable within the diol structure, with isobutyl, *n*-butyl, benzyl, and phenylpropyl side chains all affording potent inhibitors with K_i values in the range 3–11 nM. Potency is increased with increasing length and hydrophobic bulk of the P1/P1′ side chains and is markedly dependent on the identity of the P2/P2′ residue, with valine providing a distinct potency advantage over alanine. A similar trend was seen in a series of hydroxyethylene isosteres (Dreyer et al., 1992), highlighting the binding advantage of fully occupied hydrophobic specificity pockets. The large capacity of the S1/S1′ pockets is demonstrated with XII possessing the longer phenylpropyl side chains. A nearly continuous hydrophobic cleft connects the S1 and S3 (and the S1′ and S3′) binding sites, accommodating extended P1/P1′ side chains (Miller et al., 1989; Thompson et al., 1992). However, XIV, with ether-linked P1/P1′ side chains, is isosteric with its all-carbon analogue XII yet is 20-fold less potent. This potency loss may reflect the energy cost of desolvation upon transfer of the ether oxygens into the S1/S1′ cavities, further underscoring the importance of the hydrophobic effect for inhibitor binding.

Previously reported inhibition constants for IV and XI were 4.5 and 0.22 nM, respectively, obtained under conditions different from our standard assay using MENDT buffer (Erickson et al., 1990; Kempf et al., 1990; Matayoshi et al., 1990). To allow a better comparison of our data with that of Kempf, Erickson, and co-workers, we evaluated several inhibitors using a modified assay (AEDBN buffer, pH 4.7, 1.0 M NaCl, 30 °C) designed to replicate the conditions used in their previous reports. In AEDBN buffer, inhibition is improved for all compounds and the K_i values for IV and XI are similar to the previously published data (Tables I and II). These results are consistent with studies showing that peptide substrates and pepstatin A bind most favorably to HIV-1 protease at pH 4–5 and at high ionic strength (Billich et al., 1990; Hyland et al., 1991; Tropea et al., 1992). Moreover, relatively small variations in the K_i values for IV and XI are observed in the pH range 3.5–6.0 (Table III), indicating that the dominant influence is likely to be the ionic strength. As discussed by Tropea et al. (1992), high solution NaCl content may increase enzyme/inhibitor hydrophobic interactions through a salting-out effect. We have consistently used MENDT buffer (pH 6.0, 0.2 M NaCl) or similar buffers in studies of the inhibition and mechanism of action of HIV-1 protease despite the somewhat lower catalytic efficiency under these conditions, on the assumption that moderate hypertonicity and a pH slightly below neutrality may best approximate "physiological" conditions likely to pertain to maturing virions in cell culture and in vivo.

Table III: Enzyme Selectivity of Pseudosymmetric Monool and Symmetric Diol Inhibitors[a]

	K_i (nM)			
	HIV-1 protease		pepsin	
cmpd	pH 3.5	pH 4.5	pH 3.5	pH 4.5
IV	360	180	40	35
XI	2.2	2.3	110	200

[a] For HIV-1 protease, assays were conducted in GAMNT-EDT buffer (37 °C) as described in Table I; IV was shown to be a linear competitive inhibitor at pH 3.5 and 4.5 Assays for porcine pepsin were conducted in 40 mM sodium formate buffer (37 °C) using the substrate PheGlyHis[4-nitro-Phe]PheAlaPhe-OMe, and apparent K_i values were determined by Dixon analysis, assuming competitive inhibition.

Enzyme Specificity of Pseudosymmetric and Symmetric Inhibitors. We examined the monool IV and the diol XI for their ability to inhibit a prototypical cellular aspartic protease, porcine pepsin. Contrary to expectations, as shown in Table III, both compounds are potent pepsin inhibitors when evaluated using a standard spectrophotometric pepsin assay (Medzihradszky et al., 1970). IV is significantly more effective as an inhibitor of pepsin than of HIV-1 protease at both pH values examined. Only small pH effects are observed in each case, although for inhibition of pepsin by XI, and of HIV-1 protease by IV, inhibition is greatest near the pH optimum of the respective enzyme (pH 1–3 for pepsin and pH 4.5–5.0 for HIV-1 protease), as expected for an inhibitor that binds the catalytically competent form of the enzyme. Our present results are in conflict with a previous report that IV at concentrations of up to 10 μM exhibited no inhibition of pepsin (Erickson et al., 1990); the reasons for this discrepancy are unknown.

Structure of HIV-1 Protease Complexed with a Symmetric Diol. X-ray crystallographic analysis unexpectedly revealed that the symmetric diol VII binds asymmetrically to HIV-1 protease. Although the electron density for the inhibitor is highly symmetric about the pseudo 2-fold axis of the protein, it is diffuse and the diol is evidently present in two equally populated orientations, as was previously observed with a series of hydroxyethylene isostere complexes (Dreyer et al., 1992) and with acetylpepstatin (Fitzgerald et al., 1990). The two copies of the inhibitor are related by an almost perfect 2-fold axis and can be superimposed with an rms deviation of 0.06 Å, but the individual copies are not themselves C_2-symmetric. With the alternative least-squares superposition (corresponding to an approximate 180° rotation of one copy of the inhibitor), the rms deviation between the two copies is 0.62 Å. The conformation of one half of the inhibitor, which we call the P3–P1 portion,[2] closely resembles the corresponding amino-terminal (P3–P1) portion of the hydroxyethylene isosteres (Dreyer et al., 1992). The other half is more divergent, and we call this (somewhat arbitrarily) the P1′–P3′ region of the inhibitor.[2] The contact distances of the two copies of the inhibitor with the enzyme are not identical but are very similar, and for convenience we will limit the discussion to one of the copies.

The inhibitor binds in an extended conformation in the active site, and all of the inhibitor heteroatoms are involved in hydrogen bonds with the Ψ loops and flap region of the

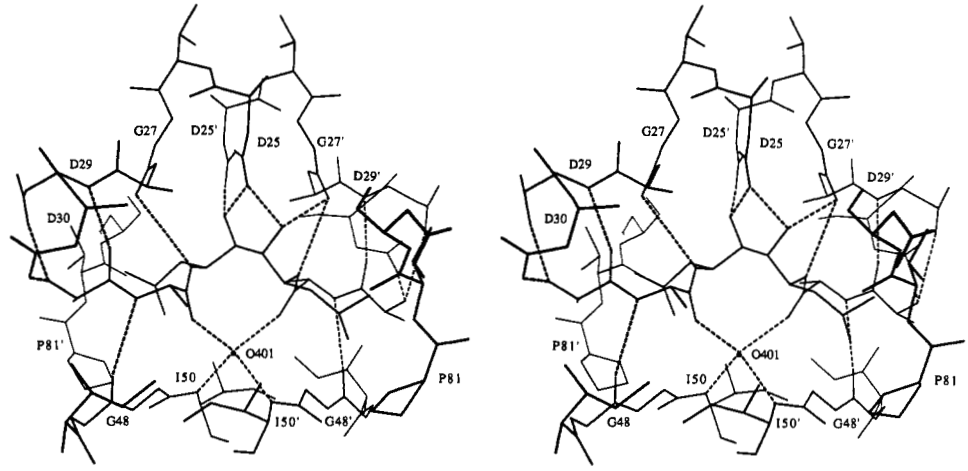


FIGURE 2: Structure of the symmetric inhibitor VII bound in the active site of HIV-1 protease. Active site amino acid residues forming potential hydrogen bonds with the inhibitor are labeled by their one-letter code designations. Potential hydrogen bonds are indicated by dotted lines.

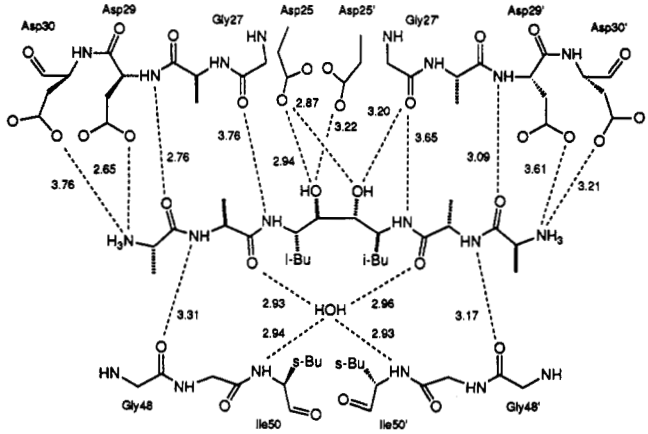


FIGURE 3: Distances for potential hydrogen bonds in the complex of VII.

protease dimer (Figure 2). One hydroxyl group of the diol is situated next to the pseudo 2-fold symmetry axis of the complex within hydrogen-bonding distance of the carboxyl groups of both Asp25 and Asp25′. The second hydroxyl group of the inhibitor is within hydrogen-bonding distance of only one of the aspartates (Asp25), but it also appears to be able to hydrogen bond with the carbonyl of Gly27′ (3.20 Å). The midpoint of the inhibitor is thus offset nearly 0.8 Å from the pseudo 2-fold axis, resulting in more intimate protein/inhibitor hydrophobic contacts within the S1′ pocket than within S1. The closest contact between the P1′ isobutyl side chain and the side chain of Pro81 is 3.74 Å, while the closest contact between P1 and Pro81′ is 5.09 Å.

The carbonyl groups of Gly27 and Gly27′ can form hydrogen bonds with the P1 and P1′ amino groups of the inhibitor, although the distances are very long (Figure 3). As observed in other HIV protease/inhibitor complexes, the P2 and P2′ carbonyl groups of the inhibitor point toward a water molecule (water 401), which then forms a bridge to the NH groups of Ile50 and Ile50′ [Miller et al. (1989), Erickson et al. (1990), and Bone et al. (1991), and references therein]. The complex is further stabilized by hydrogen bonds from the Asp29 and Asp29′ NH groups to the carbonyls of P3 and P3′, and from NH of P2 and P2′ to the carbonyls of Gly48 and Gly48′. As with the hydroxyethylene isosteres (Dreyer et al., 1992), the free terminal amino groups of the inhibitor appear to interact with the carboxyl side chains of Asp29, Asp30, Asp29′, and Asp30′.

The conformation of the protein within the diol-protease structure exhibits only minor deviations from symmetry. The coordinates of the two monomers superimpose with an rms deviation of 0.12 Å. The structure of the protein is very similar to that of previous structures, including the hydroxyethylene isostere complexes (Dreyer et al., 1992) and the complex with the monool A74704 (Erickson et al., 1990). The rms deviation between the protein backbone atoms in the complexes with VII and A74704 is 0.54 Å; the largest deviations between these two structures occur between side chains that are on the surface of the protein, specifically Glu34, Arg41, Lys55, and Arg57. The latter two residues are disordered in the diol structure, as seen by their poor electron density. The asymmetry between the protein monomers that appears in the flap region of several previous structures (Miller et al., 1989; Fitzgerald et al., 1990; Jaskólski et al., 1991), in which the carbonyl of Ile50′ forms a hydrogen bond with the amide of Gly51, is not present in our diol-protease model. As with our previous hydroxyethylene isostere complexes, the flaps of the diol-protease structure exhibit near 2-fold symmetry.

Modeling the Asymmetry. In an attempt to model the forces stabilizing the asymmetric binding mode of the symmetric inhibitor, molecular dynamics calculations were performed on two model complexes of VII with HIV-1 protease, in which the inhibitor was initially positioned either symmetrically or asymmetrically with respect to the protein. Energy stabilized after 50 ps for each system. The systems could not be distinguished by total energy, since the energetic differences due to the inhibitors are small compared to the sum of the water/water, protein/water, and protein/protein interaction energies. In order to examine the interactions of the inhibitor with the surrounding system, the "anal" module of AMBER was used to separate the total energy into contributions from five groups: those amino acids of the protein that were allowed to move in the simulation; the inhibitor; water 401 (numbering as in Figure 2); those water molecules and counterions that were allowed to move during the simulation; those amino acids, water molecules, and counterions whose positions were fixed during the simulation. A plot of the sum of the energies of interaction of the inhibitor with these five groups shows that the asymmetric trajectory is stable, maintaining a constant average energy (Figure 4A), while the symmetric system undergoes a transition to higher energy at about 200 ps (Figure 4B). A plot of the rms deviation of the central diamino diol portion of the inhibitor from its conformation in the crystal structure also shows that the asymmetric system is stable (Figure 4C) and that the

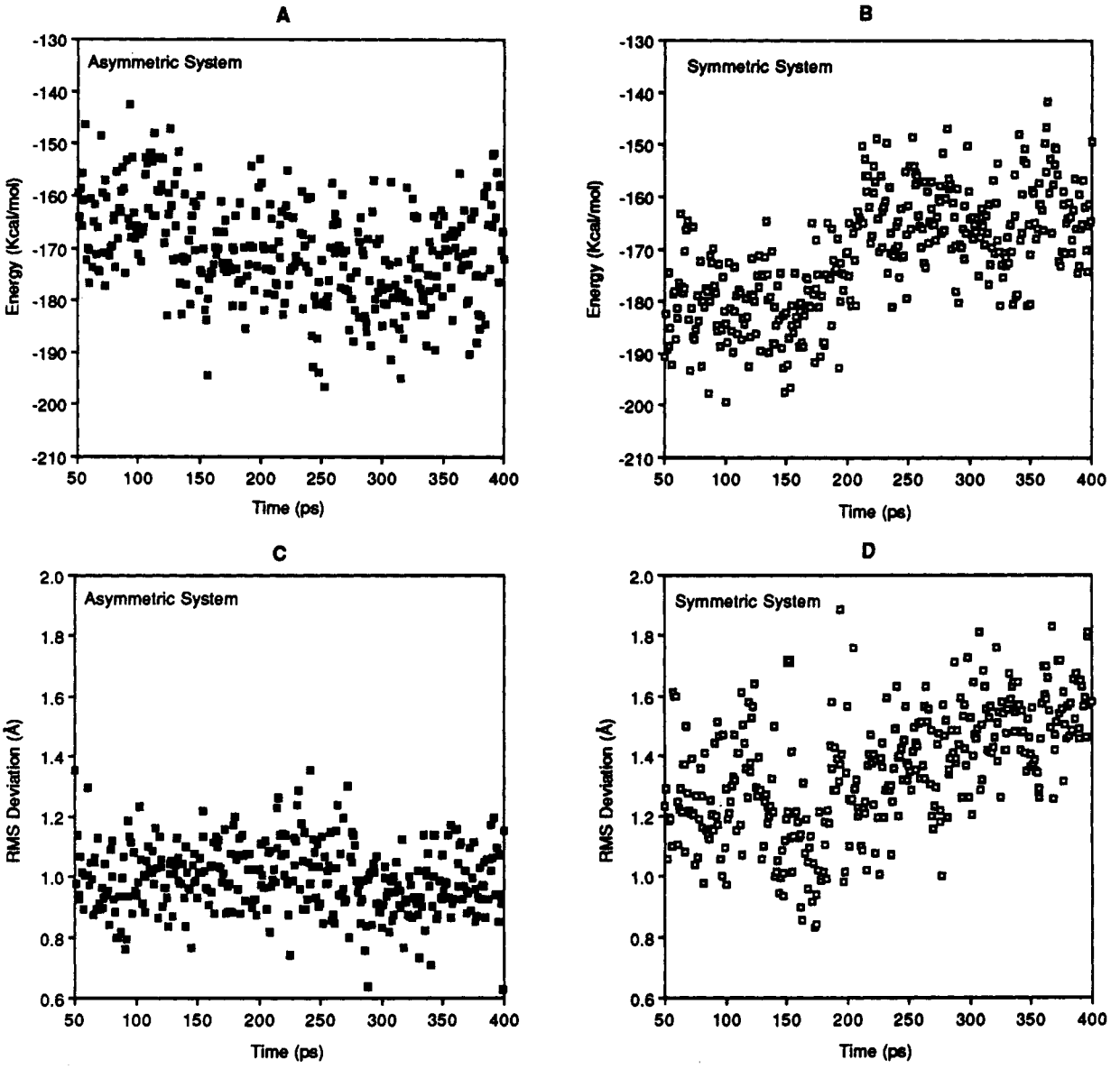


FIGURE 4: Molecular dynamics simulations of the complex of HIV-1 protease with VII. Total energy: (A) asymmetric, (B) symmetric starting structure. rms deviation between calculated and observed: (C) asymmetric, (D) symmetric starting structure.

symmetric system undergoes a transition at about 200 ps (Figure 4D). Although the differences in energies at the beginning and end of the calculations are small compared with thermal energy fluctuations on the picosecond time scale, some qualitative observations can be made. First, the symmetric system diverges from the crystal structure as the calculation progresses while the asymmetric system remains similar to the crystal structure. The transition seen with the symmetric system in panels B and D of Figure 4 corresponds structurally to the loss of a hydrogen bond between water 401 and one of the P2/P2′ (Ala) carbonyl groups of the inhibitor. The water molecule forms a new hydrogen bond to a second water molecule, but the inhibitor carbonyl group is left without a hydrogen-bonding partner.

DISCUSSION

A benefit of C_2-symmetric inhibitors from the standpoint of crystallography is that they could bind HIV-1 protease in a symmetric fashion and hence remove the 2-fold static disorder exhibited by asymmetric inhibitors (Fitzgerald et al., 1990; Dreyer et al., 1992; Thompson et al., 1992). One pseudosymmetric inhibitor was reported to bind nearly symmetrically to HIV-1 protease (Bone et al., 1991), while the complex with another pseudosymmetric inhibitor, A74704 (compound IV), showed significant deviations from symmetry (Erickson et al. 1990). As seen with the diol VII, the complex of even a completely symmetric ligand with a symmetric receptor need not possess a symmetric minimum-energy conformation. The asymmetric binding mode of VII could be induced by crystal packing forces, but this is unlikely since we observe insignificant deviations from 2-fold symmetry for the protein, and the inhibitor is well-contained within the active site. The molecular dynamics simulations further support the view that the asymmetric binding mode represents a lower energy complex in solution.

Although initially surprising, a breakdown of symmetry can be expected to arise relatively frequently in the interactions between symmetric molecules. A simple illustration is provided by the energy of a system of four identical collinear atoms interacting with a Lennard-Jones 6/12 potential, two of which are joined by a mobile rigid rod (B_1–B_1', B_2–B_2', or B_3–B_3') and two of which are fixed (A, A′) (Figure 5). In this system the rod will adopt a symmetric position only when it is sufficiently long that both ends are within a minimum

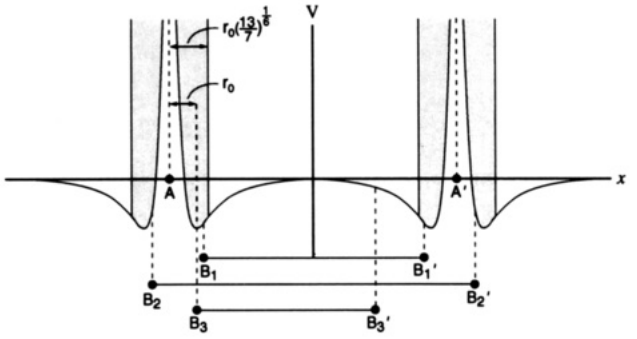


FIGURE 5: Van der Waals potential for a system of four identical collinear atoms. The solid curves represent the Lennard-Jones 6/12 potential functions for the fixed atoms A and A′: $V = \epsilon[(r_o/r)^{12} - 2(r_o/r)^6]$, where r_o is the distance to the energy minimum, ϵ is the depth for each well, and r is the distance $x - A$ (or $x - A'$). A second pair of identical atoms B and B′ joined by a rigid rod collinear with A and A′ with adopt a symmetric position if and only if B and B′ are within the shaded regions.[3] Thus, the midpoints of B_1–B_1' and B_2–B_2' are at $x = 0$, while the midpoint of B_3–B_3' is at a position $x < 0$.

distance from the fixed points, equal to $(13/7)^{1/6}$ times their van der Waals radius (i.e., B_1–B_1' and B_2–B_2' are symmetric and B_3–B_3' is asymmetric). A similar result can be derived for a model of hydrogen bonding.[3] This simple model suggests a mechanism by which the asymmetric position of the inhibitor may be favored: Optimization of some interactions occurs at the expense of other interactions. We speculate that the observed asymmetric binding mode might be an expected outcome when the inhibitor is imperfectly optimized for its binding site. VII is one of the weaker diol inhibitors (Table II), due in large part to its P2/P2′ alanine side chains which incompletely fill the S2/S2′ hydrophobic pockets, and it is possible that a more potent diol would bind more symmetrically.

Thanki et al. (1992) determined the X-ray structure of HIV-1 protease complexed with a (*R*,*R*)-dihydroxyethylene isostere inhibitor. Like the complex with VII, the P1 hydroxyl group of the isostere interacts with the carboxyl groups of both Asp25 and Asp25′, while the P1′ hydroxyl group can only form a hydrogen bond with Asp25. However, unlike VII, both hydroxyl groups of the dihydroxyethylene isostere are on the P1 side of the plane defined by the aspartate pair. The glycol moiety (HOCCOH) of each inhibitor exibits a gauche conformation, with a dihedral angle of +50° for the dihydroxyethylene isostere and −53° for VII. The dihydroxyethylene isostere glycol thus corresponds roughly to a twisted trans amide conformation, while that of VII more closely resembles a twisted cis amide. It remains to be seen whether an inhibitor bearing a central glycol will be found in which the glycol binds symmetrically. Potentially, optimization of the interaction of a single hydroxyl group with the aspartate pair represents a sufficient driving force to favor an asymmetric position of the glycol.

A major motivation for the design of symmetric inhibitors is the expectation that they should exhibit high selectivity for the C_2-symmetric retroviral proteases vs the related, unsymmetrical cellular aspartic proteases. In spite of this, little has been reported concerning the selectivity of symmetric inhibitors, although the monools and diols were reported to be highly selective (Erickson et al., 1990; Kempf et al., 1990). We were therefore surprised to find that IV and XI are potent pepsin inhibitors (Table III). In retrospect the promiscuity of inhibitors IV and XI is easily rationalized based on the structural homology that exists between the monomeric cellular aspartic proteases and the dimeric retroviral proteases (Pearl & Taylor, 1987; Rao et al., 1991). Conservation of both sequence and structure is greatest in the region of the active site Ψ loops (residues 29–42 and 212–225 in pepsin, 22–35 and 22′–35′ in HIV-1 protease) containing the pair of catalytic Asp-Thr-Gly sequences which are related by a near-perfect 2-fold symmetry axis in both enzymes. The monomeric aspartic proteases may even have evolved from an ancestral homodimeric enzyme by gene fusion (Tang et al., 1978), and inversely, as discussed by Rao et al. (1991), modern-day retroviral proteases may have arisen from the monomeric enzymes by gene deletion. The apparently close relatedness between pepsin and HIV-1 protease is also reflected in their similar substrate specificities, both enzymes exhibiting a marked preference for aromatic P1 and P1′ residues and frequent use of hydrophobic residues within the remaining subsites (Powers et al., 1977; Poorman et al., 1991).

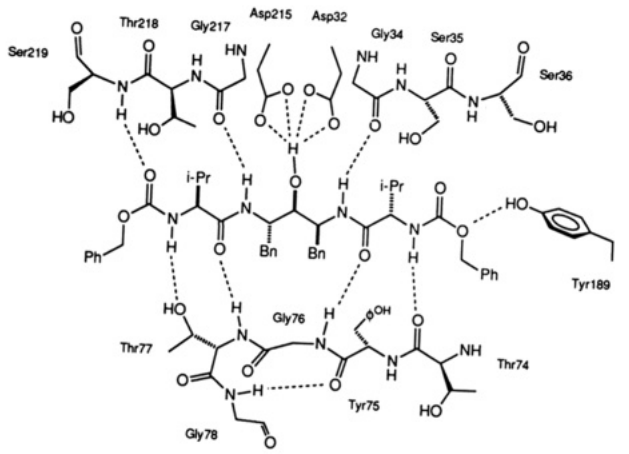


FIGURE 6: Hypothetical hydrogen-bonding interactions of the pseudosymmetric inhibitor IV with the active site of porcine pepsin.

X-ray structural data are available for porcine pepsin complexed with a peptide analogue inhibitor, A62095, composed of P4–P1′ residues (Chen et al., 1992). From this complex and several related microbial pepsin/inhibitor complexes (Bott et al., 1982; Foundling et al., 1987; Veerapandian et al., 1992; James et al., 1992), we can predict binding interactions for symmetric inhibitors with porcine pepsin (Figure 6). The structure of IV (A74704) as found in the HIV-1 protease/A74704 complex (Erickson et al., 1990) can be docked without alteration into the pepsin/A62095 structure, with reasonable distances for the potential hydrogen bonds shown in Figure 6 (<3.5 Å between heavy atoms). Several adjustments to the A74704 structure improve the fit. The main adjustments concern the Cbz groups, which we rotated substantially about the O–CH_2Ph bonds (ca. 100° for P3 and 120° for P3′) in order to take advantage of the S3 and S3′ binding pockets in pepsin. In this model, the interactions with the two Ψ loops are similar to those found in HIV-1 protease [Erickson et al. (1990); Figure 2], except for the indicated hydrogen bond to the hydroxyl group of Tyr189 (Bott et al., 1982). In Figure 6, rather than interacting with a bridging water molecule, the valine carbonyls of IV interact directly with Gly76 and Thr77 in the single flap of pepsin (Foundling et al., 1987; Chen et al., 1992). The six side chains

[3] Sketch: For pairwise van der Waals interactions, let $V(r) = \epsilon[(r_o/r)^{12} - 2(r_o/r)^6]$ as in Figure 5. Define $r_s = {}^1/_2((A' - A) - (B' - B))$ and $V_t(r) = V(r) + V(2r_s - r)$. Then $dV_t(r_s)/dr = 0$, and $d^2V_t(r_s)/dr^2 > 0$ if and only if $|r_s| < r_o(13/7)^{1/6}$. A similar result for a model of hydrogen bonding follows by using a potential of the form $V(r) = C/|r|^{12} - D/|r|$ (C, $D > 0$).

of IV are accommodated within identifiable S3/S3′ binding pockets of pepsin. Although A62095 lacks P2′-P3′ residues, cavities are readily discerned in the docked model for S2′ (Tyr75, Ile73, Ser35) and S3′ (Tyr189, Ile128, Gly188) (Sielecki et al., 1990).

Despite substantial differences in the topography of functional groups within the active sites of HIV-1 protease and pepsin, just as HIV-1 protease will accept asymmetric peptide substrates that resemble good pepsin substrates, pepsin is evidently sufficiently "pseudosymmetric" to accommodate symmetric inhibitors designed for HIV-1 protease. Hence the exploitation of symmetry is not sufficient to ensure selectivity. From an alternative point of view, the asymmetric binding mode of VII belies the intent of its design and allows us to distinguish "P" and "P′" halves of the inhibitor by analogy to hydroxyethylene isosteres.[2] The relationship of the P1′–P3′ half of VII to the corresponding region of a hydroxyethylene isostere is formally that of a retropeptide (Goodman & Chorev, 1979). The spatial orientations of the P1′–P3′ hydrogen-bonding groups and side chains are approximately maintained by simultaneously reversing the amide directionality, inverting the configuration at each α-carbon, and rotating 180° about the axis of the backbone. For both pepsin and HIV-1 protease, the inhibitory potency of hydroxyethylene isosteres is successfully maintained with these partial retro analogues. Perhaps retropeptides—either by themselves or as part of other structures—could have more general potential as novel derivatives of known receptor binding peptides.

ACKNOWLEDGMENT

We thank Steven A. Carr and the Mass Spectrometry Facility of SmithKline Beecham Pharmaceuticals for mass spectral analyses, Edith A. Reich for elemental analyses, and Michael Minnich and Jeffrey Culp for providing purified HIV-1 protease. We thank John Erickson and Alexander Wlodawer for providing the coordinates of the HIV-1 protease complexes with A74704, MVT101 and JG365 before they appeared in the Brookhaven Protein Data Bank. We also thank Sherin Abdel-Meguid and Brian W. Metcalf for helpful discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

Spectral and analytical characterization data for all intermediates in the synthesis of compounds I and IX and for all final inhibitors (4 pages). Ordering information is given on any current masthead page.

REFERENCES

Babine, R. E., Zhang, N., Jurgens, A. R., Schow, S. R., Desai, P. R., James, J. C., & Semmelhack, M. F. (1992) *Bioorg. Med. Chem. Lett. 2*, 541–546.

Bessler, B. H., Merz, K. M., Jr., & Kollman, P. A. (1990) *J. Comput. Chem. 11*, 431–439.

Billich, A., Hammerschmid, F., & Winkler, G. (1990) *Biol. Chem. Hoppe-Seyler 371*, 265–272.

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

Bott, R., Subramanian, E., & Davies, D. R. (1982) *Biochemistry 21*, 6956–6962.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Chen, L., Erickson, J. W., Rydel, T. J., Park, C. H., Neidhart, D., Luly, J., & Abad-Zapatero, C. (1992) *Acta Crystallogr. B48*, 476–488.

Chenera, B., Boehm, J. C., & Dreyer, G. B. (1991) *Bioorg. Med. Chem. Lett. 1*, 219–222.

Dreyer, G. B., Metcalf, B. W., Tomaszek, T. A., Carr, T. J., Chandler, A. C., 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.

Erickson, J., Neidhart, D. J., VanDrie, J., Kempf, D. J., Wang, X. C., Norbeck, D., Plattner, J. J., Rittenhouse, J., Turon, M., Wideburg, N., Kohlbrenner, W. E., Simmer, R., Helfrich, R., Paul, D., & 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.

Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., Sibanda, B. L., Hemmings, A., Wood, S. P., Blundell, T. L., Valler, M. J., Norey, C. G., Kay, J., Boger, J., Dunn, B. M., Leckie, B. J., Jones, D. M., Atrash, B., Hallett, A., & Szelke, M. (1987) *Nature 327*, 349–352.

Ghosh, A. K., McKee, S. P., & Thompson, W. J. (1991) *Tetrahedron Lett. 32*, 5729–5732.

Goodman, M., & Chorev, M. (1979) *Acc. Chem. Res. 12*, 1–7.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Huff, J. R. (1991) *J. Med. Chem. 34*, 2304–2314.

Hyland, L. J., Tomaszek, T. A., Jr., Roberts, G. D., Carr, S. A., Magaard, V. W., Bryan, H. L., Fakhoury, S. A., Moore, M. L., Minnich, M. D., Culp, J. S., DesJarlais, R. L., & Meek, T. D. (1991) *Biochemistry 30*, 8441–8453.

Jadhav, P. K., & Woerner, F. J. (1992) *Bioorg. Med. Chem. Lett. 2*, 353–356.

James, M. N. G., Sielecki, A. R., Salituro, F., Hakayawa, K., & Gelb, M. H. (1992) *Biochemistry 31*, 3872–3886.

Jaskólski, 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.

Kempf, D. J., Norbeck, D. W., Codacovi, L., Wang, X. C., Kohlbrenner, W. E., Wideburg, N. E., Paul, D. A., Knigge, M. F., Vasavanonda, S., Craig-Kennard, A., Saldivar, A., Rosenbrook, W., Clement, J. J., Plattner, J. J., & Erickson, J. (1990) *J. Med. Chem. 33*, 2687–2689.

Kempf, D. J., Sowin, T. J., Doherty, E. M., Hannick, S. M., Codacovi, L. M., Henry, R. F., Green, B. E., Spanton, S. G., & Norbeck, D. W. (1992) *J. Org. Chem. 57*, 5692–5700.

Lambert, D. M., Petteway, S. R., Jr., McDanal, C. E., Hart, T. K., Leary, J. J., Dreyer, G. B., Meek, T. D., Bugelski, P. J., Bolognesi, D. P., Metcalf, B. W., & Matthews, T. J. (1992) *Antimicrob. Agents Chemother. 36*, 982–988.

Le Merrer, Y., Dureault, A., Greck, C., Micas-Languin, D., Gravier, C., & Depezay, J. (1987) *Heterocycles 25*, 541–548.

Matayoshi, E. D., Wang, G. T., Krafft, G. A., & Erickson, J. (1990) *Science 24*, 954–958.

Medzihradszky, K., Voynick, I. M., Medzihradszky-Schweiger, H., & Fruton, J. S. (1970) *Biochemistry 9*, 1154–1162.

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., 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.

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.

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., Stickler, J. E., Gorniak, J. G. & Debouck, C. (1989) *Biochem. Biophys. Res. Commun. 159*, 420–425.

Nakatsuka, M., Ragan, J. A., Sammakia, T., Smith, D. B., Uehling, D. E., & Schreiber, S. L. (1990) *J. Am. Chem. Soc. 112*, 5583–5601.

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

Plattner, J. J., & Norbeck, D. W. (1990) in *Drug Discovery Technologies* (Clark, C. R., & Moos, W. H., Eds.) pp 92–126, Ellis Horwood, England.

Poorman, R. A., Tomasselli, A. G., Heinrikson, R. L., & Kezdy, F. J. (1991) *J. Biol. Chem. 266*, 14554–14561.

Powers, J. C., Harley, A. D., & Myers, D. V. (1977) in *Acid Proteases: Structure, Function, and Biology* (Tang, J., Ed.) pp 141–157, Plenum Press, New York.

Rao, J. K. M., Erickson, J. W., & Wlodawer, A. (1991) *Biochemistry 30*, 4663–4671.

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

Seibel, G. L., Singh, U. C., Weiner, P. K., Caldwell, J., & Kollman, P. A. (1989) *AMBER 3.0, Revision A*, Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, CA 94143.

Sielecki, A. R., Fedorov, A. A., Boodhoo, A., Andreeva, N. S., & James, M. N. G. (1990) *J. Mol. Biol. 214*, 143–170.

Stork, G., & Kahne, D. E. (1983) *J. Am. Chem. Soc. 105*, 1072–1073.

Strickler, J. E., Gorniak, J., Dayton, B., Meek, T., Moore, M., Magaard, V., Malinowski, J., & Debouck, C. (1989) *Proteins 6*, 139–154.

Swain, A. L., Miller, 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.

Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A., & Blundell, T. L. (1978) *Nature 271*, 618–621.

Thanki, N., Rao, J. K. M., Foundling, S. I., Howe, W. J., Moon, J. B., Hui, J. O., Tomasselli, A. G., Heinrickson, R. L., Thaisrivongs, S., & Wlodawer, A. (1992) *Protein Sci. 1*, 1061–1072.

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.

Tropea, J. E., Nashed, N. T., Louis, J. M., Sayer, J. M., & Jerina, D. M. (1992) *Bioorg. Chem. 20*, 67–76.

Veerapandian, B., Cooper, J. B., Šali, A., Blundell, T. L., Rosati, R. L., Dominy, B. W., Damon, D. B., & Hoover, D. J. (1992) *Protein Sci. 1*, 322–328.

Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230–253.