Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 4, No. 4, pp. 583-588, 1994
Elsevier Science Ltd
Printed in Great Britain
0960-894X/94 $6.00+0.00

0960-894X(94)E0013-5

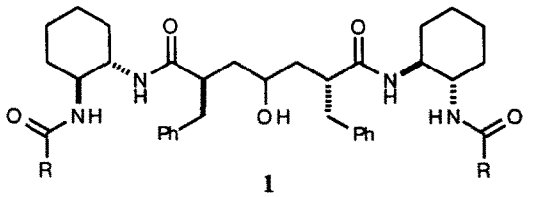
Design, Structure Activity and X-Ray Crystallographic Studies of Pseudo-symmetrical Nonpeptidyl HIV-1 Protease Inhibitors

Robert E. Babine *§, Nan Zhang *, Steven R. Schow, Zhangbao Xu, Randal A. Byrn †, Richard C. Hastings †, M. F. Semmelhack , Michael M. Wick and Suresh S. Kerwar

Exploratory Medicinal Chemistry Section
American Cyanamid Company
Medical Research Division
Lederle Laboratories
Pearl River, NY 10965

Abstract: A new class of nonamino acid derived HIV-1 protease inhibitors of structure **1** are described. Structure activity relationships are discussed in the context of a protein crystal complex.

We have previously reported[1] on the structure-assisted design, synthesis and enzyme inhibiting properties of a pseudo-symmetrical,[2] dipeptide derived HIV-1 protease inhibitor **2**. We now wish to report on the extension of this work to the design and biological evaluation of a novel end group. These studies have resulted in a new class of potent non-amino acid derived HIV-1 protease inhibitors of the generic structure **1**. The X-ray crystal structure of the complex between HIV-1 protease and **3**, the most potent member of this series, is also reported.

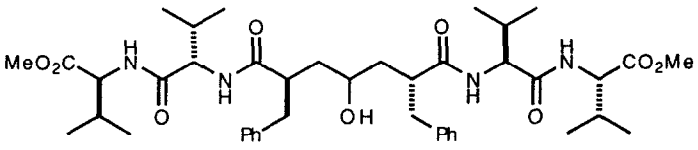


1

Our tactic was to divide the design problem into two separate stages, the design of a central unit and the design of two identical terminal groups. The first round of design identified the central unit.[1b] The second round took the peptidyl lead compound **2** as the starting point. These modeling studies identified a cyclohexyl group, substituted with an amide group, as a replacement for the Val-Val dipeptide. The role of the amide was two fold. First, modeling showed that each of the newly introduced amide groups could form two hydrogen bonds with the enzyme, thus orienting the cyclohexane rings in place. Second, it provided a handle to introduce different substituents into the S_3 sites. Thus, we had incrementally designed a new class of molecules with generic structure **1**.

---

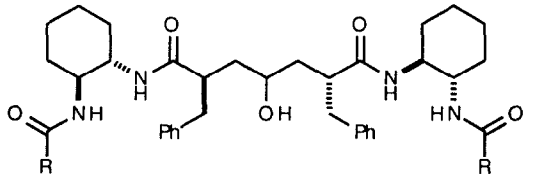
§ Agouron Pharmaceuticals Inc. 3565 General Atomics Ct. San Diego, CA 92121.



2

Table 1 lists IC_{50} values of various analogs as HIV-1 protease inhibitors[3] and as antiviral agents.[4] Of different derivatives surveyed, the most potent inhibitor contained R as a 2-quinolyl group.[5] The three pyridyl derivatives **4** -**6** were less potent by about an order of magnitude, and compound **7** having a methyl group was found to be less potent than the pyridyl derivatives by yet another order of magnitude. These compounds were prepared using methodology we have previously reported.[1b] The key step involved bidirectional amide coupling of a C_2 symmetric diacid[1b] with a large excess of the C_2 symmetric (1S, 2S)-(+)-1,2-diaminocyclohexane.

Table 1. Biological properties of various non-peptidyl HIV protease inhibitors.



Compound	R	Enzyme IC_{50}(nM)	Antiviral IC_{50}(μM)
3	2-Quinolyl	40	2
4	2-Pyridyl	400	>20
5	3-Pyridyl	600	
6	4-Pyridyl	350	>20
7	Methyl	3,000	
8	*t*-Butoxy	>10,000	
9	3-IsoQuinolyl	45	

In an iterative approach to validate our modeling, improve our understanding of molecular recognition of HIV-1 protease and further design new HIV-1 protease inhibitors, we co-crystallized this compound with HIV-1 protease enzyme. The structure of the enzyme-inhibitor complex of **3** was elucidated by an X-ray crystallography study.[6,7]

Figure 1 shows the structure of inhibitor **3** in the HIV-1 protease complex. The inhibitor is bound to the protein in a nearly symmetric fashion. The major deviation from symmetry occurs in the central three carbons of the molecule. The central hydroxyl group was determined to lie between the two catalytic aspartates (D-25 and D-25').[8] The two benzyl groups were found to be in the two equivalent S_1 subsites and the two cyclohexyl groups were found to be in the two equivalent S_2 subsites. The two carbonyl groups which link the central unit to the cyclohexyl groups were found to form hydrogen bonds to the central structural water molecule.[9] It appears likely that the inhibitor lies in two C_2-symmetric orientations as is seen in many crystal structures solved in hexagonal

space groups.[6a] In addition the amide groups linking the two cyclohexyl rings to the quinoline groups also form hydrogen bonds to the enzyme, albeit with less than optimal geometry.[10]

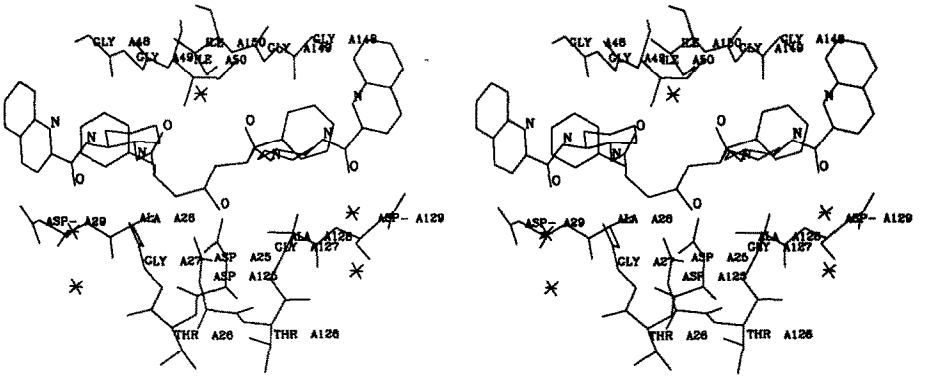


Figure 1: "Relaxed" stereo view of inhibitor **3** bound to HIV-1 protease.

In the crystal structure the nitrogen atom of the quinoline ring lies anti to the carbonyl oxygen of the linking amide group. As a result of this orientation each quinoline group makes an edge to face interaction [11] with its neighboring phenyl group in the adjacent S_1 subsite. In addition, this orientation allows for close contacts between the ring nitrogen atoms and the carbonyl oxygens of Gly-48 and Gly-48' of the enzyme.[12] These are presumably unfavorable electrostatic interactions and no evidence of bridging water molecules is evident in the electron density maps.

Compound **9**, the 3-isoquinolyl analog of **3** was prepared after the crystal structure was solved. The idea was to retain the orientation of the aromatic ring observed in the crystal structure while simultaneously removing the unfavorable electrostatic interaction between the ring nitrogens and Gly-48 and 48' and introducing a favorable electrostatic interaction between the nitrogen atoms and the guanidine groups of Arg-8 and 8'. These two compounds were found to be equally potent as enzyme inhibitors. We rationalized this result as follows. For both compounds **3** and **9**, there are two minimum energy conformations available for the carbonyl group adjacent to the fused aromatic rings. The favored conformation is the one in which the electronegative nitrogen atom in the aromatic ring lies anti to the electronegative carbonyl oxygen atom, diminishing an unfavorable electronic interaction. The disfavored conformation is the one in which these two electronegative atoms lie syn to each other. For compound **3**, the bound conformation is the favored conformation for the free ligand. Thus, the free ligand is preorganized for binding to the enzyme in its low energy local conformation. For compound **9**, the presumed bound conformation (isosteric with **3**) is the disfavored conformation. Therefore, energy has to be spent to orient it into its bound conformation. Even though when compared with compound **3**, compound **9** may have higher affinity toward the enzyme in its bound conformation due to the electrostatic reasons discussed above, part of the binding energy has to be spent to compensate organization of the ligand into a disfavored conformation. As a result, the overall binding energy for the two compounds are comparable.

One experimental observation that can be addressed with this crystal structure is that potency as enzyme inhibitors is related to the size of the groups in the P_3 positions (2-quinolyl > 2, 3, or 4-pyridyl > methyl). The complex of **3** buries a significant amount of surface area for both the protein (330 Å^2 hydrophobic surface area and 184 Å^2 hydrophilic surface area) and the ligand (688 Å^2 hydrophobic surface area and 162 Å^2 hydrophilic surface area). The observed potency trend appears to be related to the amount of buried hydrophobic surface area.[13,14] In addition, both quinolyl groups make edge to face interactions with the P_1 phenyl groups.

One powerful use of crystal structure data is the ability to compare related complexes to gain insight into the molecular recognition of protein-ligand complexes. L-700,417 (**10**) developed by Vacca *et al* [15] is structurally related to **3** since it contains the same 4-hydroxy-2,6-bis-(phenylmethyl)heptanediamide central unit. The crystal structure of the complex between **10** and HIV-1 protease has previously been reported[15] and the coordinates deposited with the Brookhaven Protein Data Bank (entry 4PHV). These two complexes were superimposed and the ligands directly compared. (See Figure 2) There are some similarities in the two structures as well as some striking differences. While both compounds have an identical heptanediamide central units they differ slightly in the bound conformation of this group. The phenyl P_1 groups, the central hydroxy and center carbonyl oxygens all overlap each other despite this difference. The major consequence of this different conformation is that the cyclohexyl groups in **3** and the phenyl groups of the indanes of **10** overlap. That these two groups overlap strongly suggests that burying complementary hydrophobic surfaces is the major driving force in complexation. In addition, the terminal carbonyl oxygens of **3** and the hydroxyl groups of indane **10** also overlap. There are no major differences in the structure of the protein in either complex.

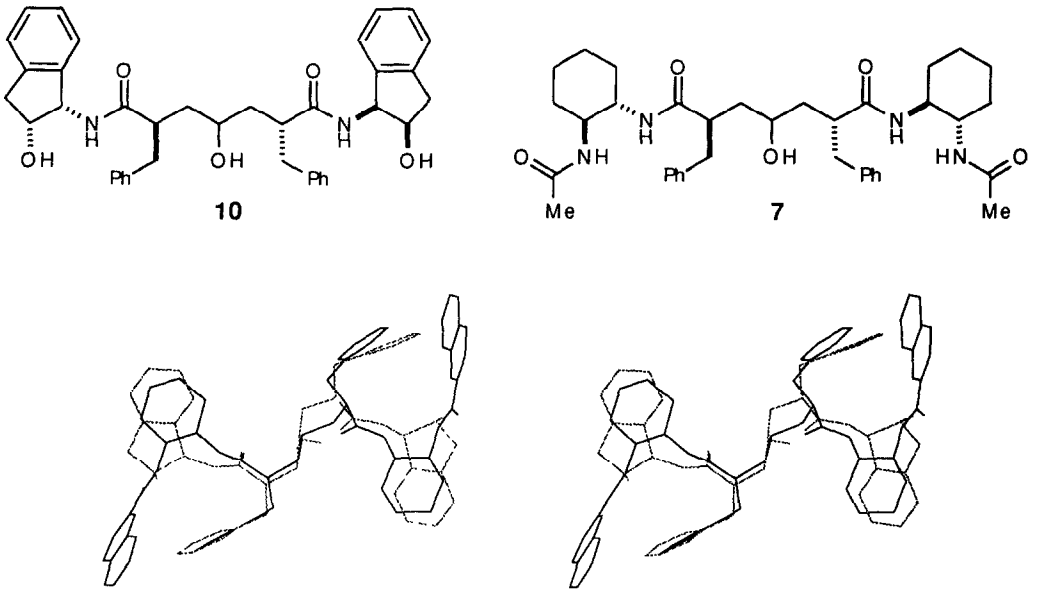


Figure 2: "Relaxed" stereo view of inhibitor **3**(dark lines) superimposed on inhibitor **10** (light lines).

One dilemma that exists is that compounds related to **3** require large hydrophobic P_3 groups to be potent enzyme inhibitors while **10**, which completely lacks a P_3 group, is a very potent enzyme inhibitor. To vividly illustrate this point, although never tested head to head in the same assay, compound **10** is approximately one thousand times more potent than **7**. At least three possibilities exist to explain these differences. (1) The different conformations of the linking heptanediamide groups. The conformation found for **3** does not appear to be strained or to make a bad steric clash with the protein. (2) Differences in desolvation energies of **10** and **7**. MacroModel calculations suggest they have similar solvation energies.[16] (3) The hydroxyl groups in the indane portions of **10** have a more favorable electrostatic interaction with the protein than do the terminal amide groups of **3**. Delphi calculations [17] were performed to access the electrostatic contributions, of the indane hydroxyl groups of **10** and the terminal amide groups of **3**, to binding. These calculations indicate that indane hydroxyl groups contributes 5.4 Kcal more to the binding affinity of **10** than the terminal amide groups contribute to the binding of **3**. We thus suggest that the large potency difference between **10** and **7** is due to less than optimal hydrogen bonds between the amide groups of **7** and the protein and to highly favorable hydrogen bonds between the indane hydroxyl groups of **10** the protein and its structural waters. [15]

Analysis of the crystal structure of **3** has allowed us to address several questions regarding the structure activity relationships of this series and a related series of compounds. This study has allowed us to gain some new insight into the molecular recognition between HIV-1 protease and inhibitors.

† Department of Medicine, Hematology/Oncology Research Laboratory, New England Deaconess Hospital, Boston, MA 02215

Acknowledgment: We wish to thank Drs. S. Plotch and E. Baum for providing the protein needed for this study.

References and Notes

1. (a) Presented in part by R. E. Babine at "Thirty-Second Annual Medicinal Chemistry Symposium", Buffalo, NY, June 3-5, 1991. (b) Babine, R. E.; Zhang, N.; Jurgens, A. R.; Schow, S. R.; Desai, P. R.; James, J. C.; Semmelhack, M. F. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 541-6. (c) Babine, R. E.; Zhang, N.; Schow, S. R.; Jirousek, M. R.; Johnson, B. D.; Kerwar, S. S.; Desai, P. R.; Byrn, R. A.; Hastings, R. C.; Wick, M. M. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1589-94 (d) Trova, M. P.; Babine, R. E.; Byrn, R. A.; Casscles, W. T.; Hsu, G. C.; Jirousek, M. R.; Johnson, B. D.; Kerwar, S. S.; Schow, S. R.; Wissner, A.; Zhang, N. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1595-1600.

2. For a recent review on symmetry-based HIV-1 protease inhibitors see: Erickson, J. W. *Perspectives in Drug Discovery and Design* **1993**, *1*, 109-128.

3. We used the following enzyme assay in this work: The enzyme inhibition studies were performed using recombinant HIV-1 protease prepared and purified by Drs. S. Plotch and E. Baum. The enzyme assay was an ^{125}I-SPA (scintillation proximity assay) using a kit provided by Amersham International plc, Bucks, England. Using the procedure described by the manufacturer, enzyme and test compounds were incubated with the bead suspension for 5 min., following termination of the reaction the radioactivity of the assay mixture was determined. Enzyme inhibition was determined as a function of inhibitor concentration. As a literature control Ro 31-8959 [5] was determined to have an IC_{50} = 30 nM. *Note:* This assay was incorrectly described in Ref 1c.

4. The anti-viral assay used in this work has previously been described. See Ref 1c.

5. Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; Lambert, R. W.; Merrett, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Ritchie, A. J.; Taylor, D. L.; Thomas, G. L.; Machin, P. J. *Science* **1990**, *248*, 358-61

6. For recent reviews on HIV-1 protease crystallography see:
(a) Appelt, K. *Perspectives in Drug Discovery and Design* **1993**, *1*, 23-48.
(b) Wlodawer, A. and Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543-85.

7. Cocrystals between **3** and HIVP were obtained by the hanging drop diffusion technique. The precipitant for crystallization consisted of satd. ammonium sulfate buffered with 0.2 M sodium phosphate, 0.1 M citric acid at pH 6.77. A 2.2 Å resolution X-ray diffraction data set (8370 reflections with I> 1.5 ∂) were collected by using the RAXII Image plate area detector at McMaster University. The merging R-factor was 6.9%. The structure of HIV-1 protease and compound **3** complex was solved by applying the Molecular Replacement method with the program MERLOT.(Fitzgerald, P. *J. Appl. Chryst.* **1988**, *21*, 273-8.) The crystal belonged to the hexagonal, Space Group P61, with cell constants of a=b=63.079 Å, c=83.004 Å, α=ß=90° and γ=120°. The HIV-1 protease dimer structure was refined by using the molecular dynamics program XPLOR. (Brunger, A. T. *J. Mol. Biol.* **1988**, *203*, 803.)The crystallographic R-factor went down from 42.0% to 29.8%. The inhibitor was located in the centric cavity of the HIV-1 protease dimer. The inhibitor fitted the electron density very well. The complex structure with 1516 non-hydrogen protein atoms and 64 non-hydrogen inhibitor atoms was refined by XPLOR again at 10.0 - 2.2 resolution data. The final refinement with XPLOR took place after the addition of five structural water molecules to give a final R-factor of 19%.

8. Distances between central OH group of inhibitor and oxygens of D-25 (2.96 and 3.18 Å) and D-25' (3.10 and 2.78 Å).

9. Distances between amide groups of heptanediamide central unit and central structural water and carbonyl oxygens of G-27 and G-27' (N-O G-27 = 3.32 Å, O-WAT = 2.78 Å, N-O G-27' = 3.38 Å, O-WAT = 2.85 Å)

10. Distances between nitrogen terminal amide groups of inhibitor and carbonyl oxygen of G-48 (3.60 Å) and G-48'(3.41 Å). Distances between carbonyl oxygen of terminal amide groups of inhibitor and backbone nitrogen of D-29 (2.82 Å) and D-29'(3.29 Å).

11. Burley, S. K.; and Petsko, G. A. *Adv. Prot. Chem.* **1988**, *39*, 125-189

12. Distances between quinoline nitrogens and carbonyl oxygens of G-48 (4.25 Å) and G-48'(3.68 Å).

13. Buried solvent accessible hydrophobic surface areas: 2-Quinolyl 1018 $Å^2$, 2-Pyridyl 908 $Å^2$, Methyl 823 $Å^2$. Buried surface areas were calculated using the solvent surface routine in Quanta 3.3. This program has been developed by Molecular Simulations Inc.

14. The observed potencies and calculated buried solvent accessible hydrophobic surface areas correlate very well based upon a 12 cal/$Å^2$ energetic function. Eisenberg, D. and McLachlan, A.D. *Nature* **1986**, *319*, 199-203

15. Bone, R.; Vacca, J. P.; Anderson, P. S.;and Holloway, M. K. *J. Am. Chem. Soc* **1991**, *113*, 9382-4.

16. Solvation Energies were calculated with MacroModel V4.0 (Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Caufield, C.; Chang, G.; Hendrickson, T.; and Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.)using the OPLS* forcefield and the GBSA solvation model (Still, W. C.; Tempczyk, A.; Hawley, R.C.; and Hendrickson,T. *J. Am. Chem. Soc.* **1990**, *112*, 6127-9). Solvation energies were taken from the average solvation energy for the small molecule during a 100 Psec molecular dynamics run. The calculated solvation energies were as follows: **10** -19.00 Kcal/mole; **7** -19.63 Kcal/mole

17. Two sets of Delphi (Gilson, M.; Sharp, K. and Honig, B. *J. Comput. Chem.* **1988**, *9*, 327) calculations were performed for each complex (**3** and **10**). One set used the complex (including 5 structural water molecules) with partial charges derived from the Discover CFF91 force field (normal charges). Hydrogen atom positions were then minimized to convergence with Discover. The second set of calculations used the same structures and charges except that for **3** the partial charges for two terminal CONHCH groups were set to 0.0 and for **10** the partial charges for two terminal HOCH groups were set to 0.0 (mutated charges) . The calculations were carried out for the complexes both in water (dielectric = 80.0 ; protein interior = 2.0) and in vacuum (dielectric = 1.0 ; protein interior = 2.0). Calculations for the free inhibitors were carried out only in water to obtain the coulombic term. The solvation energies of the complexes were calculated by subtraction of the reaction field energies (solvent - vacuum) and the coulombic terms were also calculated by subtraction (complex - inhibitor). The total electrostatic component was the coulombic term plus the solvation term. The effect of the individual group was obtained by subtraction of the mutated charge calculation from the normal charge calculation. Discover, CFF91 and Delphi are distributed by Biosym Technologies of San Diego,CA.

(*Received in USA* 23 *November* 1993; *accepted* 22 *December* 1993)