Bioorganic & Medicinal Chemistry Letters, Vol.2, No.6, pp. 541-546, 1992
Printed in Great Britain

0960-894X/92 $5.00 + .00

The Use of HIV-1 Protease Structure in Inhibitor Design[1]

Robert E. Babine*, Nan Zhang, Alex R. Jurgens, Steven R. Schow, Parimal R. Desai, John C. James and M. F. Semmelhack

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

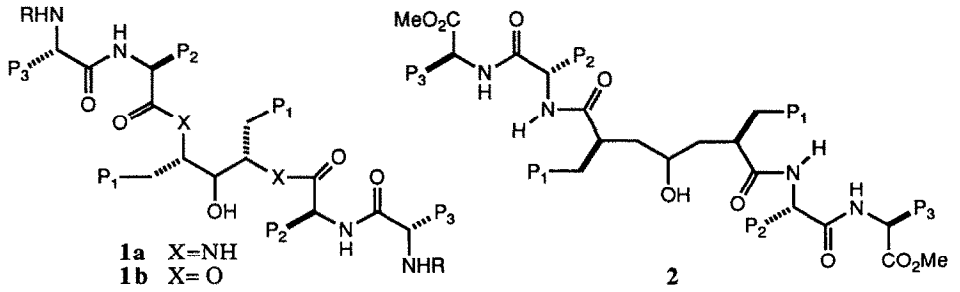
(*Received* 4 *March* 1992)

Abstract: The X-ray structure of HIV-1 protease and molecular dynamics studies were used in the design of pseudo-symmetrical enzyme inhibitors **1** and **2**. Short bidirectional syntheses of these compounds are described.

HIV-1 protease is an essential retroviral enzyme which is responsible for the processing of viral polyproteins into viral structural proteins and enzymes.[2] Structural studies using X-ray crystallography have shown that the protein is a C_2 symmetrical homodimer, and a member of the well known aspartic protease class of hydrolytic enzymes [3,4]. As soon as the enzyme structure was determined it was recognized that C_2 symmetrical molecules had potential as inhibitors of this enzyme.[5] We wish to report the use of mechanistic and structural information known about HIV-1 protease in the "structure assisted" design of potent inhibitors of this enzyme.

C_2 Symmetrical Aspartic Protease

"Structure Assisted" Design



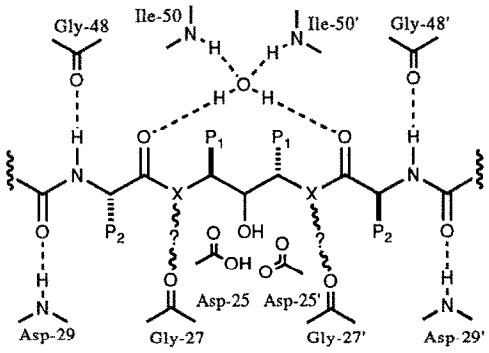
Pseudo C_2 Symmetrical Inhibitors

In contrast to most protease enzymes which have evolved separate C-terminal and N-terminal substrate binding domains, the retroviral aspartic proteases as a result of their symmetry have structurally identical C-terminal and N-terminal binding domains. The enzyme's C_2 axis lies between and perpendicular to the two catalytic aspartate residues in the active site which activate an enzyme bound water molecule. It is well know that the secondary alcohol group of statine and related protease inhibitors can displace this catalytic water molecule of aspartic proteases in an enthalpically neutral fashion, resulting in enhanced binding. This effect is proposed to be due to entropic effects.[6] Upon visual examination of the X-ray structure of HIV-1 protease it was clear that the so

called S_1 and S_1' subsites [7] are related by a C_2 symmetry operation and are thus structurally identical (or nearly identical). With regard to symmetrical retroviral proteases we prefer the term "two equivalent S_1 (S_2,S_3,etc) subsites" to S_1 and S_1' (S_2 and S_2' ,etc) subsites. This structural property of the enzyme dictates our design strategy of C_2 symmetrical or psuedosymmetrical inhibitors. It was also clear from molecular graphics examination that the S_1 and S_2 subsites are quite hydrophobic in nature. 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. These three units would be coupled together by amide linkages. A key requirement of the central unit is that its symmetry properties be aligned with the symmetry properties of the enzyme. The proposed role of the central unit is to displace the catalytic water molecule with a hydroxyl group, introduce two hydrophobic P_1 groups into the two equivalent hydrophobic S_1 subsites, and form bridging hydrogen bonds to the enzyme bound structural water molecule [8]. The initial compounds prepared use P_2 and P_3 groups derivable from simple dipeptides as the terminal groups. The choice of dipeptides as initial terminal groups was primarily due to their commercial availability.

Using this strategy and a dynamical model we have formulated several potential targets for synthesis. Molecular dynamics simulations [9,10] and retrosynthetic considerations were used in prioritizing targets for execution of synthetic studies. We used the complex between HIV-1 protease and MVT-101 (a reduced amide inhibitor) [4a] as the starting enzyme structure. The average structure during the dynamics run as well as the dynamics trajectory were then critically evaluated. A potential target for synthesis was one in which major conformational changes in the enzyme (relative to the starting X-ray structure) did not take place, the potential inhibitor did not undergo major conformational changes during the dynamics run, and its average structure essentially maintained C_2 symmetry. Structures **1** and **2** represent two classes of compounds resulting from this analysis. We now will discuss the use of dipeptides as probe reagents to evaluate the central unit and generate new leads for anti HIV-1 agents.

Figure 1

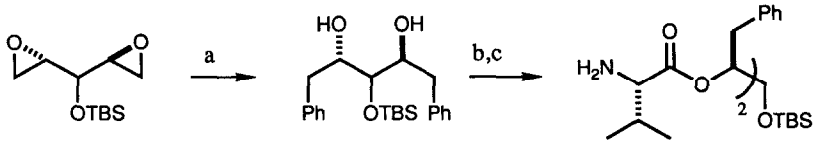


Diamide **1a** was considered one of our prime targets; however synthetic considerations led us to investigate diester **1b** first. Consideration of the potential hydrogen bonding arrangement between the inhibitors and the enzyme revealed that when X=NH there is potential for hydrogen bonding of the amide NH groups to the

amide carbonyl group of Gly-27 and the symmetry related Gly-27' (See Figure 1). We were aware of the work of Bartlett [11], Kollman [12] and Matthews [13] with thermolysin inhibitors which demonstrated that substitution of a hydrogen bond with a repulsive electrostatic interaction may have dramatic effects on the strength of a protein ligand complex. In the thermolysin case a "forced repulsion"[12] is created due to conformational rigidity of the enzyme. In the case of HIV-1 protease, conformational flexibility has been observed for Gly-27 and Gly-27' since X-ray studies have shown that upon binding inhibitors the plane of the amide between Gly-27 and Ala-28 (also the symmetry related Gly-27' and Ala-28') rotates approximately 90° to allow formation of hydrogen bonds with the inhibitor.[4a] Molecular dynamics simulations for a complex of diester **1b** and HIV-1 protease suggest that the conformation of these amide groups resemble the conformation found in the free enzyme and that forced repulsion resulting from the ester oxygen would not be a severe problem.

The bidirectional synthesis of **1b** (P_1=Ph,P_2=P_3=*i*-Pr,R=Boc) starting from the bis-epoxide **3**[14] is outlined in Scheme 1. Double cuprate opening of **3** gave **4**. Double ester coupling using FMOC-aminoacid chlorides [15] followed by deprotection then gave **5**. BOP mediated peptide coupling [16], followed by removal of the silyl protecting group afforded compound **1b**. This compound proved to be a potent inhibitor of HIV-1 protease with an IC_{50}= 5 nM.

Scheme 1

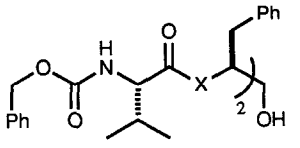


3 **4** **5**

Reagents:
a. $Ph_2Cu(CN)Li_2$, THF, 75% b. FMOC-Val-Cl, DMAP, pyridine, CH_2Cl_2, 75% c. 4-aminomethylpiperidine, $CHCl_3$
d. Boc-Val-OH, BOP, Et_3N, HOBT, CH_2Cl_2, 46%(2 steps)
e. nBu$_4$NF, THF, 91%

1b P_1=Ph,P_2=P_3=*i*-Pr, R=*t*-Boc

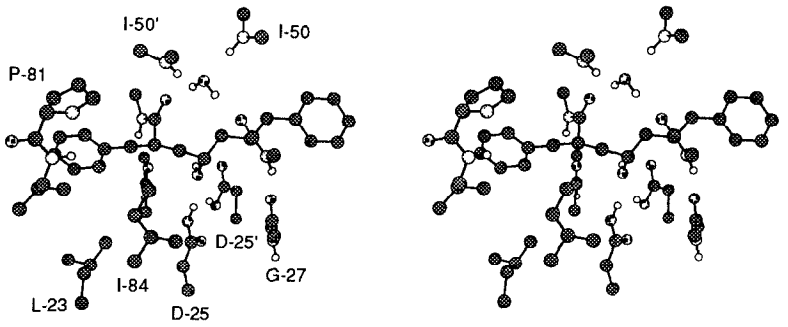
The reports from Erickson [5b] and Kempf [5c] related to compounds **1a**, including **7**, prompted us to compare the enzyme inhibiting properties of diester **6** directly with diamide **7**. Diamide **7** was determined [17] to have an IC_{50} = 25 nM and a K_i =10 nM (literature report [5b,5c] of IC_{50} = 3 nM and K_i = 4.5 nM) while diester **6** was less potent with a K_i = 100 nM. This experiment suggests that the repulsive interaction between the ester groups and the carbonyls of Gly-27 and 27' is not too severe, consistent with the model proposed based upon molecular dynamics simulations.



6 X = O
7 X = NH

Structures of the type **2** represents another class of compounds which resulted from our analysis of the structure of HIV-1 protease. Figure 2 shows a "relaxed" stereoview of the average structure from production dynamics of the central unit of inhibitor **2** (P_1=Ph) bound to HIV-1 protease. Shown is the secondary hydroxyl group bound between Asp-25 and Asp-25' (catalytic groups), the hydrogen bonding network between the structural water , Ile-50 ,Ile-50' and the two carbonyl groups of the inhibitor. Hydrogen bonds between the amide NH's of the inhibitor and the Gly-27 carbonyl (also the symmetry related Gly-27') are also shown. The orientation of the P_1 phenyl groups is shown and the hydrophobic nature (counterclockwise from top left, Pro-81, Val-82, Leu-23, Ile-84) of one of the equivalent S_1 subsites is shown (at left). For clarity the inhibitor is truncated after the two amide groups which are attached to dipeptide terminating groups. The "C-terminal" ends (not shown) are oriented in the active site by hydrogen bonds in much the same fashion as seen in crystal structures of non-symmetrical inhibitors. For compound **2** the P_1 and P_2 groups are related quite nicely by a C_2 symmetry operation while the P_3 groups show some deviation from symmetry.[18]

Figure 2

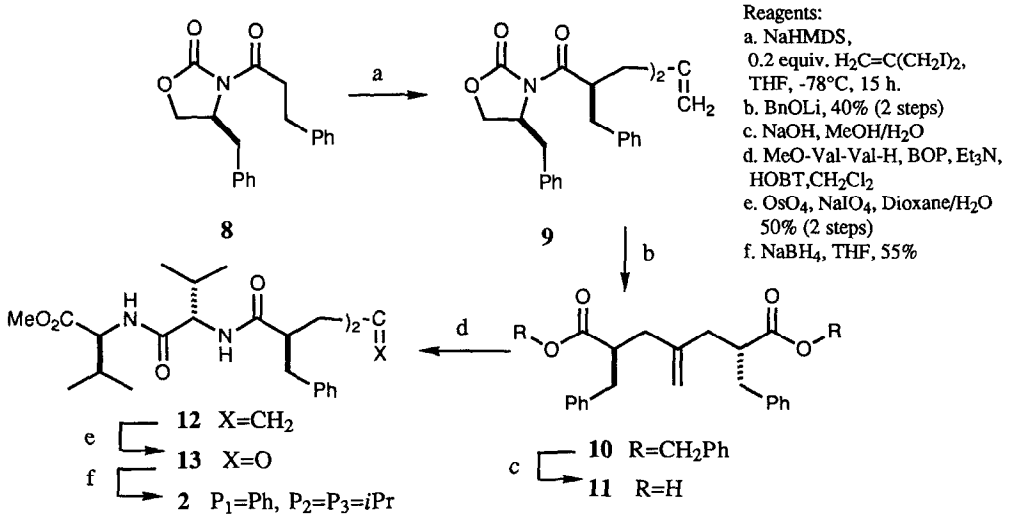


The bidirectional synthesis of **2** (P_1=Ph,P_2=P_3=*i*Pr) is shown in Scheme 2. A double alkylation of 3-iodo-2-iodomethyl-1-propene [19] using the sodium enolate of the Evans reagent **8** [20] provided **9**. This allows the introduction of C_2 symmetry in a single step with control of absolute stereochemistry. Removal of the chiral auxiliaries then gave diester **10** . Conversion to the diacid **11** and bidirectional coupling with the methyl ester of the dipeptide Val-Val afforded **12**. The final stages of the synthesis involve converting the exocyclic olefin, after it has served its key role of enhancing the electrophilicity of our diiodide, into secondary alcohol **2**. The final steps of the synthesis, coupled with the predictable stereocontrol [20] of the Evans alkylation, provide strong confidence in the structure assignment. In addition, **12** and **13** exhibit NMR spectra expected for a symmetrical molecule, and reduction of ketone **13** provides a single compound **2** (no new stereocenter is created) which has an NMR spectrum expected for a molecule without any elements of symmetry. Compound **2** proved to be a potent inhibitor of HIV-1 protease with IC_{50} = 5 nM.

The results presented demonstrate the soundness of a strategy using information about enzyme structure, mechanism and a dynamical model for inhibitor design. The fundamental, unsolved issue of calculating thermodynamic quantities from three dimensional structures (quantitative molecular modeling) forced us to use a semi-quantitative method and human analysis for evaluation of potential inhibitors. While the criteria used for evaluating putative inhibitors worked quite well in the examples reported above, we suspect that this criteria would

also allow for the selection of much weaker inhibitors. The use of a convergent bidirectional synthesis strategy and dipeptides as probe reagents allowed us to address the experimental evaluation of potential inhibitors in a timely fashion. Studies related to structure-activity relationships and *in vitro* cell data will be reported in due course.

Scheme 2



Acknowledgement: We wish to thank Dr. A. Wlodawer for providing us with coordinates for 4HVP prior to their availability in the Protein Data Base , Dr. D. Kempf for providing an NMR spectrum of **7** and Prof. S. L. Schreiber for providing us with experimental details for the preparation of **3** prior to publication.

References and Notes

1. Presented in part at "Thirty-Second Annual Medicinal Chemistry Symposium", Buffalo,NY, June 3-5, 1991

2. For recent reviews see: (a) Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A ; Heinrikson, R. L., *Chim. Oggi.* **1991**, *9*, 6-27 (b) Huff, J. R., *J. Med. Chem.* **1991**,*34*, 2305-14

3.(a) Navia, M. A.; Fitzgerald, P. M. D.; McKeever, B.M.; Leu, C.-T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P. *Nature* **1989**,*337*, 615-20 (b) Wlodawer, A.; Miller, M.; Jaskloski, M.; Sathyanarayana, B.K.; Baldwin, E; Weber, I. T.; Selk, L.; Clawson, L.; Schneider, J.; Kent, S. B. *Science*, **1989**,*245*,616-21 (c) Lapatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. *Nature* **1989**,*342*, 299-302

4.(a) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M.V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1990**, *246*, 1149-52 (b) Fitzgerald,P.M.D.; McKeever,B.M.; VanMiddlesworth, J.F.; Springer, J.P.; Heimbach, J.C.; Leu, C.-T.; Herber, W.K.; Dixon, R.A.F.; Darke, P.L. *J.Biol.Chem.* **1990**,*265*, 14209-14219 (c) Swain, A. L.; Miller, M. M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 8805-9 (d) Jaskolski, M.; Tomasselli,

A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S.B. H.; Wlodawer, A. *Biochemistry* **1991**,*30*, 1600-9

See also Ref 5b and 5g

5.(a) Blundell, T; Pearl, L. *Nature* **1989**, *337*, 596-7 See also Ref. 3c.

For examples of symmetrical and pseudo-symmetrical inhibitors see: (b) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse,J. W.; Turon, M.; Wideburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrinch, R.; Paul, D. A.; Knigge, M. *Science* **1990**, *249*, 527-33 (c) Kempf, D. J.; Norbeck, D. W.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Paul, D. A.; Knigge, M. F.; Vasavanonda, S.; Craig-Kennard, A. S.; Rosenbrook, W.; Clement, J. J.; Plattner, J. J.; Erickson, J. *J. Med. Chem.* **1990**, *33* , 2687-9 (d) Sham, H. L.; Wideburg, N. E.; Spanton, S. G.; Kohlbrenner, W. E.; Betebenner, D. A.; Kempf, D. J.; Norbeck, D. W.; Plattner, J. J.; Erickson, J. W. *J. Chem. Soc., Chem. Commun.* **1991**, 110-12 (e) Chenera, B.; Boehm, J. C.; Dreyer, G. B. *Bioorg. Med. Chem. Lett.*, **1991**,*1*, 219-22 (f) Ghosh, A. K.; McKee, S. P.; Thompson, W. J. *Tetra. Lett.* **1991**,*32* , 5729-32 (g) Bone, R.; Vacca, J. P.; Anderson, P. S.; Holloway, M. K. *J.Am.Chem.Soc.* **1991**,*113* ,9382-4

6.Rich, D. H. *J. Med. Chem.* **1985**, *28*, 263-73

7. Protease substrate nomenclature: Schecter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**,*27*, 157.

8. In all of the structures reported in Ref 4 there is a buried water molecule located near the active site which forms hydrogen bonds to two carbonyl oxygens of the inhibitor and two amide NH's of the enzyme. (Ile-50 and Ile-50'). This water molecule will be referred to the "structural water" throughout this paper.

9. For recent reviews on molecular dynamics see: (a) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992-1023 (b) Karplus, M.; Petsko, G. A. *Nature* **1990**, *347*, 631-9

10. Molecular modeling studies were performed with Quanta V2.1,V3.0 and V3.2 and CHARMm V21.1.7b, V21.2 and V21.3 distributed by Polygen Corp.

11.(a) Bartlett, P. A.; Marlowe, C. K. *Science* **1987**,*235*, 568-571 See also:(b) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. *J.Am.Chem.Soc.* **1991**,*113*, 297-307

12.(a) Bash, P. A.; Kollman, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R. *Science* **1987**,*235*, 574-6 (b) Mertz, K. M.; Kollman, P.K. *J.Am.Chem.Soc.* **1989**,*111*,5649-58

13.Tronrud, D.E.; Holden, H.M.; Matthews, B.W. *Science* **1987**,*235*,571-4

14.(a) Schreiber, S. L.; Sammakia, T.; Uehling, D. E. *J. Org. Chem.* **1989**, *54*,15-16 (b) Nakatsuka, M.; Ragan, J. A.; Sammakia, T.; Smith, D.B.; Uehling, D. E.; Schreiber, S. L. *J. Am. Chem. Soc.* **1990**, *112*, 5583-601

15.(a) Beyermann, M.; Bienert, M.; Niedrich, H.; Carpino, L. A.; Sadat-Aalaee, D. *J.Org.Chem.* **1990**, 55,721-8 (b) Carpino, L.A.; Cohen, B.J.; Stephens, K.E.; Sadat-Aalaee, S.Y.; Tien, J-H.; Langridge, D.C. *J.Org.Chem.* **1986**,*51*, 3734-6

16.Castro, B. *Synthesis* **1976**,751

17. The enzyme inhibition studies were performed using HIV-1 protease obtained from Medigenics, Omaha, NE. using a modification of the procedure in Tomaselli , A.G.; Olsen, K.K.; Hui, J.O.; Staples, D.J.; Sawyer, T.K.; Heinrikson, R.L.;Tomich, C-S.C. *Biochem.* **1990**,*29*, 264-9

18. For an X-ray structure of the complex between HIV-1 protease and a related inhibitor see Ref. 5g

19.Skell, P. S.; Doerr, R. G. *J.Am.Chem.Soc.* **1967**,*89*, 4688

20.Evans , D. A. ;Ennis, M. D.; Mathre, D. J. *J.Am.Chem.Soc.* **1982**,*104*, 1737-9