Hydroxyethylamine Isostere of an HIV-1 Protease Inhibitor Prefers Its Amine to the Hydroxy Group in Binding to Catalytic Aspartates. A Synchrotron Study of HIV-1 Protease in Complex with a Peptidomimetic Inhibitor

Jan Dohnálek,* Jindřich Hašek, Jarmila Dušková, and Hana Petroková

Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovského nám. 2, 162 06 Praha 6, Czech Republic

Martin Hradilek, Milan Souček, and Jan Konvalinka

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Praha 6, Czech Republic

Jiří Brynda, Juraj Sedláček, and Milan Fábry

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Praha 6, Czech Republic

Received July 6, 2001

A complex structure of HIV-1 protease with a hydroxyethylamine-containing inhibitor Boc-Phe-Ψ[(*S*)-CH(OH)CH_2NH]-Phe-Gln-Phe-NH_2 has been determined by X-ray diffraction to 1.8 Å resolution. The inhibitor is bound in the active site of the protease dimer with its hydroxyethylamine isostere participating in hydrogen bonds to the catalytic aspartates 25 and 25′ and glycine 27′ of the active site triads via five hydrogen bonds. The isostere amine interactions with the catalytic aspartates result in a displacement of the isostere hydroxy group in comparison with the common position known for analogous hydroxyethylamine containing inhibitors. A comparison with another inhibitor of this series shows that the change of one atom of the P_2' side chain (Glu/Gln) leads to an altered ability of creating hydrogen bonds to the active site and within the inhibitor molecule. The diffraction data collected at a synchrotron radiation source enabled a detailed analysis of the complex solvation and of alternative conformations of protein side chains.

Introduction

The aspartic protease of the HIV-1 (human immunodeficiency virus) has become a target of AIDS (acquired immunodeficiency syndrome) treatment, and numerous types of its inhibitors have been suggested as drugs. It has been proved that mutated species of the viral protease resistant to inhibition develop in patients treated with PR-inhibitor based drugs.[1] The inhibitor Boc-Phe-Ψ[(*S*)-CH(OH)CH_2NH]-Phe-Gln-Phe-NH_2 (denoted as SQ) which belongs to a novel peptidomimetic series resistant to a specific set of drug induced mutations has been subjected to an X-ray structural study of its complexing ability (K_i = 33 nM for HIV-1, BRU isolate).[2,3] Its structural properties have been compared with another member of this series.[4]

Results

Inhibitor Binding. The inhibitor binds by 16 hydrogen bonds to the protease active site (two of them mediated by the buried water molecule W301) and makes another five hydrogen bonds to surrounding water molecules and three to two glycerol molecules in the vicinity of the active site (Figures 1 and 2). The hydroxyethylamine isostere binds to the Asp 25, Asp 25′, and Gly 27′ via five hydrogen bonds. The hydroxy oxygen shares its hydrogen in two hydrogen bonds to $O_{G27'}$ and $O_{\delta2,D25}$. The amino group of the isostere binds to one oxygen of Asp 25 $O_{\delta2,D25}$ and to both carboxy oxygens of Asp 25′. The isostere amine is localized in the position closest to the protease dimer axis. As the hydrogen bonds NH−Asp25, and NH−Asp25′ dominate in the isostere−catalytic aspartates interactions, the hydroxy group is shifted away from its common position known from other complexes of HIV protease with analogous hydroxyethylamine containing inhibitors (e.g.,

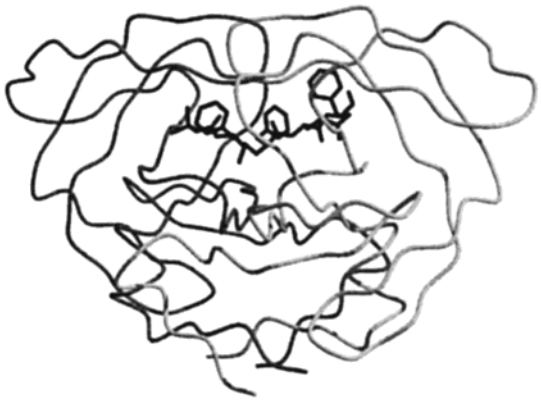

Figure 1. Position of SQ inhibitor molecule in the active site of HIV-1 protease dimer. Protein chains are represented as C_α trace, and the inhibitor is represented as a stick model of non-hydrogen atoms.

* Corresponding author. Tel: ++4202 20403205. Fax: ++4202 35357981. E-mail: dohnalek@imc.cas.cz.

10.1021/jm010979e CCC: $22.00
Published on Web 02/22/2002

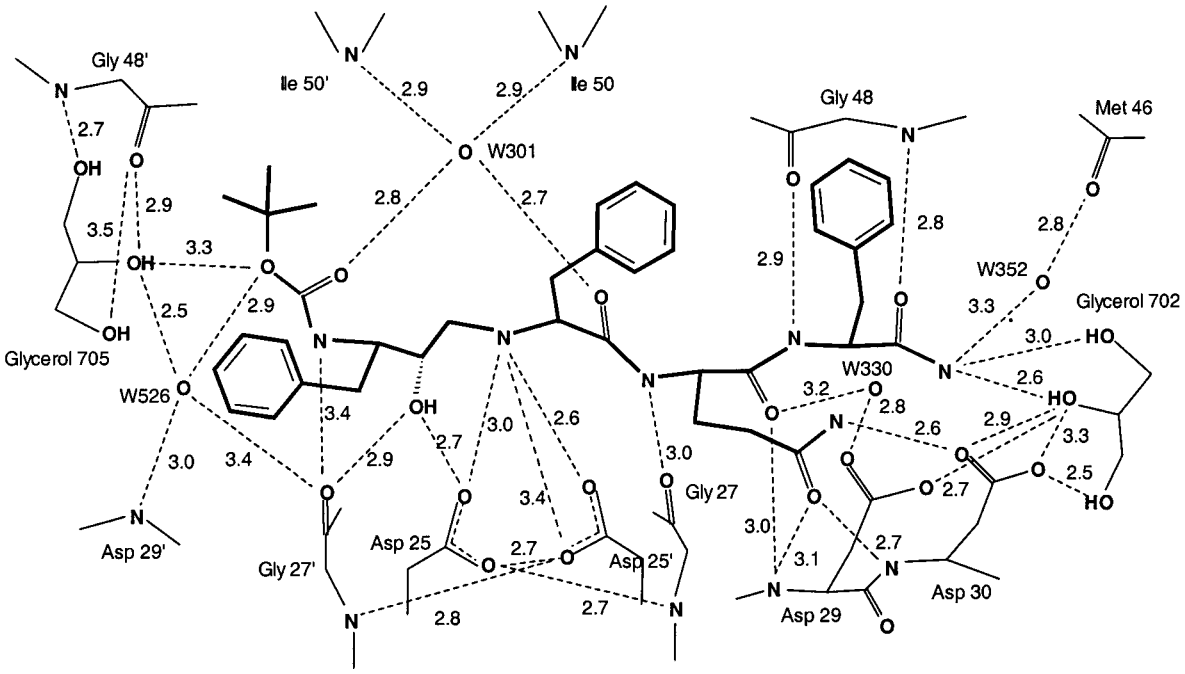

Figure 2. Schematic representation of the SQ inhibitor hydrogen bonds within 3.6 Å distance. Bonds of the type C−H···O are not shown.

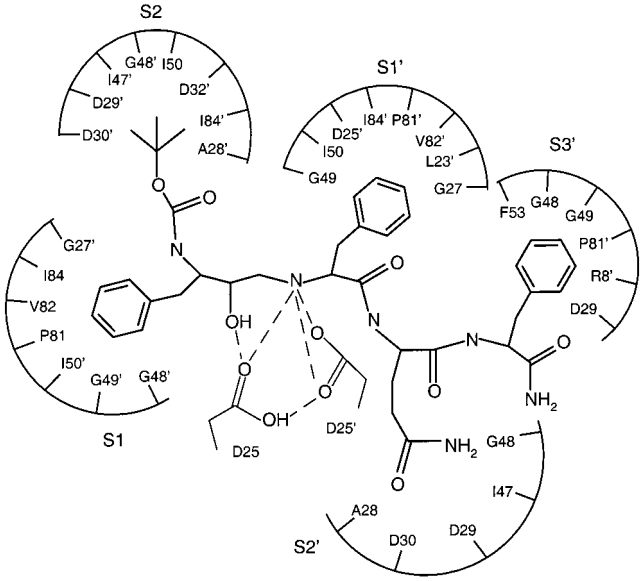

Figure 3. Schematic view of the protease active site pockets binding the inhibitor SQ. Protease residues within 4.1 Å distance from the inhibitor are shown.

Amprenavir,[5] PDB code 1HPV, or Saquinavir,[6] 1HXB). This shift propagates further along the inhibitor chain to C_α atoms of P_1 and P_1' residues and reflects in differences between S_1 and S_1' binding pockets of inhibitor side chains (Figure 3).

Water W301 mediates hydrogen bonds of P_2 and P_1' carbonyl oxygens to the main chain NH groups of isoleucines Ile 50 and Ile 50′ and is buried in a cavity between the tips of protease flaps and the inhibitor.

The P_2' glutamine makes six hydrogen bonds to Gly 27, Asp 29, and Asp 30 of the polypeptide chain and one bond to structurally important water W330 which interconnects protein chains, inhibitor, and surrounding solvent. The P_2' side chain hydrogen bonds to Asp 29 main chain and to Asp 30 both main chain and side chain demonstrate a good recognition of P_2' by the S_2' binding pocket. This leads to a very well characterized position of all P_2' side chain atoms and their significant contribution to the binding energy of the inhibitor.

The terminal groups of the inhibitor make hydrogen bonds to water and glycerol molecules bound to the surface of the complex. The glycerol 705 and water W526 bind to Boc urethane oxygen of the inhibitor. Another glycerol molecule 702 was localized near the NH_2 terminal of the inhibitor. It makes two hydrogen bonds to the end NH_2 which also participates in a hydrogen bond to water W352.

Numerous studies of HIV protease complexes have shown that the discussion of hydrogen bonding network alone is not sufficient for description of the inhibitor affinity to protease.[7−9] In an attempt to describe the other important factors influencing the affinity, all the inhibitor−protease contacts C−C closer than 4.1 Å and all C−O or C−N interactions closer than 3.6 Å are summarized in Table 1 and graphically represented in Figure 4. The number of C−C intermolecular contacts gives a rough estimate of the intensity of hydrophobic interactions between the inhibitor and the enzyme in the individual binding pockets that are schematically drawn in Figure 3. All the C−O and C−N contacts can be classified as hydrogen bonds but have not been included in Figure 2.[10]

The S_1 and S_1' sites accommodate the same groups—benzyls of P_1 and P_1'—but different amino acids of the pockets participate in contacts. C_α atoms of P_1 and P_1' are not positioned symmetrically with respect to the dimer pseudosymmetry axis. This causes a slightly different environment for accommodation of phenyl rings. As summarized in Table 1, the prevailing type of interaction for P_2, P_1, and P_1' is either hydrophobic or of type C−X (X = N, O), whereas P_2' (Gln) creates the highest number of hydrogen bonds (14, including close C−X contacts) of all inhibitor residues and the lowest number of C−C contacts, which is in agreement with

Table 1. Summarized Interaction Types between the Inhibitors SQ and SE, Protease and Solvent[a]

SQ Complex

	number of contacts					
interaction type	P_2	P_1	P_1'	P_2'	P_3'	total
C−C	12	21	18	4	11	66
C−X	16	21	10	26	22	95
H bond to protein + str. important solvent	2 + 2	9 + 1	8 + 1	13 + 1	8 + 0	40 + 5
H bond to other solvent	1	0	0	0	3	4
total (H bond/other)	5/28	10/42	9/28	14/30	11/33	49/161

SE Complex

	number of contacts					
interaction type	P_2	P_1	P_1'	P_2'	P_3'	total
C−C	11	6	13	6	6	42
C−X	7	17	12	13	17	66
H bond to protein + str. important solvent	6 + 1	8 + 0	7 + 1	14 + 0	8 + 0	43 + 2
H bond to other solvent	0	0	0	0	4	4
total (H bond/other)	7/18	8/23	8/25	14/19	12/23	49/108

[a] Structurally important solvent includes three waters: W301, W526, and W330.

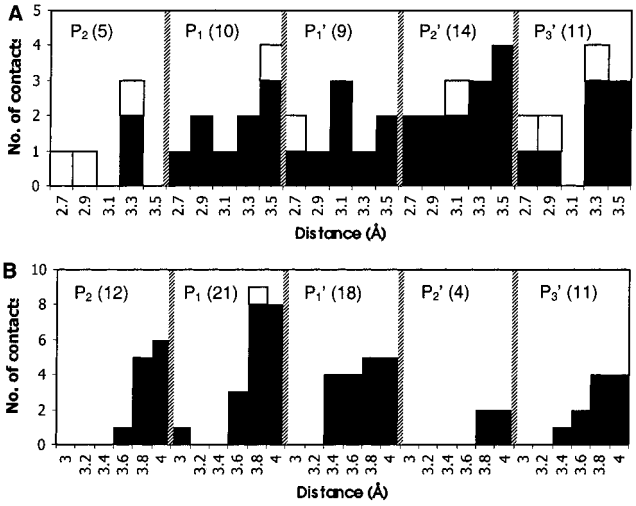

Figure 4. Distribution of (A) hydrogen bond and (B) C−C contact distances for the SQ inhibitor. All contacts within 3.6 and 4.1 Å, respectively, have been included and divided to 0.2 Å categories (2.6−2.8 Å, etc.). Filled columns represent inhibitor−HIV-1 protease contacts, blank columns inhibitor−solvent contacts. The total number of contacts of a given type is shown in parentheses.

its hydrophilic nature. The P_3' phenyl ring is partially exposed to solvent and is involved in fewer C−C interactions and creates numerous contacts of C−X type.

Atomic Displacement Parameters of the Inhibitor. The lowest values of atomic displacement parameters (temperature factors) in the inhibitor for Gln P_2' show its high stability caused by strong hydrogen bonds. In this respect, the hydroxyethylamine isostere is less firmly bound. The P_1 phenyl ring with the highest *B* values has poorly defined electron density for C_ϵ and C_ζ atoms. Its higher temperature factors can be caused by close contacts between the phenyl hydrophobic ring and surrounding solvent. The phenyl group of P_1' interacts with another phenyl group of P_3', and this disables the solvent contacts to P_1'. The *B* factor distribution on the inhibitor shows that the residue with the lowest displacement factors is not the hydroxyethylamine isostere but the tightly binding Gln P_2'.

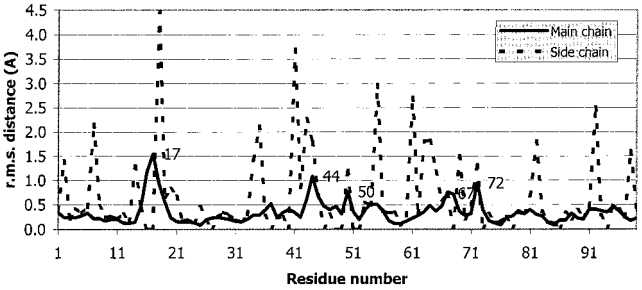

Figure 5. Rms distances between A and B chains of the protease dimer calculated for the best superposition of one subunit onto the other. The main chain rms distances close to the value of one point to the sections of the chain with major variance.

Inhibitor−Solvent Interaction. The inhibitor has been found to interact via hydrogen bonds with six solvent molecules including two glycerol molecules 702 and 705 and water W301 buried in the active site (Figure 2). The water W301 occupies its position typical for peptidomimetic inhibitor structures—stabilized by four hydrogen bonds to protein (Ile 50 and 50′) and to inhibitor carbonyls O_{P2} and $O_{P1'}$. The crystals of protease−inhibitor complex were soaked in 30% glycerol solution before flash freezing and so glycerol localized in electron density maps must have bound in the course of soaking. Water W330 participates in three hydrogen bonds to protein (Arg 8′, Asp 29, and Gly 27), one hydrogen bond to inhibitor ($O_{P2'}$), and two to water molecules W416 and W390.

Water molecules W390, W330, and W416 form a chain of solvent molecules that penetrates from the surface solvent region to the active site. It fills a cavity surrounded by protein chains A (main and side chain) and B (Arg 8′ side chain) and by the inhibitor. Water W526 occupies a site corresponding to the $O_{P2'}$ position on the other side of the tunnel. All solvent molecules in the vicinity of the inhibitor were modeled with full occupancy.

Dimer Symmetry. The asymmetric unit of the protease crystal in space group $P2_12_12$ consists of one dimer and thus allows asymmetry within the dimer with respect to its noncrystallographic 2-fold axis. Chemically identical protease chains A and B have different contacts to neighboring protease molecules. The plot of rms distances of the chains A and B superimposed in Figure 5 identifies the sections of the main chain that display the major conformational variances: 13−20, 35−55, and 63−73.

Conformational differences in β hairpins 13−20 and 63−73 are attributed to different interactions of A and B chains with neighboring protease molecules. The chain section 35−55 comprises residues of the tip of the protease flaps that interact directly with bound inhibitor (Ile 47, Gly 48, Gly 49, Ile 50, and Phe 53) as opposed to residues 13−20 and 35−55.

Protein Conformation. The protease dimer in the referred structure consists of two chemically identical polypeptide chains with 99 amino acids in each chain (sequence of the BRU isolate of the virus). Secondary structure elements assignment by program Procheck[11] is identical for A and B chains of the dimer and does

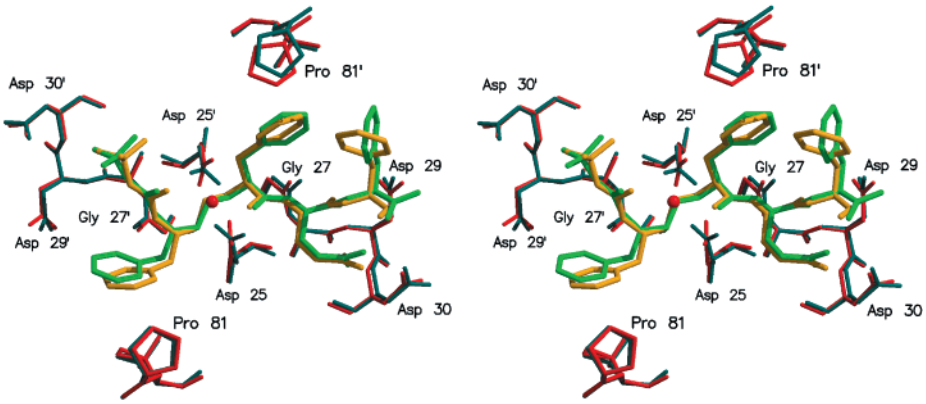

Figure 6. Stereoview of the protease active site; comparison of SE (green) and SQ (orange) inhibitors binding to the active site of HIV-1 protease. Protein chains for SE complex colored in green-blue, for SQ in red. Created with programs Molscript[23] and Raster3D.[24]

not show significant differences when compared to other 3D structures of HIV protease complexes.[12]

The protein conformation is partially stabilized by water molecules interconnecting protein chains. In total, 52 water molecules mediate interactions between protein main chains (interactions up to the neighbor residue of the same chain excluded). Another 46 water molecules mediate main chain−side chain interactions, and 23 water molecules mediate side chain−side chain interactions.

Flexibility of Side Chains. Nineteen amino acid residues have been found in alternative conformations: Thr 4, Glu 34, Ser 37 and 37′, Lys 45 and 45′, Ile 50 and 50′, Gly 51 and 51′, Ile 64, Cys 67 and 67′, Ile 72, Val 82, Leu 97, Glu 35′, Met 46′, Cys 95′.

Residues interacting with neighboring molecules include Thr 4, Glu 34, Ser 37, Lys 45 and 45′, Ile 72, Val 82, Glu 35′, and Met 46′. Alternative conformations have been observed also for residues mediating protein core interactions: Ile 64, Leu 97, and Cys 95′.

Cysteines 67 and 67′ are probably oxidized. Oxidation of cysteines of HIV-1 protease has been observed in a number of other structures. In this case, cysteines are found oxidized either as a result of chemical oxidation during protein purification or as a result of radiation damage. At least seven deposited HIV protease structures show unknown mutations at cysteine sites: PDB codes 1DMP and 1QBS mutation C97X, pdb codes 1CPI, 3HVP, 4HVP, 7HVP, and 8HVP mutations C67X and C95X. In this case, the end groups of the side chains cannot be easily resolved in the electron density maps. The alternative conformations correspond to two states: intact and oxidized. In both chains, similar geometry has been modeled as regular cysteine residues in two conformations with alternative sets of solvent molecules.

The hydrogen bond between peptide bonds 50−51 of A and B chains closes the active site by the protease flaps. As this bond is mediated by the carbonyl oxygen of one and the NH group of the other chain, the symmetry of the orientation of these peptide bonds with respect to the dimer pseudosymmetry axis is excluded. Two possible orientations have been found in this structure, which differ in the flipped peptide between residues 50−51. These alternative conformations of the main chain have been proved by the presence of a network of alternative hydrogen bonds to the surrounding solvent. The side chains of the isoleucines were not modeled alternatively.

Table 2. Comparison of Side Chain Binding Sites for SE and SQ Complexes[a]

S_2		S_1		S_1'		S_2'		S_3'	
SE	SQ	SE	SQ	SE	SQ	SE	SQ	SE	SQ
I50	I50	P81	P81	G27	G27	A28	A28	D29	D29
-	A28′	-	V82	-	G49	D29	D29	G48	G48
D29′	D29′	-	I84	I50	I50	D30	D30	-	G49
D30′	D30′	G27′	G27′	L23′	L23′	I47	I47	F53	F53
V32′	V32′	G48′	G48′	D25′	D25′	-	G48	-	R8′
I47′	I47′	G49′	G49′	P81′	P81′				P81′
G48′	G48′	-	I50′	V82′	V82′				
-	I84′	-	-	I84′	I84′				

[a] All contacts of inhibitor side chains to protease within 4.1 Å are accounted for.

Complex Hydration. A total of 301 water molecules has been localized mainly either in the surface clefts of the inhibitor−protease complex or close to the intermolecular contacts of symmetry-related protein dimers.

Continuous clusters of water molecules interconnected by hydrogen bonds have been analyzed. The 301 water molecules can be divided into 30 clusters and another 55 individual water molecules. The two largest clusters (groups of 51 and 44 water molecules interconnected via hydrogen bonds) lie close to the contacts between protease molecules in the crystal.

Discussion

SE−SQ Comparison. Least-squares fit of protease C_α atoms SE (PDB ID 1FQX)[13] to SQ gives a rms fit of 199 atoms of 0.56 Å. Positions of SE and SQ inhibitors bound in the active site are compared in Figure 6. The hydroxyethylamine isostere of both inhibitors binds almost identically with respect to the protease main chain. In contrary, all inhibitor side chains differ in their positions. Inhibitor side chains display differences in positions that cause variances in inhibitor−protease contacts as shown in Table 2.

The P_2 Boc of SE together with the P_1 residue are shifted closer to the S_2 binding site, and so the P_2 of SE loses C−X close contacts (number of contacts in SQ structure/number of contacts in SE structure: 16/7) and gains four potential hydrogen bonds to the protease (2/6).

The P_1 phenyl rings are found in similar positions with small shifts away from the protease surface of the

channel in the case of the SE complex. This shift leads to a significantly lower number of protease−inhibitor contacts P_1-S_1, mainly of C−C type (21/6).

The P_1' position and conformation is the most conserved one in the compared structures.

SE and SQ inhibitors differ only in a single non-hydrogen atom (Glu−Gln) of the P_2' residue. The structural comparison of the two complexes shows that Glu and Gln bind in a similar conformation conserving two hydrogen bonds $N_{\epsilon,P2'}-O_{\delta2,Asp30}$ (SQ, 2.6 Å), $O_{\epsilon1,P2'}-O_{\delta2,Asp30}$ (SE, 3.0 Å) and $O_{\epsilon,P2'}-N_{Asp30}$ (SQ, 2.7 Å), $O_{\epsilon2,P2'}-N_{Asp30}$ (SE, 3.0 Å). The bond $O_{\epsilon,P2'}-N_{Asp29}$ in the SQ complex (3.1 Å) does not occur in the case of the SE inhibitor.

The dihedral angles of the P_2' side chains have been compared: $\chi_1 = -178°$ (SE), 176° (SQ), $\chi_2 = -75°$ (SE), −79° (SQ), and $\chi_3 = 126°$ (SE), 157° (SQ). The values of P_2' dihedral angles display a significant difference only for χ_3 (C_β C_γ C_δ $C_{\epsilon1}$ (N_ϵ)). The carboxy group of SE P_2' is rotated by 31° with respect to the SQ amide. This rotation causes the loss of one of the hydrogen bonds to the protease main chain. On the other hand, this conformation of the Glu side chain makes possible an internal hydrogen bond, between the $O_{\epsilon1,P2'}$ and $O_{P3'}$ (distance in the SE complex of 2.5 Å). The corresponding distance $N_{\epsilon,P2'}-O_{P3'}$ in the SQ complex was found significantly longer, 3.4 Å.

Planes of the P_3' phenyl rings in SE and SQ structures are found in perpendicular orientations, and this results in significantly different protease−inhibitor contacts of these groups. This rotation results in a loss of both C−X and C−C contacts of the phenylalanine side chain in the SE complex when compared to SQ. The remarkable differences in P_3' torsion angles between SE and SQ complexes could be also attributed to the ability of the inhibitor to prefer distinct hydrogen bonding schemes of the P_2' side chain and the terminal amide group based on the nature of the P_2' side chain (polar-charged). Stronger bonds of Gln to the S_2' binding pocket result in a shift of the P_3' phenylalanine toward the surface of the enzyme, and the benzyl ring in the case of SQ complex has to adjust its orientation to avoid inconvenient close contacts to the protease surface.

The overall differences have been observed mainly in hydrophobic C−C contacts (66/42) and in C−X interactions (95/66). The P_1 residue shows the largest variance in its contacts and P_1' the smallest one. The total number of all types of hydrogen bonds has been conserved.

The carboxy groups of Asp 25 and Asp 25′ are rotated with respect to the SE complex by approximately 40° and 60° so that the $O_{\epsilon2,Asp25}$ creates closer hydrogen bond contact to the isostere hydroxy group (2.7 Å) than in the case of the SE complex (3.3 Å).

The compared structures in two different space groups show a higher degree of 2-fold dimer symmetry in $P6_1$ than in $P2_12_12$. However, the dimer interface is formed by identical networks of hydrogen bonds. Only small conformational changes of residues participating in dimer formation have been observed resulting in different interface surface areas. The dimer interface area (based on solvent accessible surface calculations) is 1853 Å^2 for the SE complex and 1833 Å^2 for SQ.

Biomedical Implications of Inhibitor Binding, Comparison within the Series. Two inhibitors with 200-fold difference in K_i values ($K_{i,SE}$ = 0.15 nM, $K_{i,SQ}$ = 33 nM) bind to the active site of HIV-1 protease in similar positions. The hydroxyethylamine hydrogen bond network and its position with respect to the protein main chain is close to identical in the two cases. Variances in benzyl groups binding probably do not contribute to the explanation of inhibition constants. Preliminary calculations show that 200 times higher inhibition constant for SQ complex in comparison with SE follows mainly from higher deformation energy for SQ inhibitor to fit the protease binding pocket.[14] P_2' hydrogen bonds vary between SE and SQ complexes. The carboxy group rotation in the case of SE inhibitor leads to a loss of one hydrogen bond when compared to the SQ structure and enables another hydrogen bond contact of P_2' carboxy oxygen to the terminal amide oxygen of the inhibitor. It has to be noted that SE and SQ structures differ significantly in the diffraction limits of experimental data (3.1 and 1.8 Å, respectively). The SE inhibitor is modeled in two disordered positions whereas the SQ molecule was found in one unique orientation. Lower *B* factors of P_2' residues in both structures indicate firm binding of this part of the molecule and its importance for the affinity of the inhibitor to the binding site.

The SQ complex structure proves that the hydroxyethylamine isostere prefers its amine in hydrogen bonding to both catalytic aspartates and thus the hydroxy group binds between Asp 25 and carbonyl oxygen of Gly 27. It has been shown in the case of SE complex that this position of inhibitor hydroxy group forces a shift of the protease Gly27 O atoms away from the enzyme dimer axis. The reported structure of the SQ inhibitor confirms this previous conclusion. The $O_{Gly27}-O_{Gly27'}$ distance of 8.8 Å is less than in the SE complex (9.6 Å) where the hydroxy oxygen was modeled on both sides of the active site, and it exceeds the distances, e.g., in the Saquinavir complex (7.8 Å) or in unliganded protease structures (in 3PHV 8.3 Å). The protease chain has to adjust on hydroxyethylamine isostere binding distinctly from, e.g., the hydroxyethylene one where the hydroxy group is not forced to bind between Gly 27 and Asp 25. The presence of a hydroxy group in the isostere concurrently with another hydrogen bond donor can imply inconvenient modification of the protease conformation.

Conclusion

The structure of the inhibitor Boc-Phe-Ψ[(*S*)-CH(OH)-CH_2NH]-Phe-Gln-Phe-NH_2 (SQ) in complex with HIV-1 protease has been determined at 1.8 Å resolution. The inhibitor binds to the active site of the enzyme in a conformation similar to the inhibitor SE.[13] The hydroxyethylamine isostere binds to Asp 25, Asp 25′, and Gly 27 preferring the amino group in contacts to the catalytic aspartates. The number of hydrogen bonds and relative *B* factor values of the P_2' Gln side chain support its significance for high affinity of the inhibitor to the enzyme.

The structure has been compared with a previously published complex with Boc-Phe-Ψ[(*S*)-CH(OH)CH_2-NH]-Phe-Glu-Phe-NH_2. These two molecules differ only

Table 3. Diffraction Data Statistics and Structure Parameters

A. Diffraction Data	
wavelength	1.00 Å
space group	$P2_12_12$
unit cell parameters	$a = 58.1$ Å, $b = 86.3$ Å, $c = 46.1$ Å $\alpha = \beta = \gamma = 90°$
diffraction limits (highest shell)	11.0−1.83 (1.88−1.83) Å
no. of measured diffraction maxima	188 100
no. of unique reflections (highest shell)	20 474 (1392)
$R_{symmetry}$ (highest shell)	0.11 (0.22)
data completeness (highest shell)	97.4% (93.7%)
I/σ_I (highest shell)	15.8 (4.9)

B. Structure Parameters	
no. of non-H atoms	1952
no. of non-H atoms in inhibitor	51
no. of refined water molecules	301
average temperature factor *B*, main chain atoms	10.7 Å^2
average temperature factor *B*, side chain atoms	16.8 Å^2
rmsd bond distances from ideal	0.008 Å
rmsd angles from ideal	1.51°
crystallographic residual *R* factor (test set, 5% of all reflections)	0.21
crystallographic residual *R* factor (working set)	0.17
crystallographic residual *R* factor (all reflections)	0.17

by one atom of the P_2' residue. The comparison yields the following conclusions: (1) The hydroxyethylamine isostere binds in its firmly determined position irrespectively of the crystallization and experimental conditions. (2) The residue of interest (P_2') differs in its third torsion angle giving the carboxy and amide groups diverse sets of hydrogen bonds. The carboxy group in the P_2' Glu participates in an internal hydrogen bond to the terminal amide oxygen.

The 200-fold difference in K_i values for these inhibitors can be explained by a suitable conformation for internal hydrogen bonds in the inhibitor both in solvent and in a complex with the enzyme.

Experimental Section

The inhibitor was prepared by alkylation of N-terminal amino group of tripeptide H-Phe-Gln(OBzl)-Phe-NH_2 with (2*R*,3*S*)-3-[(*tert*-butyloxycarbonyl)amino]-1,2-epoxy-4-phenylbutane at elevated temperature in a protic solvent followed by hydrogenolysis of benzyl protecting group. The inhibition constant was measured for the BRU isolate of HIV-1 protease, $K_i = 33$ nM.[2]

Native HIV-1 protease in 50 mmol L^{-1} sodium acetate buffer, pH 5.6, 1 mmol L^{-1} EDTA (ethylenediaminetetraacetic acid from disodium salt), and 0.05% 2-mercaptoethanol at a concentration of 3 mg mL^{-1} was inhibited with 4-fold molar excess of the inhibitor dissolved in dimethyl sulfoxide to concentration of 11 mM.[15] Hanging drop vapor diffusion method was used for crystallization. Crystals were grown with 1 mol L^{-1} NH_4 phosphate, 0.1 mol L^{-1} Na citrate (Hampton Research Crystal Screen 11)[16] as precipitant, pH 4.5, at 6−8 °C. Detailed crystallization conditions for the series of complexes have been published.[17,18]

X-ray Data Collection. Platelike crystals with dimensions approximately 0.3 × 0.3 × 0.1 mm^3 were used for X-ray diffraction data collection at the beamline 5.2 R, synchrotron source Elettra, Trieste. Two crystals diffracted at 100 K to the diffraction limit of 1.83 Å. The complex crystallized in space group $P2_12_12$ with unit cell parameters $a = 58.1$ Å, $b = 86.3$ Å, $c = 46.1$ Å, $\alpha = \beta = \gamma = 90°$. The data were collected on a MarResearch image plate 180 mm. Experimental and data processing details are summarized in Table 3.

A total of 143 images of 1° oscillation angle were processed and scaled with HKL software package[19] and resulted in 20474 experimental structure factor amplitudes.

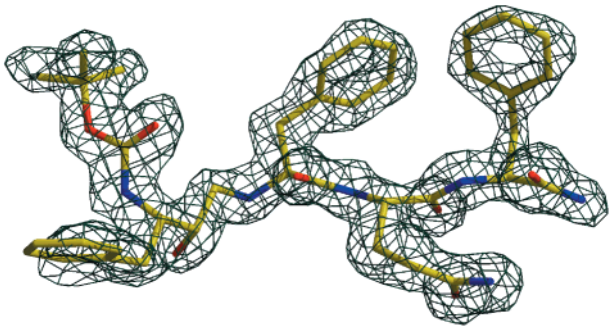

Figure 7. Final $2F_{observed} - F_{calculated}$ electron density map in the protease active site with the SQ inhibitor (contoured at 1.2σ level). Created with programs Molscript[23] and Raster3D.[24]

Structure Solution and Refinement. The starting model of the protein chains of HIV-1 protease was adopted from an isomorphous PDB entry[12] 1HSG ($P2_12_12$, 2 Å resolution, $R = 0.17$). The crystallographic residuals dropped after the initial rigid body refinement to $R = 0.44$ and $R_{free} = 0.46$. The noncrystallographic symmetry restraints were applied to 64 of 99 amino acid residues of chains A and B in the first cycles of refinement. The difference electron density maps provided clear maxima for the inhibitor molecule (Figure 7). The refinement of individual atomic positions and temperature factors lead to $R = 0.22$ and $R_{free} = 0.27$. As significant differences between A and B chains occurred both in the main chain and side chains conformation, the NCS restraints were released and further refinement of atomic positions and individual B factors yielded $R = 0.21$ and $R_{free} = 0.25$. The localized solvent was built in the model abiding by the following rules: (a) difference density maximum above 3σ level present, (b) a hydrogen bond donor/acceptor within bonding distance, (c) $2F_o - F_c$ maximum above 1σ level present after refinement, and (d) water oxygen *B* factor < 50 after refinement.

In total, 301 water oxygens and five glycerol molecules were localized in the structure. Bound glycerol is assumed to be a result of cryoprotectant soaking before flash freezing and diffraction data collection. The occupancy factor of 101 water oxygens occupying alternative solvent positions was set to 0.5. The final cycles of positional and temperature factor refinement resulted in $R = 0.17$, $R_{free} = 0.21$, and $R_{all} = 0.166$ on all structure factor amplitudes including the test set. The structural parameters are summarized in Table 3.

Alternative conformations were modeled for 19 amino acid residues of the protease dimer.

For data manipulation, refinement and model building computer programs CCP4,[20] CNS,[21] and O[22] were utilized.

The structural data and X-ray structure factors have been submitted to the Protein Data Bank under access code 1IIQ.

Acknowledgment. The work was supported by the Czech Technical University (grant no. 30497487), Grant Agency of the Academy of Sciences of the Czech Republic (project A4050811), Grant Agency of the Czech Republic (projects no. 203/97/P031, no. 203/98/K023, no. 203/00/D117, and no. 204/00/P091) and by an International Research Scholar's Award from the Howard Hughes Medical Institute (HHMI 75195-54081) to J.K. The authors thank Sincrotrone Elettra in Trieste for providing beamtime and Dr. A. Savoia for support.

References

(1) Deeks, S. G.; Kahn, J. O. *The Problem of Protease Resistance*; UCSF AIDS Program, San Francisco General Hospital, San Francisco, CA, 1997; http://www.thebody.com/hivnews/newsline/june97/protease.html.

(2) Konvalinka, J.; Litera, J.; Weber, J.; Vondrášek, J.; Hradilek, M.; Souček, M.; Pichová, I.; Majer, P.; Štrop, P.; Sedláček, J.; Heuser, A. M.; Kottler, H.; Kräusslich, H. G. Configurations of diastereomeric hydroxyethylene isosteres strongly affect biological activities of a series of specific inhibitors of human-immunodeficiency-virus proteinase. *Eur. J. Biochem.* **1997**, *250*, 559−566.
(3) Rinnová, M.; Hradilek, M.; Bařinka, C.; Weber, J.; Souček, M.; Vondrášek, J.; Klimkait, T.; Konvalinka, J. A picomolar inhibitor of resistant strains of human immunodeficiency virus protease identified by a combinatorial approach. *Arch. Biochem. Biophys.* **2000**, *382*, 22−30.
(4) Hašek, J.; Dohnálek, J.; Dušková, J.; Konvalinka, J.; Hradilek, M.; Souček, M.; Sedláček, J.; Brynda, J.; Buchtelová, E. Antiretroviral drug design based on structure analysis of proteases. *Mater. Stuct.* **1998**, *5*, 437−438.
(5) Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia M. A. Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc.* **1995**, *117*, 1181−1182.
(6) Krohn, A.; Redshaw, S.; Ritchie, J. C.; Graves, B. J.; Hatada, M. H. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (*R*)-hydroxyethylamine isostere. *J. Med. Chem.* **1991**, *34*, 3340−3342.
(7) Babine, R. E.; Bender, S. L. Molecular recognition of protein−ligand complexes: Applications to drug design. *Chem. Rev.* **1997**, *97*, 1359−1472.
(8) Wlodawer, A.; Erickson, J. W. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* **1993**, *62*, 543−585.
(9) Tossi, A.; Bonin, I.; Antcheva, N.; Norbedo, S.; Benedetti, F.; Miertus, S.; Nair, A. C.; Maliar, T.; Dal Bello, F.; Palù, G.; Romeo, D. Aspartic protease inhibitors − An integrated approach for the design and synthesis of diaminodiol-based peptidomimetics. *Eur. J. Biochem.* **2000**, *267*, 1715−1722.
(10) Desiraju, G. R.; Steiner, T. *The weak hydrogen bond*; Oxford Univerzity Press: New York, 1998.
(11) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. Procheck − a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283.
(12) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(13) Dohnálek, J.; Hašek, J.; Dušková, J.; Petroková, H.; Hradilek, M.; Souček, M.; Konvalinka, J.; Brynda, J.; Sedláček, J.; Fábry, M. A distinct binding mode of a hydroxyethylamine isostere inhibitor of HIV-1 protease. *Acta Crystallogr. D* **2001**, *57*, 472−476.
(14) Skalova, T.; Hasek, J.; Dohnalek, J.; Petrokova, H.; Buchtelova, E. Affinity of HIV-1 protease to tetrapeptide inhibitors. *Mater. Struct.* **2001**, *8*, 29−31.
(15) Sedláček, J.; Fábry, M.; Hořejší, M.; Brynda, J.; Luftig, R. B.; Majer, P. A rapid screening method for biological-activity of human-immunodeficiency-virus proteinase-inhibitors by using a recombinant DNA-derived bacterial system. *Anal. Biochem.* **1993**, *215*, 306−309.
(16) Jancarik, J.; Kim, S. H. Sparse-matrix sampling − a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **1991**, *24*, 409−411.
(17) Dohnálek, J.; Hašek, J.; Brynda, J.; Fábry, M.; Sedláček, J.; Konvalinka, J.; Hradilek, M.; Souček, M.; Adams, M. J.; Naylor, C. E. Peptidomimetic inhibitors complexed with HIV−I protease: Crystallisation for X-ray diffraction studies. *Gen. Physiol. Biophys.* **1998**, *17*, Suppl. 1, 9−11.
(18) Buchtelová, E.; Hašek, J.; Dohnálek, J.; Tykarska, E.; Jaskolski, M.; Olivi, L. Preliminary crystallographic study of native HIV-1 protease inhibited by hydroxyethylamine modified tetrapeptides of Boc-Phe-Ψ[(*R*/*S*)-CH(OH)CH_2NH]-Phe-Ile-Phe-NH_2 type. *Mater. Struct.* **1999**, *6*, 6−7.
(19) Otwinowski, Z. *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing", 29−30 January 1993*; Sawyer, L., Isaacs, N., Bailey, S., Eds.; SERC Daresbury Laboratory, 1993; pp 56−62.
(20) Collaborative Computational Project, Number 4. The *CCP*4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **1994**, *50*, 760−763.
(21) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **1998**, *54*, 905−921.
(22) Jones, T. A.; Kjeldgaard, M. *Manual for O version 5.9*. Department of Molecular Biology, BMC, Uppsala University, Sweden; Department of Chemistry, Aarhus University, Denmark 1993.
(23) Kraulis, P. J. Molscript − a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(24) Merritt, E. A.; Bacon, D. J. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **1997**, *277*, 505−524.

JM010979E