DOI: 10.1002/cmdc.200500063

A Potent HIV Protease Inhibitor Identified in an Epimeric Mixture by High-Resolution Protein Crystallography**

Silvano Geremia,*[a] Nicola Demitri,[a] Jochen Wuerges,[a] Fabio Benedetti,[a] Federico Berti,[a] Gianluca Tell,[b] and Lucio Randaccio[a]

Structure-based drug design[1] uses structural data obtained by protein crystallography,[2] NMR spectroscopy,[3] and computational chemistry[4] to guide the synthesis of potential drugs. The structural information can be used to explain the basis of their activity (SAR) and to improve the potency and specificity of new lead compounds.[5] Indeed, there are currently several drugs on the market that originated from this structure-based approach to drug design. The most commonly cited are anti-HIV drugs such as amprenavir (Agenerase) and nelfinavir (Viracept), which were developed by using the structural information obtained from the X-ray crystallographic structure of HIV-1 protease, an essential hydrolase in the retroviral life cycle.[5]

The use of X-ray and NMR techniques has been extended beyond pure structural characterization to new approaches for structure-based lead discovery.[6,7] Both NMR-[8] and X-ray-based[9] screening of mixtures have been introduced in pharmaceutical research. In general, crystallography has the advantage over NMR techniques for the definition of ligand-binding sites.[10] However, the detection of ligands in protein sites is highly dependent on the resolution limit of diffraction data.[11] For a typical protein structure with resolution worse than 2 Å, it may be very difficult to unambiguously recognize the ligand from inspection of the electron-density map of the catalytic site. To overcome this difficulty, a method has been developed recently that uses the soaking of a protein crystal in a cocktail of molecules with different shapes that can be easily distinguished in the electron density map.[9] In recent years, various technical improvements, ranging from better crystallization techniques to the use of synchrotron sources and cryogenic techniques for the measurement of diffraction data, have made it possible to improve the resolution limit[12] considerably, and biocrystallographic studies with resolution values between 1.5 and 0.9 Å are becoming more frequent.[13] At this level of resolution, the individual atoms can be clearly distinguished. This leads to the precise identification of the stereochemistry of a bound inhibitor and often provides important clues for the rationalization of its affinity and selectivity.[13] In particular, the possibility to distinguish the stereoisomer recognized by an enzyme is crucial for the development of chiral drugs. Currently, racemates and diastereomeric mixtures are almost absent in the list of FDA-approved drugs.[14]

All potent and selective inhibitors of HIV protease are single enantiomers and were designed as mimics of the viral substrates with non-cleavable structures replacing the scissile bond.[15] The HIV-1 protease specifically processes the gag and gag–pol polyproteins into mature viral replication enzymes (reverse transcriptase, integrase, and protease) and structural proteins (p6, p7, p17, and p24).[16] The HIV-1 protease, unlike other proteases, is able to cleave Tyr–Pro and Phe–Pro sequences in the viral polyprotein. As the amide bonds of Pro residues are not susceptible to easy cleavage by mammalian endopeptidases, peptidomimetic molecules that incorporate non-hydrolysable Phe–Pro isosteres are expected to have the potential advantage of higher selectivity.[17] Recently, we described a new synthetic approach to dihydroxyethylene Xaa–Pro isosteres[18] (in which Xaa = variable residue) and applied it to the synthesis of a new generation of protease inhibitors based on a novel Phe–Pro isostere (Figure 1).[19] The presence of two amino termini in this non-hydrolysable moiety permits the development of pseudosymmetric peptidomimetic inhibitors that can match the twofold symmetry of the homodimeric structure of the HIV-1 protease (Figure 2). The pseudo-hexapeptide TS-126, which has an acetylated Trp–Val dipeptide at both termini of the central unit (Figure 1), has IC_{50} values in the nanomolar range, similar to those of commercial drugs. However, during the synthesis of this new inhibitor, an epimerization of the two C^{α} (Val) stereocenters occurred and a mixture of four stereoisomers was obtained. To determine which stereoisomer is associated with the strong enzyme inhibition, biocrystallographic techniques were used to screen the epimeric mixture. The

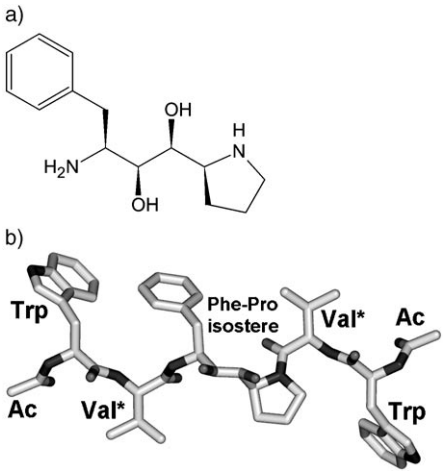

Figure 1. a) Structure of the Phe–Pro isostere; b) stick representation of the L-Val/L-Val epimer of the TS-126 inhibitor.

[a] *Prof. S. Geremia, N. Demitri, Dr. J. Wuerges, Prof. F. Benedetti, Dr. F. Berti, Prof. L. Randaccio*
Department of Chemical Sciences and
Centre of Excellence in Biocrystallography
University of Trieste
Via L. Giorgieri 1, 34127 Trieste (Italy)
Fax: +39 040 558 3903
E-mail: geremia@univ.trieste.it

[b] *Prof. G. Tell*
Department of Biomedical Sciences and Technologies and
Centre of Excellence in Biocrystallography
University of Udine
Piazzale Kolbe 4, 33100 Udine (Italy)

[**] *This work was supported by the MIUR (FIRB RBNE01TTJW).*

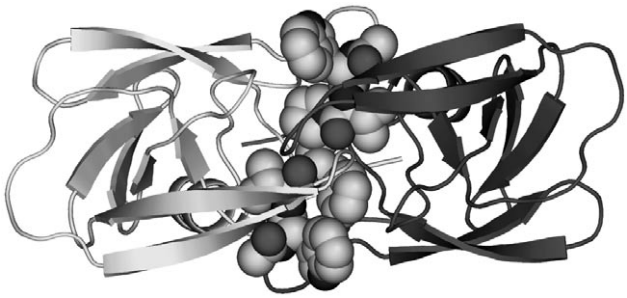

Figure 2. The dimeric structure of HIV-1 protease (ribbon/loop rendering) is viewed parallel with the twofold symmetry axis. The TS-126 inhibitor (space filling representation) inside the catalytic channel fits the pseudosymmetry of the protein structure well.

enzyme was crystallized in the presence of a mixture of the four stereoisomers. Fourier maps obtained by high-resolution diffraction data (up to 1.3 Å) from synchrotron radiation clearly show that the catalytic site is fully occupied by a single ordered isomer (Figure 3). This permitted the unambiguous identification of the natural L configuration of both Val residues of the bound inhibitor. At first glance, this could have been taken for granted. However, it is important to note that the inhibitor has two N termini, and therefore, its stereochemistry is quite different from that of a natural peptide. The electron-density map, without electron-density residuals, also suggests that the inhibition constant of this isomer should be at least one order of magnitude lower than those of the other configurational stereoisomers (assuming an equimolar mixture and the ability to detect an occupancy as low as 10% in the Fourier maps). Furthermore, the protein and inhibitor concentrations used in the crystallization drop, and the full occupancy of the site (or at least an occupancy $>90\%$) indicate that the inhibition constant is less than 1 μM.

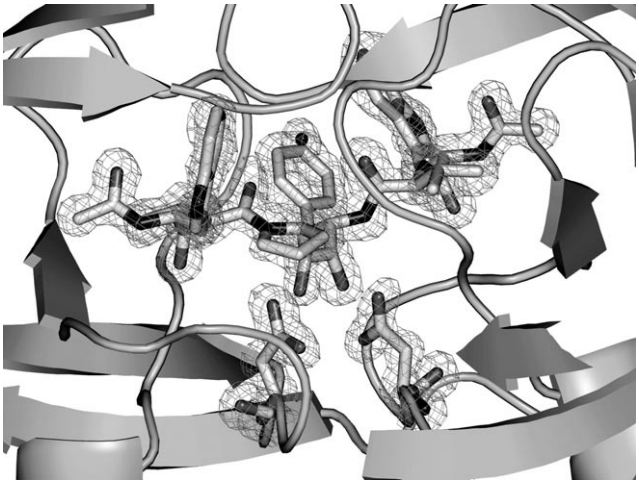

Figure 3. Electron-density map ($2F_o-F_c$, contour level 2σ) near the catalytic site of HIV-1 protease (ribbon/loop rendering). The TS-126 inhibitor and catalytic aspartate groups are shown in stick representation.

This crystallographic study allowed us to identify the stereochemistry of the inhibitor and to assess the inhibition properties of this compound without purification of the mixture and the classic activity assays that are normally carried out on each compound. The co-crystallization strategy can be applied in conjunction with synchrotron radiation and cryogenic techniques to obtain high-resolution structures that permit the discovery, by self selection, of the most potent inhibitor in a mixture of stereoisomers. Notably, atomic resolution is essential to assign the amino acid configuration and to determine the relative population of the stereoisomers in the catalytic site quantitatively. In fact, the undetectable H atoms in the electron-density map can permit inversion of the configuration of the model through a minimal shift (≈ 1 Å) of the C^α atom (umbrella mode of inversion of the stereogenic center). In this respect the co-crystallization method has an advantage over soaking techniques in that it avoids crystal stress, which decreases the resolution limit of diffraction data.

In conclusion, we have demonstrated that high-resolution X-ray crystallography can be used as a powerful method to identify the best HIV protease inhibitor present in an epimeric mixture. This study represents the first example of a successful application of protein crystallography for screening a mixture of stereoisomers.

Experimental Section

Expression and purification of HIV-1 protease: HIV-1 protease (Genbank HIVHXB2CG), stabilized by five mutations (Q7K, L33I and L63I to minimize auto-proteolysis, and C67A and C95 to prevent aggregation by the formation of disulfide bonds) was expressed in *Escherichia coli* and purified from inclusion bodies.[20]

Synthesis of the inhibitor: The diaminodiol Phe–Pro core unit of TS-126 was synthesized in enantiomerically pure form according to refs. [18] and [19]. The core unit was then coupled with 2 equiv Ac-NH-Trp-Val-OH in the presence of 1-benzotriazolyloxy-tris(pyrollidino)phosphonium hexafluorophosphate (PyBOP) and 1-hydroxy-1*H*-benzotriazole (HOBT) as activating agents. Coupling resulted in the formation of four stereoisomers, presumably through the well-known peptide oxazolone racemization[21] at C^α (Val) stereocenters.

Enzyme assay: Inhibition of HIV-1 protease activity was evaluated as previously reported.[22] In particular, assays of HIV protease activity were performed with the fluorogenic substrate Abz-Thr-Ile-Nle-(*p*-nitroPhe)-Gln-Arg-NH_2 (Bachem, Budendorf, Switzerland). Titration of protease was carried out with the reference inhibitor ritonavir.

Crystallographic analysis: The stabilized HIV-1 protease was crystallized in the presence of an epimeric mixture of TS-126. Crystals were grown at room temperature by vapor diffusion using the hanging drop method. The reservoir contained 0.25 M citrate buffer (pH 6), 10% DMSO, and 25–50% saturated ammonium sulfate as precipitant. The protein (2–5 mg mL^{-1}) was preincubated with a 5-fold molar excess of the epimeric mixture of the TS-126 inhibitor. The crystallization drops were formed using 1 μL of reservoir solution and 1 μL of a solution of protein with inhibitors. Crystals of typical dimension $0.5\times0.3\times0.1$ mm^3 grew in 2–7 days. Preliminary X-ray diffraction data at 1.8 Å resolution were collected on a rotating anode source with a KappaCCD detector. Final data were collected on diffraction beam-line at the Elettra Synchrotron,

Table 1. X-ray data collection and refinement statistics.

unit cell	$a = 57.27$ $b = 84.96$ $c = 46.06$
space group	$P2_12_12$
resolution range [Å]	10–1.30
total reflections	288 250
unique reflections	55 106
completeness [%]	98.6 (98.3)[a]
I/σ (I)	10.0 (2.5)[a]
multiplicity	2.7 (2.7)[a]
R_{merge} (I) [%]	6.1 (40.8)[a]
refinement:	
R_{factor} [%]	13.48
R_{free} [%]	17.77
protein atoms	1512
inhibitor atoms	66
other atoms[b]	34
water molecules	272
B factor	
average	19.49
protein main chain	13.86
protein side chains	19.69
inhibitor main chain	15.30
inhibitor side chains	24.30
other molecules[c]	30.24
water molecules	33.07
RMS deviation	
bond lengths [Å]	0.013
bond angles [Å]	0.033

[a] Value in parentheses: highest resolution shell. [b] 4 DMSO, 1 glycerol, 1.5 CH_3COO^-, 1 Na^+, and 1 Cl^-. [c] DMSO, glycerol, CH_3COO^-, Na^+, and Cl^- ions.

Trieste Italy (Table 1). Data were collected on MAR-CCD detector from 100 K frozen crystal with 20% glycerol as cryoprotecting agent. Data were processed using the Mosflm suite. Protease coordinates (PDB code: 1S6G) were used as the starting model for molecular replacement with AmoRe. The structure was refined by using SHELXL and refitted with the program O. Alternative conformations for residues were modeled where appropriate. The coordinates for the structure have been submitted to the Protein Data Bank (code: 2A1E).

Keywords: configuration determination · HIV protease · peptidomimetics · protein structures · X-ray diffraction

[1] T. L. Blundell, *Nature* **1996**, *384*, 23 – 26.
[2] R. S. Rowland, *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 613 – 619.
[3] S. W. Homans, *Angew. Chem.* **2004**, *43*, 290 – 300; *Angew. Chem. Int. Ed.* **2004**, *43*, 290 – 300.
[4] W. L. Jorgensen, *Science* **2004**, *303*, 1813 – 1818.
[5] J. Greer, J. W. Erickson, J. J. Baldwin, M. D. Varney, *J. Med. Chem.* **1994**, *37*, 1035 – 1054.
[6] C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855 – 861.
[7] M. S. Congreve, D. J. Davis, L. Devine, C. Granata, M. O'Reilly, P. G. Wyatt, H. Jhoti, *Angew. Chem.* **2003**, *42*, 4479 – 4482; *Angew. Chem. Int. Ed.* **2003**, *42*, 4479 – 4482.
[8] S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, *274*, 1531 – 1534.
[9] V. L. Nienaber, P. L. Richardson, V. Klighofer, J. J. Bouska, V. L. Giranda, J. Greer, *Nat. Biotechnol.* **2000**, *18*, 1105 – 1108.
[10] T. L. Blundell, H. Jhoti, C. Abell, *Nat. Rev. Drug Discovery* **2002**, *1*, 45 – 54.
[11] A. M. Davis, S. J. Teague, G. J. Kleywegt, *Angew. Chem.* **2003**, *42*, 2718 – 2736; *Angew. Chem. Int. Ed.* **2003**, *42*, 2718 – 2736.
[12] A. Sharff, H. Jhoti, *Curr. Opin. Chem. Biol.* **2003**, *7*, 340 – 345.
[13] L. Coates, D. A. Myles, *Curr. Drug Targets* **2004**, *5*, 173 – 178.
[14] H. Caner, E. Groner, L. Levy, I. Agranat, *Drug Discovery Today* **2004**, *9*, 105 – 110.
[15] D. D. Richman, *Nature* **2001**, *410*, 995 – 1001.
[16] M. Hill, G. Tachedjian, J. Mak, *Curr. HIV Res.* **2005**, *3*, 73 – 85.
[17] A. Wlodawer, J. Vondrasek, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249 – 284.
[18] F. Benedetti, F. Berti, F. Dinon, G. Nardin, S. Norbedo, *Org. Lett.* **2004**, *6*, 1017 – 1019.
[19] F. Benedetti, F. Berti, P. Campaner, F. Dinon, A. Tossi, PD2004A000151, **2004**.
[20] J. M. Louis, G. M. Clore, A. M. Gronenborn, *Nat. Struct. Biol.* **1999**, *6*, 868 – 875.
[21] M. Goodman, W. J. McGahren, *Tetrahedron* **1967**, *23*, 2031 – 2050.
[22] A. Tossi, F. Benedetti, S. Norbedo, D. Skrbec, F. Berti, D. Romeo, *Bioorg. Med. Chem.* **2003**, *11*, 4719 – 4727.

Received: October 13, 2005
Published online on December 13, 2005

