



Two Solutions for the Same Problem: Multiple Binding Modes of Pyrrolidine-Based HIV-1 Protease Inhibitors

Andreas Blum[†], Jark Böttcher[†], Stefanie Dörr, Andreas Heine, Gerhard Klebe* and Wibke E. Diederich*

Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany

Received 26 January 2011;
received in revised form
20 April 2011;
accepted 21 April 2011

Edited by M. F. Summers

Keywords:

HIV-1 protease inhibitors;
structure-based drug design;
pyrrolidine-based inhibitors;
crystallography;
deviating binding modes

Structure-based drug design is an integral part of industrial and academic drug discovery projects. Initial lead structures are, in general, optimized in terms of affinity using iterative cycles comprising synthesis, biological evaluation, computational methods, and structural analysis. X-ray crystallography commonly suggests the existence of a single well-defined state, termed binding mode, which is generally assumed to be consistent in a series of similar ligands and therefore used for the following optimization process. During the further development of symmetrically disubstituted 3,4-amino-pyrrolidines as human immunodeficiency virus type 1 protease inhibitors, we discovered that, by modification of the P1/P1' moieties of our lead structure, the activity of the inhibitors towards the active-site mutation Ile84Val was altered, however, not being explainable with the initial underlying structure–activity relationship. The cocrystallization of the most potent derivative in complex with the human immunodeficiency virus type 1 protease surprisingly led to two different crystal forms ($P2_12_12_1$ and $P6_122$). Structural analysis revealed two completely different binding modes; the interaction of the pyrrolidine nitrogen atom with the catalytic aspartates remains as the only similarity. The study presented clearly demonstrates that structural biology has to escort the entire lead optimization process not to fail by an initially observed binding orientation.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

The first successful application of structure-based drug design (SBDD) was reported in 1976 by

Beddell *et al.*, who utilized the three-dimensional structure of hemoglobin for ligand development.¹ Since then, the number of protein structures publicly available in the Protein Data Bank (PDB) has increased exponentially, resulting in more than 70,000 entries to date.² Due to the steadily growing amount of available protein structures and methodological advances in structural analysis, SBDD has meanwhile become an integral part of drug discovery projects.³

The analysis of a protein–ligand complex gives detailed information about protein–ligand interactions. Once an initial lead has been identified and the target protein is accessible for structural analysis, it can further be optimized using iterative cycles comprising synthesis, biological evaluation, computational methods, and structural analysis.⁴ Structural

*Corresponding authors. E-mail addresses:

klebe@mail.uni-marburg.de;

wibke.diederich@staff.uni-marburg.de.

[†] A.B. and J.B. contributed equally to this work.

Abbreviations used: SBDD, structure-based drug design; PDB, Protein Data Bank; HIV, human immunodeficiency virus; SAR, structure–activity relationship; HIV-1, human immunodeficiency virus type 1; vdW, van der Waals; DMSO, dimethyl sulfoxide; MS, mass spectrometry; ESI, electrospray ionization.

analysis using X-ray crystallography usually suggests the existence of a single well-defined state, anticipated as a unique binding mode, which is then utilized to propose the most promising modifications of the initial scaffold to further optimize a given lead. Numerous successful projects of SBDD have been published, leading at best to the development of an approved drug as, for example, in the case of the human immunodeficiency virus (HIV) protease inhibitor tipranavir.⁵

However, the iterative process can face a lot of obstacles and surprises, and numerous examples of a nonunique behavior within a series of similar ligands have been described, of which only the following three should be mentioned as representative examples: a reversed binding mode was detected by X-ray crystallography in the due course of the SBDD process for nonpeptidic HIV protease inhibitors. In the reported case, alkylation of the inhibitor's amide functionalities resulted in a completely different occupation of individual specificity subpockets.⁶ The small energy deviations of such opposing binding modes have been investigated in detail (e.g., by the pH dependency of the binding mode observed in the case of trypsin inhibitor complexes).⁷ An extraordinary observation was made in the case of protein kinase C: the inhibitor binds in different orientations to the two crystallographically independent monomers present in the asymmetric unit.⁸ Observations like those mentioned above suggest that the occurrence of several binding orientations within a series of similar compounds or even for one certain ligand is not a rare case. Such phenomena can be indicated, for example, from a complex and not straightforward structure-activity relationship (SAR), which cannot be explained using the broadly accepted hypothesis that similar ligands bind in a similar fashion.⁹

Recently, we reported our design and synthesis of C_2 -symmetric 3,4-disubstituted pyrrolidines as a new class of human immunodeficiency virus type 1 (HIV-1) protease inhibitors.¹⁰ Our structure-guided optimization was based on the initially observed cocrystal structure of the *N*-benzyl-substituted inhibitor (3*S*,4*S*)-3,4-bis[benzenesulfonylbenzylamino]pyrrolidine (1) (PDB ID: 2PQZ; Fig. 1 and Table 1).

In the due course of the project, additional cocrystal structures of the further developed inhibitors were determined to elucidate their interactions explaining the fully consistent SAR. The conserved binding mode observed in the five crystal structures (PDB IDs: 2PQZ, 2QNP, 2PWR, 2PWC, and 2QNN) is depicted in Fig. 1. All complexes between the topologically C_2 -symmetric inhibitors and the sequentially C_2 -symmetric protein adopt an unsymmetric binding mode with respect to the ligand. However, each subpocket related by an

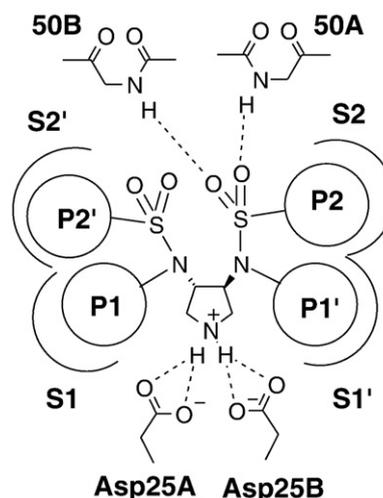
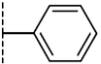
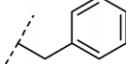
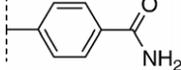
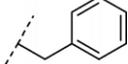
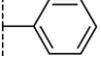
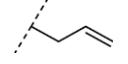
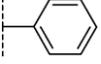
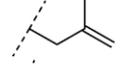
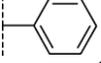
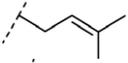
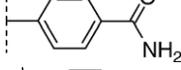
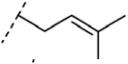
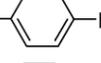
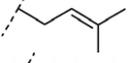
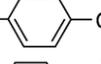
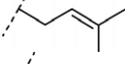
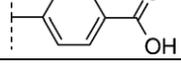
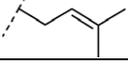


Fig. 1. Schematic representation of the conserved binding mode of the pyrrolidine-3,4-bis-*N*-benzyl-sulfonamides. Hydrogen bonds are indicated by broken lines.

imposed C_2 symmetry is occupied by substituents at the core skeleton of the inhibitor that mutually correspond by topological symmetry. The benzyl moieties occupy the S1 and S1' subpockets, whereas the sulfonamide side chains of the inhibitor address the S2 and S2' subpockets. The *N*-benzyl-substituted derivatives were additionally analyzed with respect to their susceptibility towards the active-site mutants Ile50Val (PR_{I50V}) and Ile84Val (PR_{I84V}), and revealed promising properties towards the Ile84Val mutation located at the borders of the S1/S2' and S2/S1' subpockets, respectively. Particularly, compounds bearing polar substituents in the S2/S2' pockets showed a significantly improved affinity for the mutation of isoleucine by valine (e.g., compound 3*S*,4*S*)-3,4-bis[benzyl-(4-carbamoylbenzenesulfonyl)amino]pyrrolidine (2)).¹¹ This observation was further analyzed by X-ray crystallography of the corresponding complexes of two derivatives, together with the protease variants (PDB IDs: 2R38, 2R3T, 2R3W, and 2R43), and could finally be explained by improved van der Waals (vdW) contacts, enhanced polar interactions mediated by an interstitial water molecule, and relief of steric strain likely given in the wild-type complexes.

Interestingly enough, determination of the binding affinities for the PR_{I84V} and PR_{I50V} mutants of inhibitors bearing smaller *N*-alkyl substituents ((3*S*,4*S*)-3,4-bis-(allyl-benzenesulfonyl-amino)-pyrrolidine (3), (3*S*,4*S*)-3,4-bis-[benzenesulfonyl-(2-methyl-allyl)-amino]-pyrrolidine (4), (3*S*,4*S*)-3,4-bis-[benzenesulfonyl-(3-methyl-but-2-enyl)-amino]-pyrrolidine-hydrochloride (5)) revealed a different selectivity profile (Table 1). In order to further elucidate this surprising difference in the SAR, we designed a new small series of inhibitors. Inhibitor 5,

Table 1. K_i values of the inhibitors towards the HIV protease

	P2/P2'	P1/P1'	K_i [μ M]		
			PR _{WT}	PR _{I50V}	PR _{I84V}
1			2.2	11	1.1
2			0.26	1.4	0.036
3			12	NI ^a	84
4			75	340	53
5			1.6	10	5.8
6			0.48	13	1.9
7			0.69	2.0	0.91
8			2.3	7.3	4.9
9			3.3	11	2.9

^a NI = $K_i > 500 \mu$ M.

comprising *N*-dimethylallyl substituents, which showed the best potency of the series, was picked as a reference. Its properties were further varied by decorating the sulfonamide benzyl group with para-substituents of different physicochemical properties, such as polar carboxylate and amide groups, and nonpolar bromine and cyano groups. Based on the previously observed binding modes, this moiety was thought to occupy the S2 and S2' pockets.

Results

Biological evaluation

Determination of affinities for the protease variants gave rise to a SAR that surprisingly showed no clear preference for one type of substituents (Table 1).

Whereas (3*S*,4*S*)-3,4-bis-[(4-carbamoyl-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (**6**) bearing a polar amide functionality and (3*S*,4*S*)-3,4-bis-[(4-bromo-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (**7**) bearing a hydrophobic bromide substituent gain similarly in their affinities for PR_{WT} and PR_{I84V}, substitution with a carboxylate functionality in (3*S*,4*S*)-3-[(4-carbamoyl-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-4-[(4-carboxy-benzenesulfonyl)-

(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (**9**) or with the corresponding nitrile moiety in (3*S*,4*S*)-3,4-bis-[(4-cyano-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (**8**) did not lead to an improvement in any of the cases. In comparison to PR_{WT} and PR_{I84V} binding, all compounds within this series reveal a reduced affinity for PR_{I50V}, similar to the previously studied series of *N*-benzyl-substituted derivatives. In contrast to the *N*-benzyl-substituted inhibitors, no significant improvement in affinity for PR_{I84V} is observed for all the inhibitors comprising 3,3-dimethylallyl substituents. The substitution with polar substituents in the para-position of the sulfonyl aryl moieties improves the binding affinities for this mutation in the case of the *N*-benzyl derivatives. However, in the case of dimethylallyl-substituted inhibitors, particularly for the most potent derivative **6**, no change in the relative affinity difference can be observed when compared to the unsubstituted parent structure **5**. Compared with the corresponding benzyl derivative **2**, which showed a 7-fold improvement with respect to PR_{I84V}, the now 4-fold decrease can hardly be explained regarding only the steric demand of the different *N*-alkyl moieties. To rationalize these observations in structural terms and in order to compare with the known binding mode of **2**, we crystallized compound **6** in complex with the wild-type HIV protease and consecutively determined its crystal structure.

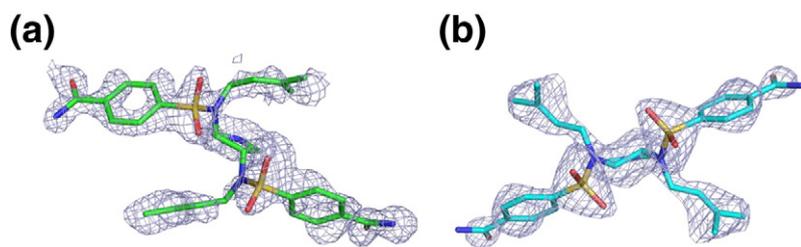


Fig. 2. The $F_o - F_c$ densities for the ligands are displayed at a σ level of 2.0 as a blue mesh. (a) Space group $P2_12_12_1$ (green; color coded by atom type). (b) Space group $P6_122$ (blue; color coded by atom type).

Structural analysis

Crystals of PR_{WT} in complex with **6** were obtained by cocrystallization of the enzyme with an inhibitor concentration of 100 μ M and a final dimethyl sulfoxide (DMSO) concentration of 10%, following our standard protocol. Crystals grew within a week and showed a deviating appearance: cubic shaped at a precipitant concentration of 3.5 M NaCl and needle-like shaped at a concentration of 3 M NaCl. Whereas the cubic crystals exhibited space group $P2_12_12_1$, the needle-shaped crystal form corresponded to space group $P6_122$. Both structures could be determined with a resolution of 1.65 \AA , and ligand geometries were clearly visible in the $F_o - F_c$ omit map at a σ level of 2 (Fig. 2a and b). However, two completely different binding modes were observed and will be described separately: the geometry in space group $P2_12_12_1$ is referred to as orthorhombic binding mode, whereas the binding mode in $P6_122$ is named hexagonal binding mode (Table 2).

Table 2. Crystallographic data

	$P2_12_12_1$	$P6_122$
Resolution (\AA)	25–1.65	25–1.65
Cell dimensions a, b, c (\AA)	52.0, 57.6, 61.4	62.5, 62.5, 82.3
Highest-resolution shell (\AA)	1.68–1.65	1.68–1.65
Number of measured reflections	99,069	80,672
Number of independent reflections	22,526	11,774
Completeness (%) ^a	98.8 [89.4]	97.8 [96.2]
I/σ^a	26.1 [2.7]	23.2 [4.3]
R_{sym} (%) ^a	5.2 [35.0]	8.0 [36.8]
Resolution in refinement (\AA)	10–1.65	10–1.65
R_{cryst} ($F > 4\sigma F_o$; F_o)	18.3; 20.2	22.4; 23.4
R_{free} ($F > 4\sigma F_o$; F_o)	23.9; 26.2	27.4; 29.7
Mean B -factor (\AA^2) (peptide chain A; peptide chain B)	18.3; 17.3	30.5
Main chain (\AA^2) (peptide chain A; peptide chain B)	15.6; 15.5	27.5
Side chain (\AA^2) (peptide chain A; peptide chain B)	21.2; 19.4	33.9
Ligand (\AA^2)	32.4	38.2
Cl (\AA^2)	19.2	—
Water (\AA^2)	26.1	33.5
Ramachandran plot		
Most favored geometry (%)	96.2	94.9
Additionally allowed (%)	3.8	5.1

Values in brackets refer to the highest-resolution shell.

Binding mode in the orthorhombic form

In the orthorhombic space group, inhibitor **6** exhibits a binding mode closely related to the one observed previously for all our pyrrolidine-3,4-bis-*N*-benzyl-sulfonamides. The endocyclic amino functionality addresses the catalytic dyad and forms a hydrogen-bond network with aspartic acid residue 25A (2.8 \AA /2.8 \AA) and aspartic acid residue 25B (2.8 \AA /2.9 \AA). Flap interactions are established via hydrogen bonding of only one of the inhibitor's sulfonyl oxygen atoms to the main-chain NH of Ile50A (2.6 \AA). The second sulfonyl oxygen atom of this sulfonamide group remains uninvolved in any polar interaction (Fig. 3).

Also, the subsite occupancy is comparable to the previously determined complexes. Whereas the dimethylallyl substituents occupy the S_1/S_1' subpockets establishing contacts with Gly27A, Gly48A, Gly49A, Leu23B, Pro81B, Val82B, Ile84B (S_1) and Leu23A, Gly27B, and Ile84A (S_1'), the S_2/S_2' pockets are addressed by the phenyl moieties of **6** forming numerous vdW contacts with Ala28A, Val32A, Ile47A, Gly48A (S_2); and Ile50A, Ala28B, Gly48B, Ile84B (S_2'). The amide functionalities form hydrogen bonds with the main-chain NH of Asp30A (2.6 \AA) and Asp30B (2.7 \AA), and additionally in the S_2' pocket to the corresponding Asp30B side chain (2.8 \AA).

The C^α superposition of this structure with the crystal structure of the corresponding *N*-benzyl-substituted inhibitor (PDB ID: 2PWR) reveals that the overall binding mode is only slightly affected by the exchange of the *N*-benzyl moiety with the *N*-dimethylallyl moiety, which is reflected by a root-mean-square deviation (RMSD) of 0.53 between the C^α atoms of the complexes and by an RMSD of 0.60 \AA for the identical ligand atoms of **2** and **6**. The ligand is deeply buried in the binding pocket (90% of its solvent-accessible surface is buried).

Binding mode in the hexagonal form

The binding mode observed in the hexagonal space group deviates dramatically from the previously described binding mode, in agreement with all examples yet observed for our pyrrolidine-based HIV protease inhibitors. In the hexagonal case, the complex of the C_2 -symmetric protease and the C_2 -

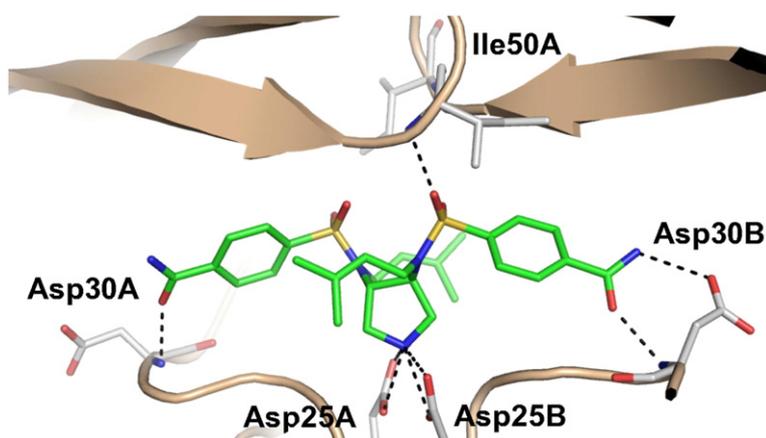


Fig. 3. Crystal structure of **6** (green; color coded by atom type) in complex with HIV protease ($P2_12_12_1$). The protein backbone trace is schematically illustrated in wheat, and the catalytic Asp25A and Asp25B, Ile50A, Ile50B, Asp30A, and Asp30B are displayed in gray (color coded by atom type). Hydrogen bonds are indicated by broken lines.

symmetric inhibitor **6** fully agrees with the C_2 symmetry (Fig. 4).

In this space group, the asymmetric unit comprises one monomer of HIV-1 protease. The functional dimer is generated via a 2-fold crystallographic symmetry. For purposes of comparison, amino acids of protein chain A and their symmetry equivalents will be arbitrarily referred to as 1A–99A and 1B–99B, respectively. According to the imposed symmetry, any ligand interactions with the A and B chains are identical. Similar to the orthorhombic case, the catalytic aspartates Asp25A and Asp25B are addressed by the endocyclic pyrrolidine nitrogen of **6** (2.8 Å/2.9 Å). However, this is the only conserved feature shared by the two binding modes. In contrast to the orthorhombic mode, each sulfonyl group of **6** forms a direct hydrogen bond with Ile50A and Ile50B, respectively, with one sulfonyl oxygen atom of each sulfonamide group (3.0 Å). The *para*-carboxamido-benzenesulfonamide moieties occupy the S1 and S1' pockets forming vdW interactions with Leu23, Gly27, Gly49, Pro81, Val82, and Ile84, and polar interactions with Arg8, with the amide

substituent (3.0 Å/3.2 Å) of chains A and B, respectively. The S2 and S2' pockets are now occupied by the dimethylallyl moieties establishing vdW contacts with Ala28, Val32, Ile47, and Ile50 of chains A and B. Comparison of the complex with the corresponding complex of the *N*-benzyl-substituted inhibitor using a C^α superposition reveals a high level of similarity of the protein structures reflected by an RMSD of 0.60 Å, but a high dissimilarity for the identical ligand atoms (RMSD=5.4 Å). Similarly to the orthorhombic case, the ligand is deeply buried in the binding pocket (86% of its solvent-accessible surface is buried).

Discussion

Investigation of the inhibitory potency of a series of pyrrolidine-based inhibitors against mutant HIV protease variants resulted in a remarkably different selectivity profile of the *N*-allyl-substituted derivatives compared to the *N*-benzyl-substituted initial lead structure. To elucidate the reasons for this opposing

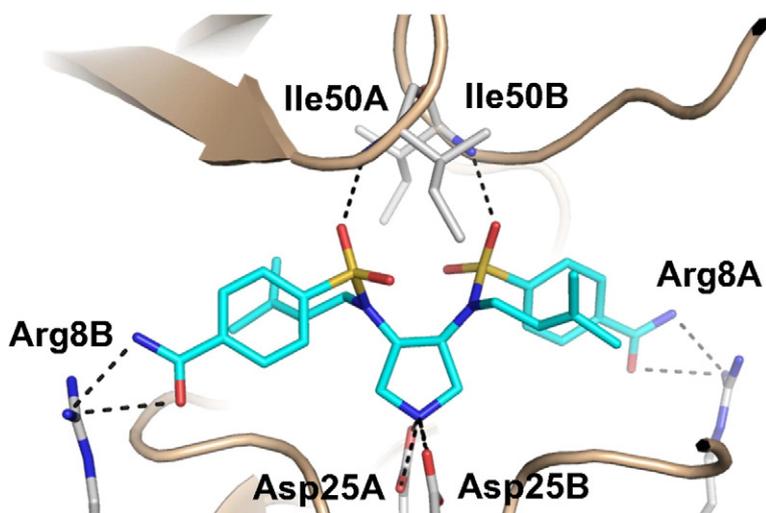


Fig. 4. Crystal structure of **6** (light blue; color coded by atom type) in complex with HIV protease ($P6_12_2$). For purposes of comparison, amino acids of protein chain A and their symmetry equivalents will be arbitrarily referred to as 1A–99A and 1B–99B, respectively. The protein backbone trace is schematically illustrated in wheat, and the catalytic Asp25A and Asp25B, Ile50A, Ile50B, Arg8A, and Arg8B are displayed in gray (color coded by atom type). Hydrogen bonds are indicated by broken lines.

behavior especially against the Ile84Val mutant, we synthesized a small series of inhibitors based on the most potent derivative **5** of the initial series.

However, the nonuniform SAR obtained within this series, comprising substituents with different physicochemical properties in the para-position of the inhibitor's sulfonamide aryl groups, did not match the activity pattern expected from the binding mode of the *N*-benzyl derivatives. For a detailed analysis of this phenomenon and for a comparison with the corresponding *N*-benzyl derivative **2**, **6** in complex with HIV protease was selected for crystal structure determination. To our surprise, protein crystals were obtained in two crystal forms under very similar crystallization conditions. Both exhibited a space group deviating from the former series, in which 10 complexes of *N*-benzyl-substituted inhibitors revealed identical crystal packing in $P2_12_12_1$.

The analysis of the complex in space group $P2_12_12_1$ revealed a binding mode closely resembling the one observed previously for all examples of the *N*-benzyl series. Due to the similarity of the surface complementarities observed for this orthorhombic binding mode to those previously determined for **2**, particularly in the contact area to Ile84 (Fig. 5a and c), a convincing explanation for the deviating selectivity profile is difficult to provide.

However, the structural evidence found in the second hexagonal crystal form suggests that a different binding mode can possibly be adopted

also under biologically relevant conditions. This binding mode reveals an inhibitor orientation much closer to that found for all approved peptidomimetic inhibitors, which are all known to exhibit susceptibility towards the I84V substitution (Fig. 5b and d). The commonly accepted assumption that similar ligands bind in a similar fashion can definitely be ruled out in this case, at least considering the crystalline state. Regarding the two different binding modes, it is not evident which orientation will be preferred under biological conditions. The nature of the substituent and the different binding pocket shape of the mutant enzymes might induce different preferences, resulting in an overall nonuniform SAR. Whether the adopted space groups impose special symmetry constraints onto the deviating binding modes or whether the preferred binding mode is reinforced by the crystal packing cannot be resolved at this point. A comparison of protein-protein contacts due to crystal packing, as well as the additional observation of two chloride ions present in the $P2_12_12_1$ complex but missing in the $P6_122$ packing, which exhibit at least a distance of 10 Å to the closest ligand atom, however, does not satisfactorily explain why the binding mode observed in the orthorhombic case should not be possible in the hexagonal $P6_122$ space group and vice versa. Both binding modes seem to possess a very close-by energy content; however, they obviously trigger spontaneously by crystallization.

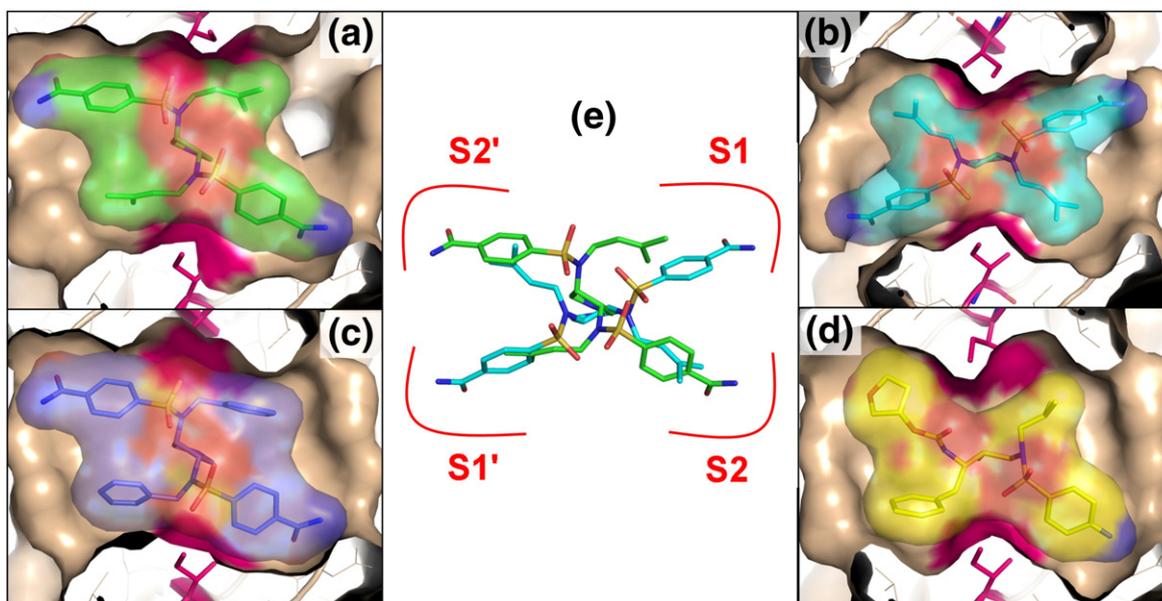


Fig. 5. Crystallographically observed binding modes in HIV protease in equal orientation and size. (a–d) Surface representation of the ligand (color coded by atom type) and the corresponding protein structure (wheat); view from the top with the flap region clipped off. Ile84 is highlighted in hot pink. (a) Orthorhombic binding mode of **6**. (b) Hexagonal binding mode of **6**. (c) Binding mode of **2** (PDB ID: 2PWR). (d) Binding mode of amprenavir (PDB ID: 1HPV). (e) Superposition of the ligand orientations observed in the corresponding orthorhombic (green) and hexagonal (blue) crystal structures of **6** aligned by a C^α fit of the protein coordinates.

Summary and Conclusion

Recently, we developed *N*-benzyl-substituted HIV protease inhibitors and investigated in detail their behavior toward the active-site mutants PR_{I50V} and PR_{I84V}. The observation of an overall increase in affinity for PR_{I84V} was, among others, attributed to relief of steric strain present in the wild-type complexes. The series 3–9 presented in this contribution was intended to elucidate whether smaller hydrophobic residues might be better suited to address the S1 and S1' pockets appropriately. However, an inconsistent SAR demanded the determination of a cocrystal structure of at least one derivative bearing such modifications. Remarkably, two different crystal structures of the most potent derivative 6, resulting from two different crystal forms found serendipitously, could be determined. Remarkably, they could be grown under very similar crystallization conditions, exhibiting only a tiny concentration difference in the precipitating NaCl concentration. Both crystal forms exhibited comparable diffraction qualities; in each case, a well-defined but distinct inhibitor orientation could be identified. To our surprise, the complexes of identical overall compositions experience different protein–inhibitor interactions. Only the interactions with the catalytic dyad are similar, along with a comparable overall burial of the ligands. The observation of two different binding modes for 6 suggests that also other members of this series could occur with more than one binding mode, thus explaining the nonconsistent SAR within the series. Are the results obtained by this study a peculiar, rarely occurring phenomenon or are they of fundamental importance also for other structure-based drug discovery projects? There might be some statistical evidence for the widely accepted opinion that similar ligands bind in similar fashion; however, this is by no means a “truism.” The question remains: How often does a similar behavior, as found in this study, remain unnoticed in a congeneric ligand series? Thoroughly and exhaustively performed structural biology has to be an integral part of drug discovery projects and has to escort the entire optimization process. The same holds for biological evaluation, which has to be interpreted in a very careful manner to identify possible changes in binding orientations during lead optimization. In fact, the observation of a different binding mode for members of a series might even provide unexpected opportunities to optimize ligands towards a particular binding profile. Furthermore, such ligands might exhibit a remarkable advantage to escape resistance development. Likely, both binding modes are close in energy. Mutational changes can affect and influence either of the two modes and can change the energetic balance. However, an inhibitor that is capable of binding in two close-by alternative modes might

swap from one to the other in order to escape resistance development. A similar concept has been followed during the development of TMC125-R165335 (etravirine) and TMC120-R147681 (dapivirine), two novel allosteric site inhibitors of the reverse transcriptase with remarkable resistance profile.¹² The authors describe such an advantage to overcome resistance development by the ability to bind the target enzyme in multiple conformations and to more efficiently escape structural modifications of the resistance-mutated protein.

Experimental

Synthesis

Inhibitors 1–5 have been synthesized as previously described.¹⁰ Inhibitors 6–9 were prepared from the common intermediate (3*S*,4*S*)-3,4-bis-[(4-bromo-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidine-1-carboxylic acid *t*-butyl ester (12) following the sequence outlined below. The reaction of 3,4-diamino-pyrrolidine (10) with 4-bromo-phenylsulfonyl chloride renders the corresponding sulfonamide (3*S*,4*S*)-3,4-bis-(4-bromo-benzenesulfonylamino)-pyrrolidine-1-carboxylic acid *t*-butyl ester (11), which is then alkylated with 3,3-dimethylallyl bromide, thus leading to the key intermediate 12 (Fig. 6).

Deprotection of 12 under acidic nonaqueous conditions renders 7. Substitution of the 4-bromo substituent employing Zn(CN)₂, Pd[P(C₆H₅)₃]₄ under microwave irradiation required prolonged reaction times and high temperatures, which resulted in a concomitant cleavage of the BOC-protecting group and thus directly furnished inhibitor 8.¹³ Intermediate 8 was further utilized for the preparation of the inhibitors 6 and 9. Saponification of the nitrile functionality in 8 under basic conditions and microwave irradiation yielded the corresponding acid 9. Mild hydrolysis with 30% hydrogen peroxide in DMSO/K₂CO₃ gave rise to the carboxamido derivative 6.¹⁴

General

Reported yields refer to the analytically pure product obtained by column chromatography. Flash chromatography was performed using silica gel 60 (0.04–0.063 mm) purchased from Macherey-Nagel. For reversed-phase chromatography, prepacked RP-18 columns (Redisep-C-18, 13 g) were used. Commercially available solvents and reagents were used without further purification. Microwave reactions were performed in the laboratory microwave system Discover (LabMate) from CEM. All proton and carbon NMR spectra were recorded on a Jeol Eclipse+ Spectrometer. ¹H NMR spectra were referenced to CDCl₃ (7.26 ppm) or DMSO-*d*₆ (2.50 ppm). ¹³C NMR spectra were referenced to CDCl₃ (77.00 ppm) or DMSO-*d*₆ (39.70 ppm). The values of chemical shifts (δ) are given in parts per million, and coupling constants (*J*) are given in hertz (br=broad; ps=pseudo; s=singlet; d=doublet; t=triplet; q=quartet; sm=symmetric multiplet; m=multiplet). Mass spectra were obtained from a double-

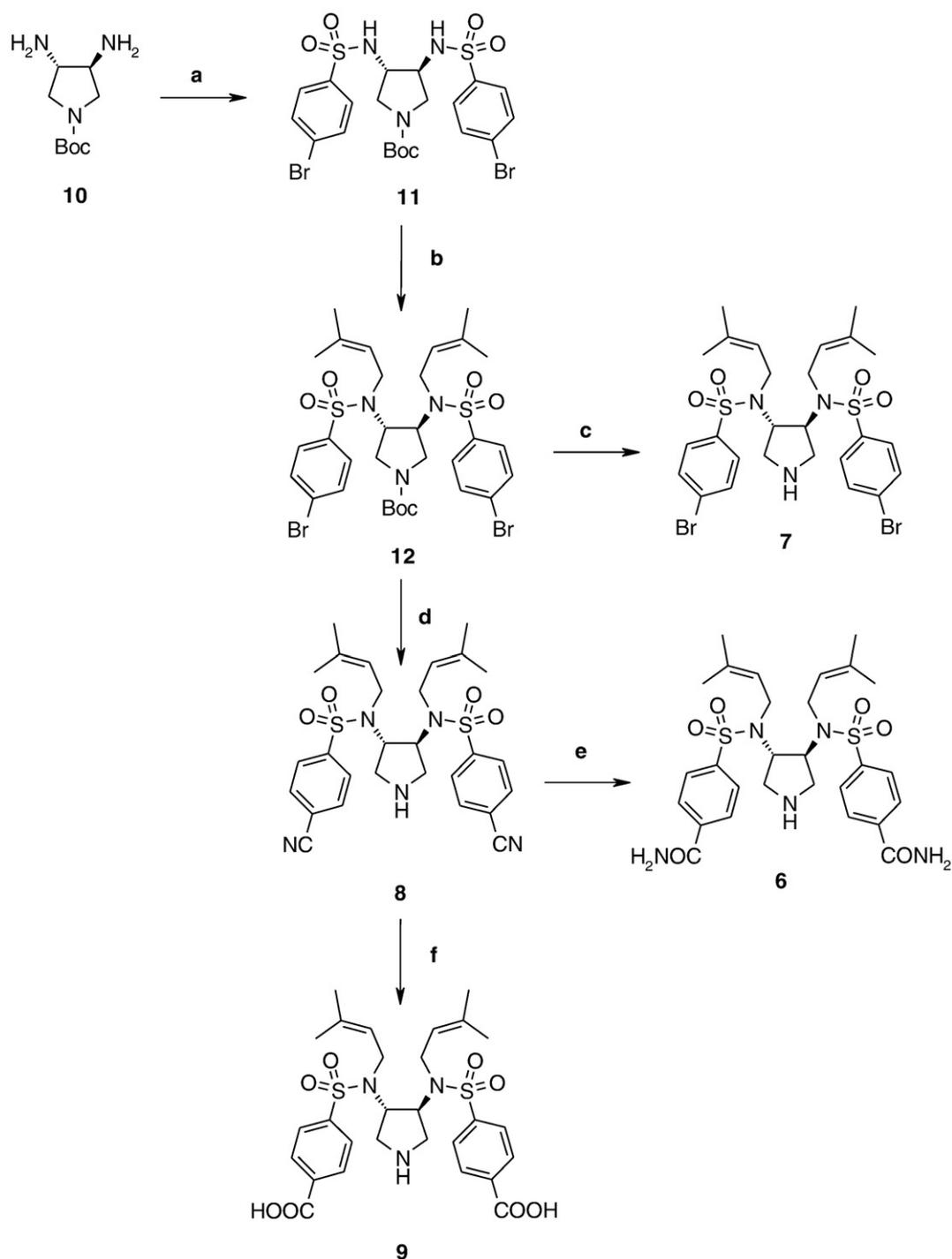


Fig. 6. Preparation of the inhibitors. The BOC-protected 3,4-diamino-pyrrolidine was prepared from D-(-)-tartaric acid, as previously described.¹⁰ (a) 4-Br-PhSO₂Cl, NEt₃, CH₂Cl₂, 65%. (b) 3,3-Dimethylallyl bromide, K₂CO₃, CH₃CN, 0–20 °C, 79%. (c) CH₂Cl₂/CF₃COOH 1:1, 74%. (d) Zn(CN)₂, Pd(PPh₃)₄, dimethylformamide, μW, 51%. (e) 30% H₂O₂, K₂CO₃, DMSO, 38%. (f) NaOH, H₂O, μW, 46%.

focusing sector field Micromass VG-Autospec spectrometer. Combustion analyses were determined on a Vario Micro Cube (Elementar Analysen GmbH). Infrared spectra were obtained on a Nicolet FT-IR Spectrometer 510 P.

(3*S*,4*S*)-3,4-bis-(4-Bromo-benzenesulfonylamino)-pyrrolidine-1-carboxylic acid t-butyl ester (11)

A solution of triethylamine (1.87 g, 18.5 mmol, 2.4 Eq) was added to an ice-cold solution of 4-bromo-

benzenesulfonyl chloride (4.33 g, 17.0 mmol, 2.2 Eq) in dichloromethane (40 mL), followed by the dropwise addition of **10** (1.64 g, 7.71 mmol, 1.0 Eq) dissolved in dichloromethane (5 mL), with the temperature maintained. The reaction mixture was allowed to reach ambient temperature and was stirred for an additional 12 h. The reaction was quenched with water and washed three times with 10 mL of an aqueous solution of HCl (1%). The combined organic layers were dried over MgSO₄, filtered, and concentrated under reduced pressure. Column chromatography (3:1 t-butyl methyl ether/hexanes containing 0.1% of triethylamine) rendered **11** (3.23 g, 65%): ¹H NMR (400 MHz, CDCl₃) δ=7.86–7.62 (brd, 8H), 5.43 (brs, 1H), 5.26 (brs, 1H), 3.72–3.44 (brs, 4H), 3.08–2.94 (brs, 2H), 1.39 (s, 9H); ¹³C NMR (100 MHz, CDCl₃) δ=153.2, 140.2, 132.5, 128.7, 128.6, 126.6, 78.9, 56.8, 56.0, 48.7, 48.3, 28.1; mass spectrometry (MS) [electrospray ionization (ESI)] *m/z* (%): 662 (100) [M+Na]⁺; analysis calculated for C₂₁H₂₅Br₂N₃O₆S₂: C 39.45, H 3.94, N 6.57; found: C 39.87, H 4.21, N 6.62.

(3S,4S)-3,4-bis-[(4-Bromo-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidine-1-carboxylic acid t-butyl ester (12**)**

Anhydrous K₂CO₃ (1.10 g, 8.0 mmol, 1.6 Eq) was added to a solution of sulfonamide **11** (3.20 g, 5.0 mmol, 1.0 Eq) in acetonitrile (40 mL), and the resulting mixture was cooled to 0 °C, upon which dimethylallyl bromide (1.79 g, 12.0 mmol, 1.2 Eq) was carefully added dropwise. The reaction mixture was allowed to reach ambient temperature, and stirring was continued for 12 h. After addition of 20 mL of t-butyl methyl ether, the reaction mixture was filtered, and the filtrate was concentrated under reduced pressure. Column chromatography (5:1 hexanes/t-butyl methyl ether) of the oily residue gave rise to **12** (3.08 g, 79%): ¹H NMR (400 MHz, DMSO-*d*₆) δ=7.81 (brd, *J*=8.7 Hz, 4H), 7.76–7.70 (brd, 4H), 4.89 (brs, 1H), 4.74 (brs, 1H), 4.61 (brs, 2H), 3.83 (brs, 2H), 3.72 (brs, 2H), 3.25 (dd, *J*=10.3 Hz, 8.2 Hz, 2H), 3.04–2.86 (brs, 2H), 1.62 (s, 6H), 1.56 (s, 6H), 1.35 (s, 9H); ¹³C NMR (100 MHz, DMSO-*d*₆, rotamers) δ=153.3, 140.2, 134.7, 132.3, 129.2, 126.8, 121.4, 121.0, 79.1, 56.9, 56.4, 44.4, 44.0, 41.8, 28.2, 25.5, 17.7; MS (ESI) *m/z* (%): 798 (100) [M+Na]⁺; analysis calculated for C₃₁H₄₁Br₂N₃O₆S₂: C 48.01, H 5.33, N 5.42; found: C 48.23, H 5.42, N 5.31.

(3S,4S)-3,4-bis-[(4-Bromo-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (7**)**

Neat trifluoroacetic acid (17.1 mg, 11.2 μL, 0.15 mmol, 1.0 Eq) was added to a solution of **12** (117 mg, 0.15 mmol, 1.0 Eq) in dichloromethane (3 mL), and the resulting reaction mixture was stirred under a positive stream of argon for 12 h at ambient temperature, upon which the reaction mixture was concentrated *in vacuo*. Reversed-phase column chromatography using a gradient of water/CH₃CN containing 0.1% of CF₃COOH (100% water within 5 min, linear gradient to 100% CH₃CN within 55 min, then additional 10 min of pure CH₃CN) afforded **7** (88 mg, 74%) as its trifluoroacetate salt: ¹H NMR (500 MHz, DMSO-*d*₆) δ=7.83 (dt, *J*=8.7 Hz, 2.5 Hz, 4H), 7.70 (dt, *J*=8.9 Hz,

2.3 Hz, 4H), 4.84 (brt, *J*=6.6 Hz, 2H), 4.66 (brt, *J*=6.1 Hz, 2H), 3.79 (d, *J*=6.6 Hz, 4H), 3.26–3.19 (sm, 2H), 3.01–2.92 (sm, 2H), 1.64 (s, 6H), 1.56 (s, 6H); ¹³C NMR (100 MHz, DMSO-*d*₆) δ=158.7 (q, ²*J*_{C,F}=31.6 Hz), 139.7, 135.3, 132.5, 129.2, 127.1, 120.6, 117.2 (q, ¹*J*_{C,F}=298.1 Hz), 56.1, 42.4, 42.0, 25.5, 17.7; MS (ESI) *m/z* (%): 676 (34) [M+H]⁺; analysis calculated for C₂₆H₃₃Br₂N₃O₄S₂-CF₃COOH: C 42.60, H 4.34, N 5.32; found: C 42.88, H 4.63, N 5.22.

(3S,4S)-3,4-bis-[(4-Cyano-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (8**)**

Zinc cyanide (130 mg, 1.1 mmol, 1.0 Eq) and tetrakis (triphenylphosphine)palladium (38 mg, 0.033 mmol, 0.03 Eq) were added to a solution of **12** (854 mg, 1.1 mmol, 1.0 Eq) in dimethylformamide (5 mL) under a positive stream of argon. The reaction vessel was sealed, and the reaction mixture was heated in a laboratory microwave for 30 min at 175 °C and 300 W. After having cooled down to ambient temperature, the reaction mixture was diluted with t-butyl methyl ether (30 mL) and washed twice with water (20 mL) and brine. The organic layer was dried over MgSO₄, filtered, and concentrated under reduced pressure. Reversed-phase column chromatography of the oily residue using a gradient of water/CH₃CN containing 0.1% of CF₃COOH (100% water within 5 min, linear gradient to 100% CH₃CN within 55 min, then additional 10 min of pure CH₃CN) afforded **8** (385 mg, 51%) as its trifluoroacetate salt: ¹H NMR (500 MHz, CD₃OD) δ=8.02 (psd, *J*=8.7 Hz, 4H), 7.97 (psd, *J*=8.7 Hz, 4H), 4.86–4.79 (m, 4H), 3.95 (dd, *J*=16.3 Hz, 7.3 Hz, 2H), 3.87 (dd, *J*=16.4 Hz, 6.4 Hz, 2H), 3.54–3.48 (m, 2H), 3.29–3.25 (m, 2H), 1.70 (s, 6H), 1.60 (s, 6H); ¹³C NMR (100 MHz, DMSO-*d*₆) δ=158.8 (q, ²*J*_{C,F}=33.6 Hz), 144.7, 135.8, 133.5, 127.8, 120.2, 117.8, 116.6 (q, ¹*J*_{C,F}=295.2 Hz), 115.5, 56.3, 42.4, 42.1, 25.4, 17.7; MS (ESI) *m/z* (%): 568 (100) [M+H]⁺; infrared (KBr): ν=2200 (CN) cm⁻¹.

(3S,4S)-3,4-bis-[(4-Carbamoyl-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (6**)**

A solution of **8** (171 mg, 0.25 mmol, 1.0 Eq) in DMSO was cooled to 0 °C. An aqueous H₂O₂ solution (30%, 1.5 mL) and K₂CO₃ (507 mg, 0.38 mmol, 1.5 Eq) were added to this frozen mixture, and the reaction mixture was allowed to slowly reach ambient temperature. Stirring was continued for an additional 30 min. The reaction mixture was again cooled to 0 °C and quenched by the addition of water (50 mL), upon which the product precipitated. The precipitate was collected by filtration, dissolved in a small amount of methanol, and further purified by reversed-phase chromatography using the same gradient as described above, giving rise to **6** (68 mg, 38%): ¹H NMR (500 MHz, DMSO-*d*₆) δ=9.01 (brs, 2H), 8.20 (brs, 2H), 8.07 (d, *J*=8.7 Hz, 4H), 7.87 (d, *J*=8.5 Hz, 4H), 7.63 (brs, 2H), 4.86 (pst, *J*=6.7 Hz, 2H), 4.73 (brt, *J*=6.2 Hz, 2H), 3.81 (brd, *J*=6.6 Hz, 4H), 3.27–3.19 (m, 2H), 3.01–2.92 (m, 2H), 1.63 (s, 6H), 1.54 (s, 6H); ¹³C NMR (100 MHz, DMSO-*d*₆) δ=166.5, 158.8 (q, ²*J*_{C,F}=31.6 Hz), 142.6, 138.2, 135.3, 128.5, 127.2, 120.6, 117.1 (q, ¹*J*_{C,F}=299.0 Hz), 56.2,

42.4, 42.0, 25.5, 17.7; MS (ESI) m/z (%): 604 (100) $[M+H]^+$; analysis calculated for $C_{28}H_{37}N_5O_6S_2 \cdot CF_3COOH \cdot H_2O$: C 48.97, H 5.48, N 9.52; found: C 49.12, H 5.44, N 9.31.

(3S,4S)-3-[(4-Carbamoyl-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-4-[(4-carboxy-benzenesulfonyl)-(3-methyl-but-2-enyl)-amino]-pyrrolidinium trifluoroacetate (9)

An aqueous NaOH solution (25%, 2 mL) was added to a solution of **8** (123 mg, 0.18 mmol, 1.0 Eq) in 2-propanol (2 mL). The reaction vessel was sealed, and the reaction mixture was heated in a laboratory microwave for 90 min at 100 °C and 300 W. After having cooled down to ambient temperature, the reaction mixture was neutralized using aqueous HCl. The solution was concentrated *in vacuo* to a residual volume of 0.5 mL. Reversed-phase column chromatography using the same gradient as described above afforded **9** (60 mg, 46%) as its trifluoroacetate salt: 1H NMR (500 MHz, $DMSO-d_6$) δ =8.12 (d, J =8.5 Hz, 4H), 7.90 (d, J =8.5 Hz, 4H), 4.82 (t, J =6.5 Hz, 2H), 4.73 (t, J =6.2 Hz, 2H), 3.82 (d, J =6.6 Hz, 4H), 3.24 (dd, J =11.3 Hz, 7.5 Hz, 2H), 2.98 (dd, J =11.1 Hz, 9.0 Hz, 2H), 1.63 (s, 6H), 1.53 (s, 6H); ^{13}C NMR (100 MHz, $DMSO-d_6$) δ =166.3, 158.9 (q, $^2J_{C,F}$ =31.6 Hz), 144.0, 135.4, 134.8, 130.2, 127.4, 120.5, 117.2 (q, $^1J_{C,F}$ =298.1 Hz), 56.2, 42.4, 41.9, 25.4, 17.7; MS (ESI) m/z (%): 606 (64) $[M+H]^+$; analysis calculated for $C_{28}H_{35}N_5O_8S_2 \cdot CF_3COOH$: C 50.06, H 5.04, N 5.84; found: C 50.10, H 5.23, N 5.82.

Kinetic assay

Enzymatic assays were performed in 172 μ L of assay buffer [100 mM 4-morpholineethanesulfonic acid, 300 mM KCl, 5 mM ethylenediaminetetraacetic acid, and 1 mg/mL bovine serum albumin (pH 5.5)] by the addition of substrate dissolved in 4 μ L of DMSO, distinct inhibitor concentrations dissolved in 4 μ L of DMSO, and 20 μ L of HIV-1 protease in assay buffer, to a final volume of 200 μ L (final DMSO concentration, 4%). The fluorogenic anthranilyl-HIV protease substrate (Abz-Thr-Ile-Nle-(*p*-NO₂-Phe)-Gln-Arg-NH₂) was purchased from Bachem. The hydrolysis of the anthranilyl-HIV protease substrate was recorded as the increase in fluorescence intensity (excitation wavelength, 337 nm; emission wavelength, 410 nm).¹⁵ The kinetic parameters of PR_{WT} (K_m =14.6 μ M), PR_{I50V} (K_m =139 μ M), and PR_{I84V} (K_m =70 μ M) were determined by the Lineweaver-Burk method. IC₅₀ values were generated by a nonlinear regression analysis from plots of v_i/v_0 versus inhibitor concentration, in which v_i is the velocity in the presence of an inhibitor, and v_0 is the velocity in the absence of an inhibitor. K_i values were calculated from the following equation: $K_i = IC_{50}/[1 + (S/K_m)]$.

Crystallization of HIV-1 protease

Crystals of HIV-1 protease in complex with **6** were obtained by cocrystallization using the sitting-drop vapor diffusion method. The well buffer [0.1 M 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol (pH 6.5) and 3–3.5 M NaCl] was mixed with 1 μ L of protein solution [50 mM NaAc (pH 6.5), 1 mM ethylenediamine-

traacetic acid, and 1 mM DTT] at an HIV protease concentration of 7 mg/mL. Crystals corresponded to space group $P2_12_12_1$ at a precipitant concentration of 3.5 M NaCl and to space group $P6_122$ at 3 M NaCl. For cryoprotection, both crystal forms were briefly soaked in mother liquor containing 25% glycerol.

Data collection, phasing, and refinement

The data sets were collected at the synchrotron BESSY II (Berlin, Germany) on PSF beamline 14.2. Data were processed and scaled with Denzo and Scalepack, as implemented in HKL2000.¹⁶ The structures were determined by the molecular replacement method with Phaser;¹⁷ one monomer of the 1.5-Å structure of the HIV-1 protease in complex with a pyrrolidine-based inhibitor (PDB ID: 2PQZ) was used as the search model. Refinement was continued with SHELXL-97;¹⁸ for each refinement step, at least 10 cycles of conjugate gradient minimization were performed, with restraints on bond distances, angles, and B -values. Intermittent cycles of model building were performed with the program Coot.¹⁹

Surface calculations

The buried surface area was calculated with the program MS developed by Connolly.²⁰ The program calculates the molecular surface from atom coordinates; the probe radius was set to 1.4 Å.²⁰

PDB accession numbers

The coordinates have been deposited in the PDB[‡] with access codes 3CKT (orthorhombic) and 2ZGA (hexagonal).

Acknowledgements

We gratefully acknowledge financial support from the joint DFG research projects KL1204/5-3 and DI827/3-2, and from the Fonds der Chemischen Industrie (A.B.).

We acknowledge the support of the beamline staff at BESSY II, and we thank the Bundesministerium für Bildung und Forschung (05 ES3XBA/5) for financing the travel costs to BESSY II. The pET11a plasmid containing the HIV-1 protease gene was kindly provided by Professor Helena Danielson (University of Uppsala, Uppsala, Sweden).

[‡] <http://www.rcsb.org/pdb/>

References

1. Beddell, C. R., Goodford, P. J., Norrington, F. E., Wilkinson, S. & Wootton, R. (1976). Compounds designed to fit a site of known structure in human haemoglobin. *Br. J. Pharmacol.* **57**, 201–209.
2. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R. *et al.* (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
3. Hubbard, R. E. (2005). 3D structure and the drug-discovery process. *Mol. BioSyst.* **1**, 391–406.
4. Anderson, A. C. (2003). The process of structure-based drug design. *Chem. Biol.* **10**, 787–797.
5. Suvit Thaisrivongs, J. W. S. (1999). Structure-based discovery of tipranavir disodium (PNU-140690E): a potent, orally bioavailable, nonpeptidic HIV protease inhibitor. *Pept. Sci.* **51**, 51–58.
6. Reich, S. H., Melnick, M., Davies, J. F., II, Appelt, K., Lewis, K. K., Fuhry, M. A. *et al.* (1995). Protein structure-based design of potent orally bioavailable, nonpeptide inhibitors of human immunodeficiency virus protease. *Proc. Natl Acad. Sci. USA*, **92**, 3298–3302.
7. Stubbs, M. T., Reyda, S., Dullweber, F., Möller, M., Klebe, G., Dorsch, D. *et al.* (2002). pH-dependent binding modes observed in trypsin crystals: lessons for structure-based drug design. *ChemBioChem*, **3**, 246–249.
8. Gassel, M., Breitenlechner, C. B., König, N., Huber, R., Engh, R. A. & Bossemeyer, D. (2004). The protein kinase C inhibitor bisindolyl maleimide 2 binds with reversed orientations to different conformations of protein kinase A. *J. Biol. Chem.* **279**, 23679–23690.
9. Bostrom, J., Hogner, A. & Schmitt, S. (2006). Do structurally similar ligands bind in a similar fashion? *J. Med. Chem.* **49**, 6716–6725.
10. Blum, A., Böttcher, J., Heine, A., Klebe, G. & Diederich, W. E. (2008). Structure-guided design of C₂-symmetric HIV-1 protease inhibitors based on a pyrrolidine scaffold. *J. Med. Chem.* **51**, 2078–2087.
11. Böttcher, J., Blum, A., Heine, A., Diederich, W. E. & Klebe, G. (2008). Structural and kinetic analysis of pyrrolidine-based inhibitors of the drug-resistant Ile84Val mutant of HIV-1 protease. *J. Mol. Biol.* **383**, 347–357.
12. Das, K., Clark, A. D., Lewi, P. J., Heeres, J., deJonge, M. R., Koymans, L. M. H. *et al.* (2004). Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. *J. Med. Chem.* **47**, 2550–2560.
13. Alterman, M. & Hallberg, A. (2000). Fast microwave-assisted preparation of aryl and vinyl nitriles and the corresponding tetrazoles from organo-halides. *J. Org. Chem.* **65**, 7984–7989.
14. Katritzky, A., Pilarski, B. & Urogdi, L. (1989). Efficient conversion of nitriles to amides with basic hydrogen peroxide in dimethyl sulfoxide. *Synthesis*, **12**, 949–950.
15. Toth, M. V. & Marshall, G. R. (1990). A simple, continuous fluorometric assay for HIV protease. *Int. J. Pept. Protein Res.* **36**, 544–550.
16. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology* (Carter, C. W., ed.), Vol. 276, pp. 307–326. Academic Press, New York, NY.
17. Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr. Sect. D*, **60**, 432–438.
18. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. In *Methods in Enzymology* (Charles, W. C. J. & Robert, M. S., eds), Vol. 277, pp. 319–343. Academic Press, New York, NY.
19. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D*, **60**, 2126–2132.
20. Connolly, M. (1983). Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548–558.