Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# On the role of the *R* configuration of the reaction-intermediate isostere in HIV-1 protease-inhibitor binding: X-ray structure at 2.0 Å resolution

**Jarmila Dušková,[a]* Jan Dohnálek,[a] Tereza Skálová,[a] Hana Petroková,[a] Eva Vondráčková,[a] Martin Hradílek,[b] Jan Konvalinka,[b] Milan Souček,[b] Jiří Brynda,[c] Milan Fábry,[c] Juraj Sedláček[c] and Jindřich Hašek[a]**

[a]Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovského nám. 2, 162 06 Praha 6, Czech Republic, [b]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 37 Praha 6, Czech Republic, and [c]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 37 Praha 6, Czech Republic

Correspondence e-mail: duskova@imc.cas.cz

Received 13 December 2005
Accepted 23 February 2006

**PDB References:** HIV-1 protease–inhibitor RE complex, 1zsf, r1zsfsf; HIV-1 protease–inhibitor SE complex, 1zsr, r1zsrsf.

Peptidomimetic inhibitors of human immunodeficiency virus-1 protease are successful lead substances for the development of virostatic drugs against HIV as the causative agent of acquired immunodeficiency syndrome (AIDS). The hydroxyethylamine isostere of the proteolytic cleavage intermediate provides a suitable replacement for the peptide bond. A series of acyclic pseudopeptide inhibitors with the hydroxyethylamine isostere varying in chiral carbon configuration and $P_2'$ residue type were structurally analysed by single-crystal X-ray crystallography. The compounds inhibit HIV protease with subnanomolar inhibition constants and block viral replication in tissue cultures. Here, the structure of such a complex with the *R* configuration of the isosteric group (PDB code 1zsf) is presented together with newly available synchrotron data for a complex with the *S* stereoisomer of the inhibitor (PDB code 1zsr). Comparison of the structure and binding with other complexes of HIV-1 protease and similar inhibitors contributes to the understanding of how these molecules bind to the wild-type form of this enzyme. The hydroxy group of the *R* stereoisomer interacts with one of the catalytic aspartic acids by a short hydrogen bond with rather extreme geometry. The change of configuration of the chiral carbon bearing the hydroxyl from *S* to *R* does not influence the inhibition efficiency in this case.

## 1. Introduction

In the last ten years, the epidemic of HIV infection has grown to catastrophic dimensions. At the end of 2005, the number of people living with HIV was estimated at between 36.7 and 45.3 million (UNAIDS, 2005) with the total number of AIDS-caused deaths during 2005 estimated at around 3.1 million. One of the several strategies for treatment is based on inhibition of HIV-1 protease, which is essential in the virus life cycle and is responsible for the maturation of new virions and the infectivity of the virus. The function of aspartic protease is to cleave the viral Gag and Gag-Pol polyproteins during the replication cycle of the virus. The viral protease cuts the long chain into its individual enzyme components, which then facilitate the production of new viruses. This fact is exploited in substrate-based inhibitor design, in which the scissile bond is replaced by a non-cleavable isostere. Various types of inhibitors have been proposed, with the isostere based on hydroxyethylene (Jaskolski *et al.*, 1991), hydroxyethylamine (Kervinen *et al.*, 1996) or reduced amide (Beaulieu *et al.*, 1997). Other types of inhibitors were based on derivatives of urea (Jadhav *et al.*, 1997), penicillin (Jhoti *et al.*, 1994), phosphinate (Abdel-Meguid *et al.*, 1993), sulfonamide (Backbro *et al.*, 1997) and difluoroketone (Silva *et al.*, 1996). Clinical studies showed that application of inhibitor-based drugs

© 2006 International Union of Crystallography
Printed in Denmark – all rights reserved

causes mutations in the protease sequence and/or in the Gag cleavage sites and thus leads to viral resistance (Deeks & Kahn, 1997). Owing to the nature of retroviral replication, selection pressure caused by these compounds readily leads to mutant forms of the virus with increased resistance to the drugs used. The high frequency of drug-related mutations makes the case for further development of potent HIV protease inhibitors. Peptidomimetics play an important role as initial lead compounds on which some fundamental principles of inhibition can be relatively easily tested.

Here, we report the structures of complexes of wild-type HIV-1 protease (BRU isolate) with two peptidomimetic inhibitors RE and SE of identical chemical composition that are *R* and *S* stereoisomers with respect to the chiral carbon of the isostere bearing a hydroxy group. The structure formula of the tetrapeptide inhibitor with hydroxyethylamine isostere RE, Boc-Phe-Ψ[(*R*)-CH(OH)$CH_2$NH]-Phe-Glu-Phe-$NH_2$, differs from that of the inhibitor SE in the replacement of (*R*) by (*S*). The respective complexes are denoted WT-RE (*i.e.* wild-type HIV-1 protease complexed with the *R* stereoisomer of the inhibitor with Glu in the $P_2'$ position) and WT-SE. Furthermore, these two structures were compared with that of WT-OE (Petroková *et al.*, 2004), the complex of HIV-1 protease with a similar tetrapeptide inhibitor OE lacking the hydroxy group on the formerly chiral carbon of the isostere and with the structure formula Boc-Phe-Ψ[$CH_2CH_2$NH]-Phe-Glu-Phe-$NH_2$ (tetrapeptide inhibitor with ethyleneamine isostere).

Systematic analysis of structural differences within the group of HIV protease complexes with several highly similar inhibitors and correlation of their efficacy with structure stands out as an important tool for rational drug design of second-generation protease inhibitors. Such analysis brings a unique opportunity to observe the individual effects of each chemical modification of these inhibitors separately.

**Table 1**
Diffraction data statistics and structure parameters of complexes of HIV-1 protease with RE and SE inhibitors.

Values in parentheses are for the highest resolution shell.

| | | |
|---|---|---|
| Diffraction data | | |
| PDB code | 1zsf | 1zsr |
| Inhibitor | RE | SE |
| Space group | $P6_1$ | $P6_1$ |
| Unit-cell parameters (Å, °) | $a$ = 62.86, $b$ = 62.86, $c$ = 82.38, $\alpha$ = 90.0, $\beta$ = 90.0, $\gamma$ = 120.0 | $a$ = 62.80, $b$ = 62.80, $c$ = 82.22, $\alpha$ = 90.0, $\beta$ = 90.0, $\gamma$ = 120.0 |
| Diffraction limits (Å) | 26.0–2.0 (2.05–2.0) | 40.0–2.06 (2.11–2.06) |
| Radiation source | Elettra beamline 5.2R, Trieste | Elettra beamline 5.2R, Trieste |
| Temperature of measurement (K) | 100 | 100 |
| No. of unique reflections | 12461 | 10981 |
| $R_{merge}$ | 0.060 (0.26) | 0.075 (0.30) |
| Data completeness (%) | 97.1 (85.0) | 97.6 (69.8) |
| $I/\sigma(I)$ | 22.6 (2.8) | 26.7 (5.0) |
| Redundancy | 3.4 (1.9) | 6.98 (4.5) |
| Structure parameters | | |
| No. of refined atoms | 1765 | 1792 |
| No. of non-H atoms in disordered inhibitor | 2 × 51 | 2 × 51 |
| No. of refined water molecules | 129 | 152 |
| Average temperature factor $B$ (Å$^2$) | 29.7 | 29.1 |
| R.m.s.d. bond distances from ideal (Å) | 0.004 | 0.011 |
| R.m.s.d. bond angles from ideal | 0.02 Å (as 1–3 distance) | 1.7° |
| $R_{free}$ (5%) | 0.276 | 0.284 (0.289) |
| $R_{all}$ | 0.222 | 0.202 (0.241) |
| Refinement program | *SHELXL* | *CNS* |

## 2. Materials and methods

### 2.1. Crystallization

**2.1.1. WT-RE complex**. A solution of native HIV-1 protease (BRU isolate) was prepared as described in Sedláček *et al.* (1993). The inhibitor RE, Boc-Phe-Ψ[(*R*)-CH(OH)$CH_2$NH]-Phe-Glu-Phe-$NH_2$ (molecular weight $M_r$ = 664.74), a subnanomolar inhibitor of HIV PR ($K_i$ = 0.12 n*M*; Konvalinka *et al.*, 1997) was prepared by alkylation of the N-terminal amino group of the tripeptide with a pure diastereoisomer of *N*-(*t*-butyloxycarbonyl)-1-oxiranyl-2-phenylethan-1-amine at elevated temperature in a protic solvent as described previously (Konvalinka *et al.*, 1997).

The protease was concentrated to 3 mg ml$^{-1}$ and inhibited with a four-molar excess of an 11 m*M* solution of inhibitor dissolved in dimethyl sulfoxide. The protease–inhibitor complex was stored in sodium acetate buffer pH 5.8, 0.05% 2-sulfanylethanol (2-mercaptoethanol) and 1 m*M* ethylenediaminetetraacetic acid. Crystals of the complex were grown by the hanging-drop vapour-diffusion method with reservoir solution composed of 0.5 *M* ammonium phosphate and 0.1 *M* sodium citrate. The crystal used for X-ray diffraction data collection was grown at pH 4.3 and a temperature ranging from 279 to 281 K (Dohnálek *et al.*, 1998).

**2.1.2. WT-SE complex**. Details of the crystallization of the complex with SE inhibitor {Boc-Phe-Ψ[(*S*)-CH(OH)$CH_2$NH]-Phe-Glu-Phe-$NH_2$} were similar to those of the WT-RE complex and have been published elsewhere (Dohnálek *et al.*, 1998, 2001).

### 2.2. Data collection and processing

**2.2.1. WT-RE complex**. A needle-like crystal 0.75 mm in length was cryoprotected with 30% glycerol and X-ray diffraction data were collected at 100 K at a wavelength of 1.0 Å at beamline 5.2R of the Elettra synchrotron-radiation source, Trieste equipped with a MAR 30 cm image-plate detector. The data were processed using *DENZO* and scaled using *SCALEPACK* (*HKL* package; Otwinowski & Minor, 1997). The diffraction pattern revealed hexagonal symmetry, with an $R_{merge}$ of 0.062 for the data set in space group $P6_1$ (Table 1).

**Table 2**
HIV-1 protease amino-acid residues forming the protease subsites corresponding to inhibitor residues $P_2$–$P'_3$.

Lists contain the protease residues of the RE complex with at least one close contact to the inhibitor side chains at a distance of up to 4.1 Å.

| | |
|---|---|
| $S_2$ | Gly148, Asp130, Ala128, Val132, Ile50, Ile184 |
| $S_1$ | Leu23, Gly127, Gly148, Gly149, Ile150, Val82, Pro81 |
| $S'_1$ | Asp125, Ile184, Leu123, Val182, Ile50, Gly49, Thr180, Pro181 |
| $S'_2$ | Ala28, Val32, Asp30, Ile47, Asp29, Gly48 |
| $S'_3$ | Asp29, Arg108, Gly48, Asp30, Phe53 |

**2.2.2. WT-SE complex**. A needle-like crystal cryoprotected with 30% glycerol was used to collect X-ray diffraction data at 100 K at a wavelength of 1.0 Å at beamline 5.2R of the Elettra synchrotron-radiation source. The data were processed in space group $P6_1$ at 2.06 Å resolution, as in the case of WT-RE, with an $R_{\rm merge}$ of 0.075 (Table 1).

### 2.3. Structure refinement and validation

**2.3.1. WT-RE complex**. The starting model for rigid-body refinement consisted of the protein chains of a 2.5 Å resolution structure of HIV-1 protease in complex with a hydroxyethylene inhibitor (PDB code 1aaq; Dreyer *et al.*, 1992). Refinement was carried out using *SHELXL* (Sheldrick, 1997) with data in the resolution range 12–2.0 Å in space group $P6_1$. 5% of the data were used for the $R_{\rm free}$ statistic for validation of the course of refinement. After each refinement cycle, the model was manually adjusted in *O* (Jones, 2004). During this procedure alternative conformations of the side chains of three residues were found (Phe53, Phe153 and Glu35). In the cases of Cys67 and Cys167, difference Fourier maxima at the 4.5$\sigma$ level were found at a distance of $\sim$1.65 Å from the $S^{\delta}$ atoms of the residues. These maxima were interpreted as cysteine oxidation and the residues were modelled and named as singly oxidized states of cysteine (residue name CEA).

Positive maxima of the $F_{\rm o} - F_{\rm c}$ Fourier map in the centre of the active site indicated the presence of the inhibitor. The molecule of the inhibitor was built in *O* based on the geometry of fragments obtained from the Cambridge Structural Database (Allen, 2002). The bond distances and angles of these fragments were used in the *SHELXL* input file to define geometry restrictions. Two orientations of the inhibitor were found in the active site, related by a pseudo-twofold axis. They were refined with the initial occupancy set to 0.5 and this value did not change significantly during the refinement, indicating that both orientations are present equally in the crystal.

The O atoms of water molecules were added based on difference electron-density peaks above the 3$\sigma$ level and on their contacts with other hydrogen-bond acceptors or donors. Bulk-solvent correction was applied as implemented in *SHELXL*, with parameters $g$ = 0.66, $U$ = 1.84 (Moews & Kretsinger, 1973). The geometry of the protein was monitored using *PROCHECK* (Laskowski *et al.*, 1993). 94.3% of all residues occur in the most favoured regions of the Ramachandran plot and the remainder in allowed regions. Statistical and quality parameters of the structure determination are listed in Table 1.

**2.3.2. WT-SE complex**. Protein chains of the structure of another complex from this series (PDB code 1iiq) were placed in the correct position for space group $P6_1$ and then used as a starting model for refinement using *CNS* (Brünger *et al.*, 1998). The $F_{\rm o} - F_{\rm c}$ electron density showed orientational disorder of the inhibitor and further refinement confirmed 1:1 occupancy of both orientations. Disordered side chains were modelled at Phe53 and Phe153 and main-chain disorder was modelled for residues Ile50–Gly51 and Ile150–Gly151. As in the case of WT-RE, Cys67 and Cys167 were found to be oxidized to *S*-hydroxy-cysteine. A total of 152 sites for water molecules were localized in the electron-density maps. Three water molecules with occupancy 0.5 lie near the alternative conformations of protein and inhibitor. Bulk-solvent correction was applied throughout refinement and in the last refinement cycle the parameters were optimized to values of $k$ = 0.40 and $B$ = 101. The crystallographic residual factors fell to $R$ = 0.20 and $R_{\rm free}$ = 0.28 for the 5% of reflections used for testing. The final refinement cycle was carried out against all measured reflections.

## 3. Results and discussion

### 3.1. Structure of the complex of HIV-1 protease with RE inhibitor (WT-RE complex)

The molecule of HIV-1 protease consists of two chemically identical chains containing 99 residues (denoted *A*1–*A*99 and *B*101–*B*199). The secondary structure of the complex (mainly $\beta$-strands, turns and loops) is similar to other previously determined HIV protease structures found in the PDB (Berman *et al.*, 2002). The pseudo-symmetrical dimer of the protease possesses one active site in the centre of the molecule formed by equal contributions from both protein chains. The inhibitor molecule was found at the active site in two opposite and exclusive orientations (denoted as *A*201–*A*206 and *B*201–*B*206). The inhibitor is bound to two catalytic aspartic acid residues Asp25 and Asp125 by the hydroxyethylamine isostere. The hydroxy group of the isostere binds asymmetrically to only one O atom: $O^{\delta 2}$ of aspartate Asp25 (Fig. 1). The amino group participates in hydrogen bonds to both carboxylic groups of the catalytic residues. The $P_1$ and $P_2$ residues (phenyl group and *t*-butyloxycarbonyl, Boc) of the inhibitor are bound in the corresponding binding pockets (Table 2). The $P_2$ Boc group only partially interacts with the common $S_2$ residues. The $P_1$ phenylalanine successfully mimics the binding of the favoured hydrophobic residues in this position by completely filling the $S_1$ site.

In order to evaluate the differences between the *A* and *B* chains of the molecule, the equivalent main-chain atoms were superimposed in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) with a resulting r.m.s.d. of 0.33 Å. The angle of rotation to superimpose subunit *A* on subunit *B* is 179.97°. The most significant distortion of the pseudo-twofold symmetry of the dimer is caused by large

differences in the $\varphi$, $\psi$ angles of Ile50 and Ile150 (a distance of 3.3 Å between 'superimposed' carbonyl O atoms caused by a flipped peptide bond). This asymmetry is required to allow the formation of the hydrogen bond between the carbonyl group of Ile50 and the amine of Ile150. The remaining main chain is in good agreement with the pseudo-twofold symmetry.

When all side-chain atoms of the *A* and *B* chains are superimposed, the resulting r.m.s.d. is 1.36 Å and the largest differences are found for residues Phe53 and Phe153, which have alternative conformations. Glu35 was the only other residue with alternative conformations of its side chain. Regions Gly40–Ile46 and Gly140–Ile146 show high levels of disorder.

**3.1.1. Inhibitor–protein interaction**. The position of the inhibitor mimics substrate binding with the non-cleavable isostere placed instead of the scissile peptide bond. A detailed view of hydrogen-bond interactions between the inhibitor and the protease is presented in Fig. 1. The inhibitor interacts with the protein *via* 17 hydrogen bonds so that all possible hydrogen-bond donors and acceptors of the inhibitor find partners in the protease dimer or in the surrounding water molecules.

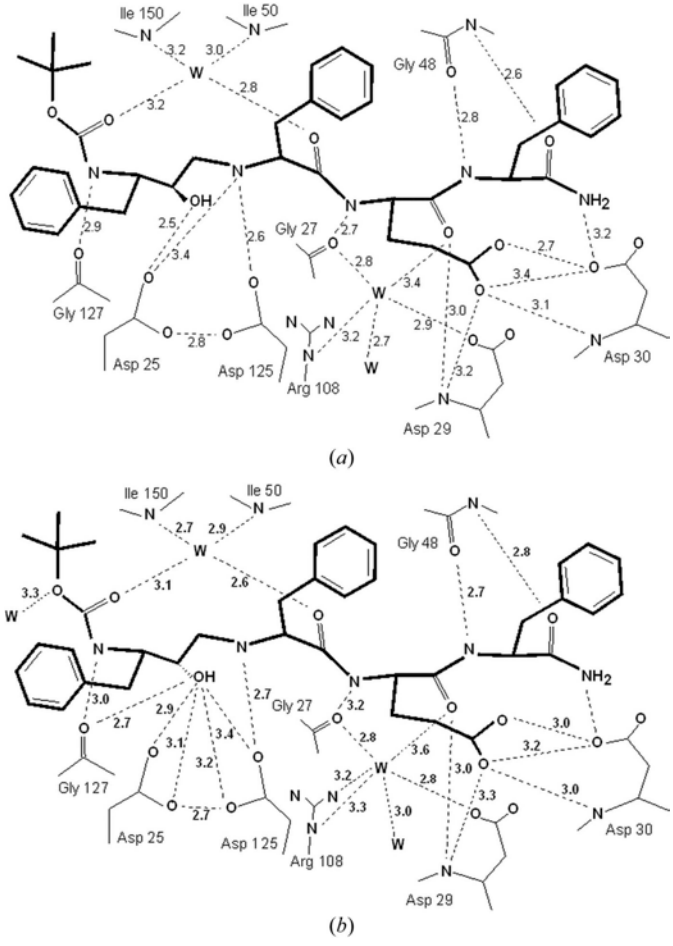


**Figure 1**
Schematic representation of the hydrogen-bonding pattern between (*a*) the RE inhibitor (PDB code 1zsf) and (*b*) the SE inhibitor (PDB code 1zsr) and protein molecules. Distances are given in angstroms.

The structurally important water molecule W301 mediates contacts between Ile50, Ile150 and the carbonyl O atoms of the $P_2$ and $P_1'$ residues. Another water molecule, W327, mediates contact between the $P_2'$ O atom and Arg8 $N^{\varepsilon}$.

The bonding pattern for the pseudosymmetrical molecule of the inhibitor (chain *B*201–*B*206) is very similar, with differences in contact distances of up to 0.2 Å. The inhibitor–protease interactions are supported by numerous van der Waals contacts. Residues forming the hydrophobic sites up to a distance of 4.1 Å from the inhibitor side chains are listed in Table 2.

**3.1.2. Interaction of the isostere**. The hydroxyethylamine isostere —(*R*)-CHOH—$CH_2$—NH— provides two groups for classical hydrogen bonds: the hydroxy and the amino groups. It should be noted that the *R* configuration of the chiral carbon of the isostere could resemble a real intermediate stabilized by interaction with the catalytic aspartic acids, one of which provides a proton for the intermediate. The isostere interacts *via* three hydrogen bonds with Asp25 and Asp125. The amino group plays a 'central' role in the binding to the catalytic residues. It forms two hydrogen bonds: a strong 2.6 Å bond to Asp125 $O^{\delta 2}$ and a weak 3.4 Å bond to Asp25 $O^{\delta 2}$. As this is a single strong bond to the Asp125 carboxylate with the NH group as a donor, it can be assumed that Asp125 $O^{\delta 2}$ is not protonated in this complex. In this way, the isosteric NH group takes the role of the hydroxyl group of hydroxyethylene-based inhibitors.

The hydroxy group of the studied isostere interacts *via* a short hydrogen bond with Asp25 $O^{\delta 2}$ (Fig. 2). The contact

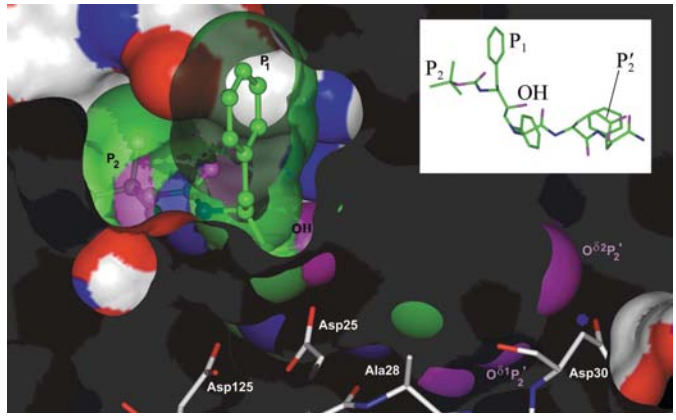


**Figure 2**
A section through the active site of the protease (WT-RE complex) illustrating the tight interaction between the hydroxyethylamine isostere and Asp25. The hydrophobic pockets $S_1$ and $S_2$ are filled with the inhibitor side chains of $P_1$ and $P_2$. The isosteric OH group penetrates through the surface of the catalytic centre which corresponds to the short contact with Asp25 $O^{\delta 2}$. Note also the tight contacts between the ligand and the enzyme between the $P_2'$ Glu carboxyl O atoms and Asp30 of the protease. Connolly solvent-accessible surface is coloured according to atom type (probe radius 1.4 Å) as implemented in *INSIGHTII* (Accelrys). Some residues of the enzyme are shown as sticks coloured according to atom type and the RE inhibitor is shown as both sticks and as a van der Waals surface (carbon, green; nitrogen, blue; oxygen, magenta). For better understanding, the insert (top right) shows the complete molecule of the inhibitor in the same orientation, representation and colour coding.

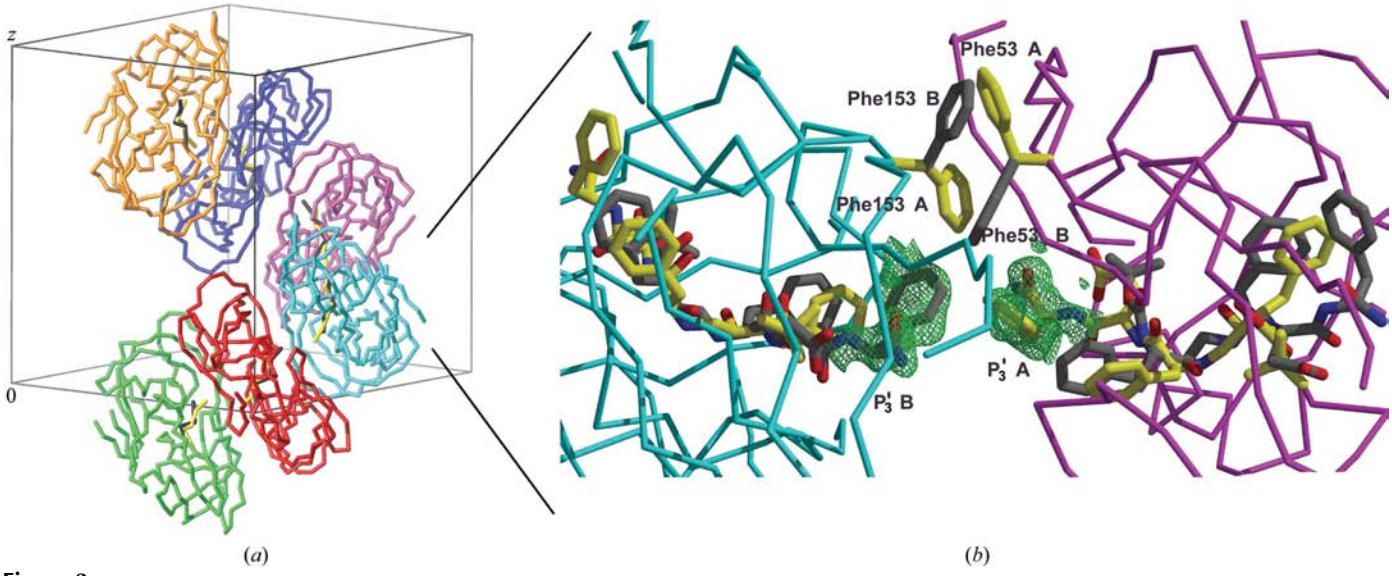


**Figure 3**
Crystal packing and disorder in the structure. (*a*) Crystal packing of six molecules in the hexagonal unit cell: $C^{\alpha}$ atoms. Molecules are arranged in helices around the crystallographic sixfold screw axis *z*. (*b*) Orientation of disordered inhibitors in two neighbouring molecules (stick model). Electron density is plotted for the phenylalanines in $P_3'$ positions (1.4$\sigma$ level). A close contact of 2.1 Å between the Phe side chains in the $P_3'$ positions of the inhibitor in the two alternatives excludes the simultaneous presence of opposite orientations of the inhibitor in neighbouring molecules. Consequently, all inhibitors in one helix of molecules of the protease should have the same orientation. The orientations of inhibitors in the individual helical columns in the crystal are independent. Also, the conformations of residues Phe53 and Phe153 alternate depending on the presence of $P_3'$ phenyl rings owing to steric requirements. This figure was created with *O* (Jones, 2004), *MOLSCRIPT* (Kraulis, 1991), *RASTER3D* (Merritt & Bacon, 1997) and *WebLabViewerPro* (Accelrys).

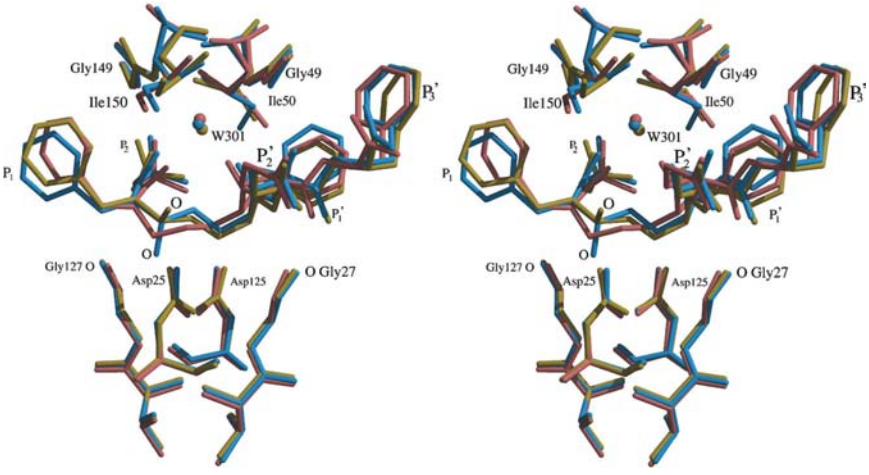


**Figure 4**
Stereoview of the HIV-1 inhibitors RE (yellow), SE (blue) and OE (light coral) in the binding pocket of HIV-1 protease. The side-chain conformations of the inhibitors differ mainly in the $P_1$ residues. The OH groups in the isosteres of RE and SE appear in different locations with respect to the catalytic aspartates. The isostere of the OE inhibitor adopts a different conformation in comparison to RE or SE, despite the fact that the amine group of all these inhibitors binds to both catalytic aspartates in a similar way. This figure was created with *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997).

distance between these two O atoms of 2.54 Å and also the angle of ∼75° between the vectors of the C—OH bond in the isostere and the $C^{\gamma}$—$O^{\delta 2}$ bond of Asp25 indicate a very tight contact with somewhat unusual geometry. *ISOSTAR* scatterplots (Allen, 2002) for this type of interaction (carboxylate–alcohol OH group) based on the Cambridge Structural Database indicate shortest $OH_{carboxylate}$–$O_{alcohol}$ interaction distances of 2.4 Å and shortest $O_{carboxylate}$–$O_{alcohol}$ distances of 2.6 Å. Much shorter contact distances were found based on data from the PDB, but the individual data points are of doubtful significance owing to the quality of the originating structures. Also, the CSD-based scatterplots indicate a wide range of the angle between the donor- and acceptor-bearing bonds in the case of protonated carboxylate oxygen. Owing to the resolution of the reported structure, the positions of H atoms cannot be experimentally observed and so the question remains open whether the hydroxy hydrogen is shared in this short bond in a way that its donor can no longer be distinguished or whether Asp25 donates its hydrogen and the hydroxy hydrogen is not involved in this interaction.

**3.1.3. Crystal packing and inhibitor disorder**. Molecules of the WT-RE complex form helical columns along the *z* axis in the crystal (Fig. 3). There is a channel with diameter ∼10 Å for the passage of solvent molecules which runs along the *z* axis in the centre of the column. The orientation of inhibitor molecules in neighbouring complexes is shown in detail in Fig. 3(*a*). If opposite orientations of the disordered inhibitor were present in two neighbouring complexes, the C atoms of the Phe $P_3'$ side chains of the neighbouring inhibitors would come into close contact (∼2.1 Å). This implies that only one orientation of the inhibitor can be present in one helical column in the crystal. In the real crystal, helical columns with

**Table 3**
Structure-formula differences in the series of studied inhibitors with general formula $P_2$–$P_1$–isostere–$P'_1$–$P'_2$–$P'_3$.

All structures were determined from data collected from flash-frozen crystals except for 1fqx, which was measured at room temperature.

| Inhibitor symbol | HIV-1 protease† | $P_2$ | $P_1$ | Isostere | $P'_1$ | $P'_2$ | $P'_3$ | $K_i$‡ (n*M*) | PDB code | Reference |
|---|---|---|---|---|---|---|---|---|---|---|
| RE | WT | Boc | Phe | (*R*)-CH(OH)$CH_2$NH | Phe | Glu | Phe | 0.12 | 1zsf | This work |
| SE | WT | Boc | Phe | (*S*)-CH(OH)$CH_2$NH | Phe | Glu | Phe | 0.15 | 1zsr | This work |
| SE | WT | Boc | Phe | (*S*)-CH(OH)$CH_2$NH | Phe | Glu | Phe | 0.15 | 1fqx | Dohnálek *et al.* (2001) |
| OE | WT | Boc | Phe | $CH_2CH_2$NH | Phe | Glu | Phe | 1.53 | 1m0b | Petroková *et al.* (2004) |
| OE | I8 | Boc | Phe | $CH_2CH_2$NH | Phe | Glu | Phe | 4.14 | 1lzq | Skálová *et al.* (2003) |
| SQ | WT | Boc | Phe | (*S*)-CH(OH)$CH_2$NH | Phe | Gln | Phe | 33.0 | 1iiq | Dohnálek *et al.* (2002) |

† WT, wild-type HIV-1 protease, BRU isolate; I8, triple mutant A71V, V82T, I84V HIV-1 protease. ‡ $K_i$, inhibition constant.

alternative orientations of the inhibitor are distributed randomly.

The alternative conformations of Phe53 and Phe153 are also closely related to crystal packing and the residues participate in the formation of a hydrophobic surface interacting with $P'_3$ of the inhibitor from a crystallographically related neighbouring complex. The side chain of Phe53 forms crystal contacts with the side chain of Phe153 from a neighbouring molecule. Identical conformations of these side chains would cause steric clashes between the phenyls. Owing to the conformational freedom of this surface residue, the intermolecular contact can be realised in two ways (Phe53 alternative *A*, Phe153 alternative *B* and *vice versa*). Moreover, conformation *A* of Phe53 would clash with conformation *B* of the inhibitor (close contact of the Phe $P'_3$ and Phe53 side-chain C atoms) and conformation *B* of Phe153 with the *A* conformation of the inhibitor. Consequently, the helical columns of the protease–inhibitor complexes always contain only one orientation of the inhibitor and one combination of Phe53/153 alternatives (either the combination inhibitor *A*, Phe53 *A*, Phe153 *B* or the combination inhibitor *B*, Phe53 *B*, Phe153 *A*).

### 3.2. Structure of the complex of HIV-1 protease with SE inhibitor (WT-SE complex)

Many features of the WT-RE structure can be found in the WT-SE structure, including the solvent channel along the *z* axis of the crystal, the alternative conformations of the Phe53 and Phe153 residues and the presence of only one orientation of the inhibitor in a helical column of molecules in the crystal. The alternative conformations of phenylalanines 53 and 153 and of the inhibitor are not necessarily correlated as in WT-RE.

**3.2.1. Protease dimer in the WT-SE complex**. The main chain of chain *A* was superimposed with that of chain *B* with an r.m.s.d. of 0.21 Å and the largest difference was registered for the flipped peptide bond Ile50–Gly51 and Ile150–Gly151. This transformation corresponds to a rotation of chain *B* around the dimer axis by 179.9°. The regions Arg41–Pro44 and Arg141–Pro144 show larger than average differences within the dimer owing to slightly different crystal-packing contacts ('shoulder' 41–44 is always involved in crystal contacts to residues from chain *A* of a neighbouring molecule and 141–144 to chain *B*). All other differences within the dimer can be ascribed to relatively insignificant variations in conformation of side chains.

**3.2.2. Room- and low-temperature WT-SE structures**. A structure of the WT-SE complex based on in-house source data collected at room temperature has been published previously (PDB code 1fqx; Dohnálek *et al.*, 2001).

The differing quality and diffraction limits of the two data sets (number of reflections to number of refined parameters $N_r/N_p = 2$ and resolution 3.1 Å for 1fqx and $N_r/N_p = 6.1$ and 2.06 Å for 1zsr) resulted in similar but not identical structures of the same complex. 28 water sites were identified in the low-resolution structure and 152 in the current structure. The low-temperature high-resolution data provide much better maps of electron density and consequently a better quality structure. The main chain of 198 protein residues of the room-temperature complex was superimposed with that of the high-resolution structure with an r.m.s.d. of 0.37 Å. Large conformational differences in proteins are mainly found for surface and flap residues, but some 'core' amino-acid residues also differ significantly in their conformations.

The most pronounced differences for the SE inhibitor were found in the torsion angles of the $P_1$ and $P'_3$ phenyl rings. The overall binding mode of the inhibitor is in good agreement and the high-resolution structure confirms the initial findings on the binding of these inhibitors from data of as low resolution as 3.1 Å.

The high-resolution structure (1zsr) is used in the following structural comparisons.

## 4. Structural differences between the WT-RE, WT-SE and WT-OE complexes

Complexes with similar inhibitors of this series have been published previously (Dohnálek *et al.*, 2001, 2002; Skálová *et al.*, 2003; Petroková *et al.*, 2004) and Table 3 summarizes the structural differences in this group of inhibitors. These inhibitors vary in their isosteres which contain either an (*S*)-CH(OH) (SE) or (*R*)-CH(OH) (RE) group or a $CH_2$ group without a hydroxy group (OE). The complexes were superimposed based on the most rigid part of the dimer (the $C^\alpha$ atoms of residues 1–11 and 87–99 of both chains). The resulting overlap of inhibitors RE, SE and OE in conformations corresponding to the bound state in protease is shown in Fig. 4.

### 4.1. Comparison of WT-RE and WT-SE complexes

The structures of the two complexes are very similar. The overall position of the *R* and *S* isosteres remains the same. However, as the pattern of inhibitor binding in these complexes is mostly driven by interactions between the $P_2$–$P_3'$ side chains and the protease pockets $S_2$–$S_3'$, the difference in isostere chirality leads to a markedly changed interaction with the catalytic centre. The hydroxy group of (*S*)-hydroxyethylamine binds *via* hydrogen bonds to the side chain of the catalytic Asp25 and to the carbonyl group of Gly127 and there are a number of other potential hydrogen-bond contacts (Fig. 1*b*). The tight placement of the OH group between the Asp25 side chain and the Gly27 carbonyl imposes an additional steric constraint for the *S* isostere conformation and orientation when compared with the *R* isostere. As the (*S*)-CH(OH) group is wedged between Gly27 and Asp25, a rotation of the carboxylic group by ∼25° in its $\chi_2$ torsion compared with the RE complex relaxes the strain. These interactions result in a small rotation of the remaining part of the (*S*)-hydroxyethylamine isostere with respect to the binding site in such way that the $CH_2$ group of the isostere is forced into a position closer to the protease flaps by ∼0.4 Å. The $CH_2$ group comes into close contact with the structurally conserved water molecule W301 (3.6 Å) and therefore in the SE complex the water oxygen is found in a position slightly shifted away from the 'ideal' central location for its tetrahedral arrangement of hydrogen bonds. The small rotation of the *S* isostere with respect to the *R* isostere also affects the exact position of the phenylalanine carbon $P_1$ $C^{\beta}$. The $P_1$ phenyl groups of the SE and RE inhibitors adopt different conformations in the complexes (SE $P_1$ $\chi_1 = -64.8°$, RE $P_1$ $\chi_1 = -84.3°$). Many other atoms of the protease molecule in the flaps and the binding tunnel change their positions in concert with the phenyl-group conformational difference. The side chain of Val23 exhibits the most pronounced difference, with a change of $\chi_1$ from the SE to the RE complex of 58°. The direct influence of the isostere configuration on this set of changes around the $P_1$ side chain cannot be unambiguously deduced from the compared structures.

As a consequence of the slight rotation of the isostere between the two structures, the isosteric NH group of the SE inhibitor is shifted by 0.3 Å to a more distant position with respect to the catalytic aspartates Asp25 and Asp125, making the hydrogen bonds with the corresponding carboxyl O atoms somewhat weaker. The remaining parts of the inhibitors bind in a very similar manner, with a small shift originating from the different isostere accommodation. The side chain of $P_2'$ Glu undergoes a small conformational change from the SE to the RE complex (SE $\chi_3 = -32.4°$ and RE $\chi_3 = 9.3°$).

### 4.2. Comparison of the WT-RE and WT-OE complexes

The OE inhibitor (Petroková *et al.*, 2004), which lacks the isosteric OH group, has more conformational freedom than the RE inhibitor. The *R* isostere, which is more restricted by the OH interaction with the catalytic aspartic acids, exhibits small overall changes in the inhibitor binding compared with OE. A strong hydrogen bond between the ethyleneamine NH and Asp25 $O^{\delta2}$ (distance 2.5 Å) determines the binding mode of the central part of the inhibitor. The phenylalanine side chain of $P_1$ in OE does not follow the RE inhibitor in forming tight hydrophobic interactions with the $S_1$ pocket and there is a difference in $\chi_1$ between these two cases of ∼28°. The different position of the ethyleneamine isostere also causes a shift of the $P_1'$ $C^{\alpha}$ atom towards the flaps, which propagates throughout the rest of the inhibitor. Despite this, the OE $P_1'$ side chain fills the $S_1'$ subsite in a similar fashion.

### 4.3. Comparison of the WT-RE complex with other structures

The overall binding mode of the inhibitor is similar to that of the *C*2 symmetrical aminodiol inhibitor Boc-NH-CH($CH_2$Ph)-CH(OH)-$CH_2$-NH-$CH_2$-CH(OH)-CH($CH_2$Ph)-NH-Boc (BMS-1821193; Kervinen *et al.*, 1996; PDB code 1odx). The aminodiol group contains two hydroxyl groups on two chiral C atoms with *R* configuration which bind to the $O^{\delta2}$ atoms of the catalytic residues Asp25 and Asp125 with distances of 2.6 and 2.7 Å, respectively. The geometry of the hydrogen bond, with angles Asp25 $O^{\delta2}\cdots$HO—(*R*)—C of 102.2° and Asp125 $O^{\delta2}\cdots$HO—(*R*)—C of 104.2°, seems to be much more relaxed compared with the RE inhibitor, where this value is around 90°. The positions of the $P_1$ and $P_2$ residues in subsites $S_1$ and $S_2$ are approximately the same in both complexes. However, in the currently reported structure, the isostere exhibits a shift of ∼1.2 Å so that the amino group is placed less symmetrically with respect to Asp25 and Asp125. The larger aminodiol isostere results in a significant shift of the rest of the inhibitor compared with RE. This becomes apparent mainly in the misfit of $P_1'$ in its hydrophobic pocket and could also be the reason for the better inhibition properties of the hydroxyethylamine inhibitors ($K_i$ = 0.12 n*M* for RE *versus* $K_i$ = 100 n*M* for BMS-1821193).

The character of the hydroxyethylamide isostere of the HIV-2 protease inhibitor, QNC-THR-PHM-RHS (Beaulieu *et al.*, 1997), resembles that of our isosteric group. However, compared with RE, the hydrogen bonds and hydrophobic interactions have a completely different scheme and the molecules are shifted against each other (based on the superposition of protein $C^{\alpha}$ atoms) by 1.68 Å. The QNC-THR-PHM-RHS inhibitor ($K_i$ = 1.2 n*M*) is bound by its hydroxy group located between the catalytic aspartates similarly to hydroxyethylene-containing inhibitors without amide or amine.

Hydroxyethylamine isosteres containing cyclic nitrogen (Swain *et al.*, 1990; Martin *et al.*, 1999) or hydroxyethylene isosteres (Wlodawer & Vondrasek, 1998) always strongly prefer either the *S* or *R* configuration. All structures of complexes of HIV protease with the clinically used inhibitor Saquinavir (Krohn *et al.*, 1991), containing a hydroxyethylamine isostere with nitrogen as a part of a cycle of the $P_1'$ residue, available in the PDB contain *R* stereoisomers of the inhibitor. The hydroxy group of Saquinavir, as in other cyclic hydroxyethylamine isosteres, is involved in interaction with the catalytic aspartates. The isosteres of RE and Saquinavir

are relatively shifted by about 1.5 Å (measured for the chiral C atom). Even with this shift, there is a great similarity in the way the hydrophobic side chains of these inhibitors interact with the respective protease-binding pockets.

The influence of stereochemistry on activity and binding modes has been widely discussed in Hosur *et al.* (1994) for *C*2 symmetry-based ethylenediol inhibitors, where the configuration of the carbon bearing the hydroxyl bound between aspartates does not influence inhibition properties whereas the *R* stereoisomer with respect to the second hydroxyl leads to a tenfold decrease in inhibition potency. However, the binding mode of these inhibitors, mainly of the isostere, differs significantly from that of our series.

## 5. Conclusion

The structures of HIV-1 protease in complex with Boc-Phe-Ψ[(*R*)-CH(OH)$CH_2$NH]-Phe-Glu-Phe-$NH_2$ and Boc-Phe-Ψ[(*S*)-CH(OH)$CH_2$NH]-Phe-Glu-Phe-$NH_2$ were determined at 2.0 Å resolution. In spite of the very simple changes in stereochemistry, each inhibitor has a specific binding pattern in the active site of the protease. The positions of the inhibitors in the protease-binding tunnel in the WT-RE, WT-SE and WT-OE structures are very similar. In the cases of the RE and SE inhibitors, both the amino and hydroxy groups of the hydroxyethylamine isostere participate in strong hydrogen bonds to the catalytic aspartates Asp25 and Asp125.

However, in comparison to the ethyleneamine isostere it is evident that the presence of the isosteric hydroxy group in the RE and SE inhibitors is important in spite of spatial requirements. Even though the SE and RE inhibition constants do not indicate a significant difference between the affinities of the two molecules for HIV-1 protease, structural analysis reveals that the active site of the enzyme accommodates the *R* isomer in a more relaxed conformation, leading to less strain on the surrounding environment than in the case of the *S* isomer. Inhibitor molecules of this type find their binding positions mainly by the interactions of the phenyl groups of the $P_1$ and $P_1'$ and $P_2'$ side chains with their corresponding subsites as long as the interaction of the isostere with the active centre allows. The *R* isostere in this sense enables a more 'convenient' positioning of the inhibitor side chains and in particular a more effective filling of the hydrophobic subsite $S_1$. From the structural point of view, the *R* configuration of the hydroxyethylamine type of isostere seems to contribute significantly to an 'easier' fit of the inhibitor in the protease active site.

To the best of our knowledge, this is the first evidence of hydroxyethylamine-based inhibitors with a negligible change of inhibition potency related to their stereochemistry ($K_i$ = 0.12 n*M* for RE and 0.15 n*M* for SE) but with experimentally confirmed substantial differences in inhibitor binding. Other hydroxyethylamine isosteres containing cyclic nitrogen (Swain *et al.*, 1990; Martin *et al.*, 1999) or hydroxyethylene isosteres (Wlodawer & Vondrasek, 1998) either strongly prefer the *S* or *R* configuration or in cases of equal potency of both stereoisomers are not accompanied by experimental confirmation of differences in binding. Moreover, the results presented here indicate that peptidomimetics-derived inhibitors of HIV-1 protease with similar types of isosteres will very likely prefer *R* stereoisomers to *S* stereoisomers if assessed by quality of the fit in the active-site tunnel.

We would like to thank Sincrotrone Elettra in Trieste for the beam time provided and the Academy of Sciences of the Czech Republic for financial support (project KJB4050312).

## References

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

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.

Backbro, K., Lowgren, S., Osterlund, K., Atepo, J., Unge, T., Hulten, J., Bonham, N. M., Schaal, W., Karlen, A. & Hallberg, A. (1997). *J. Med. Chem.* **40**, 898–902.

Beaulieu, P. L., Wernic, D., Braham, A. A., Anderson, P. C., Bogri, T., Bousquet, Y., Croteau, G., Guse, I., Lamarre, D., Liard, F., Paris, W., Thibeault, D., Pav, S. & Tong, L. (1997). *J. Med. Chem.* **40**, 2164–2176.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D. & Zardecki, C. (2002). *Acta Cryst.* D**58**, 899–907.

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. (1998). *Acta Cryst.* D**54**, 905–921.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

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

Dohnálek, J., Hašek, J., Brynda, J., Fábry, M., Sedláček, J., Konvalinka, J., Hradílek, M., Souček, M., Adams, M. J. & Naylor, C. E. (1998). *Gen. Physiol. Biophys.* **17**, *Suppl.* 1, 9–11.

Dohnálek, J., Hašek, J., Dušková, J., Petroková, H., Hradílek, M., Souček, M., Konvalinka, J., Brynda, J., Sedláček, J. & Fábry, M. (2001). *Acta Cryst.* D**57**, 472–476.

Dohnálek, J., Hašek, J., Dušková, J., Petroková, H., Hradílek, M., Souček, M., Konvalinka, J., Brynda, J., Sedláček, J. & Fábry, M. (2002). *J. Med. Chem.* **45**, 1432–1438.

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. (1992). *Biochemistry*, **31**, 6646–6659.

Hosur, M. V., Bhat, T. N., Kempf, D. J., Baldwin, E. T., Liu, B., Gulnik, S., Wideburg, N. E., Norbeck, D. W., Appelt, K. & Erickson, J. W. (1994). *J. Am. Chem. Soc.* **116**, 847–855.

Jadhav, P. K., Ala, P., Woerner, F. J., Chang, C. H., Garber, S. S., Anton, E. D. & Bacheler, L. T. (1997). *J. Med. Chem.* **40**, 181–191.

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

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

Jones, T. A. (2004). *Acta Cryst.* D**60**, 2115–2125.

Kervinen, J., Thanki, N., Zdanov, A., Tino, J., Barrish, J., Lin, P. F., Colonno, R., Riccardi, K., Samanta, H. & Wlodawer A. (1996). *Protein Pept. Lett.* **3**, 399.

Konvalinka, J., Litera, J., Weber, J., Vondráček, J., Hradílek, M., Souček, M., Pichová, I., Majer, P., Štrop, P., Sedláček, J., Heuser, A. M., Kottler, H. & Kräusslich, H. G. (1997). *Eur. J. Biochem.* **250**, 559–566.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Krohn, A., Redshaw, S., Ritchie, J. C., Graves, B. J. & Hatada, M. H. (1991). *J. Med. Chem.* **34**, 3340–3342.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Martin, J. L., Begun, J., Schindeler, A., Wickramasinghe, W. A., Alewood, D., Alewood, P. F., Bergman, D. A., Brinkworth, R. I., Abbenante, G., March, D. R., Reid, R. C. & Fairlie, D. P. (1999). *Biochemistry*, **38**, 7978–7988.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Moews, P. C. & Kretsinger, R. H. (1973). *J. Mol. Biol.* **91**, 201–228.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Petroková, H., Dušková, J., Dohnalek, J., Skálová, T., Vondráčková-Buchtelová, E., Souček, M., Konvalinka, J., Brynda, J., Fábry, M., Sedláček, J. & Hašek, J. (2004). *Eur. J. Biochem.* **271**, 4451–4461.

Sedláček, J., Fábry, M., Hořejší, M., Brynda, J., Luftig, R. B. & Majer, P. (1993). *Anal. Biochem.* **215**, 306–309.

Sheldrick, G. M. (1997). *SHELXL*97. University of Göttingen, Germany.

Silva, A. M., Cachau, R. E., Sham, H. L. & Erickson, J. W. (1996). *J. Mol. Biol.* **255**, 321–346.

Skálová, T., Hašek, J., Dohnálek, J., Petroková, H., Buchtelová, E., Dušková, J., Souček, M., Majer, P., Uhlíková, T. & Konvalinka, J. (2003). *J. Med. Chem.* **46**, 1636–1644.

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

UNAIDS (2005). *AIDS Epidemic Update 2005*, http://www.unaids.org/epi/2005/doc/report_pdf.asp.

Wlodawer, A. & Vondrasek, J. (1998). *Annu. Rev. Biophys. Biomol. Struct.* **27**, 249–284.