Volume 269, number 1, 269–272 FEBS 08784 August 1990

Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap

Alla Gustchina* and Irene T. Weber

Crystallography Laboratory, NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, MD 21701, U.S.A.

Received 4 June 1990; revised version received 28 June 1990

The crystal structure of HIV-1 protease with an inhibitor has been compared with the structures of non-viral aspartic proteases complexed with inhibitors. In the dimeric HIV-1 protease, two 4-stranded β-sheets are formed by half of the inhibitor, residues 27–29, and the flap from each monomer. In the monomeric non-viral enzyme the single flap does not form a β-sheet with an inhibitor. The HIV-1 protease shows more interactions with a longer peptide inhibitor than are observed in non-viral aspartic protease-inhibitor complexes. This, and the large movement of the flaps, restricts the conformation of the protease cleavage sites in the retroviral polyprotein precursor.

Retroviral protease; Aspartic protease; Enzyme-substrate interaction; Retroviral polyprotein precursor

1. INTRODUCTION

The crystal structure of HIV-1 protease (PR) has been determined in several laboratories [1–3]. More recently, the structure of HIV-1 PR complexed with an inhibitor has been determined [4] and a second cocrystal structure with a different inhibitor is available [5]. The catalytically active dimer of retroviral PR is structurally similar to the monomeric nonviral aspartic proteases which consist of two domains. The PR subunit and each domain of the pepsin-like proteases have a common structural core consisting of two layers of β-sheets with approximately orthogonal directions of the chains [2,6–8] and a characteristic distribution of the hydrophobic residues which form the hydrophobic nucleus of the structure [9]. The interdomain region of non-viral aspartic proteases consists of a 6-stranded β-sheet [10], in contrast to the equivalent intersubunit region in the PR dimer which consists of a 4-stranded β-sheet formed by the four termini. The catalytic triplet, Asp-Thr-Gly, has a conserved sequence in all retroviral and non-viral aspartic proteases [11] and almost identical conformation in all known crystal structures. The surface β-hairpin called the flap, residues 42 to 58 of HIV-1 PR, has the structural and functional analog in non-viral aspartic proteases of residues 70 to 83 (pepsin numbering). The flap is a flexible structural unit in all aspartic proteases; for example, binding of the pepstatin fragment to penicillopepsin alters the position of the flap's tig by 2.2 Å [12]. The function of the flap, by analogy to pepsin-like proteases, is to bind inhibitor and possibly also to exclude water from the catalytic site. The inhibitor binding interactions in the crystal structures of retroviral and non-viral aspartic proteases have been compared.

Correspondence address: A. Gustchina, Crystallography Laboratory, NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, MD 21701, USA

*On leave from the V.A. Engelhardt Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow, USSR

2. MATERIALS AND METHODS

The crystal structure of the fungal aspartic protease, rhizopuspepsin, with a reduced peptide inhibitor, D-His-Pro-Phe-His-Phe-Ψ[CH_2-NH]-Phe-Val-Tyr, [13] was obtained from the Protein Data Bank (3APR). The structure of HIV-1 PR is PDB entry 3HVP [2]. The complex of HIV-1 PR with reduced peptide inhibitor, MVT-101, (Ac-Thr-Ile-Nle-Ψ[CH_2-NH]-Nle-Gln-Arg-amide, where Nle is norleucine and K_i = 780 nM) [4] is PDB 4HVP, while the coordinates of a complex with inhibitor JG-365 (Ac-Ser-Leu-Ans-Phe-Ψ[CH(OH)CH_2N]-Pro-Ile-Val-OMe, with K_i = 0.66 nM for a mixture of both R and S diastereomers) [13] were provided by Dr. Swain. The structures were superimposed on α-carbon atoms using the programs, ALIGN [14], or COORDTRANS (R.W. Harrison, unpublished), and were examined on an Evans and Sutherland computer graphics system using the program, FRODO [15].

3. RESULTS AND DISCUSSION

In the two cocrystal structures of HIV-1 PR [4,5], the inhibitor, as proposed from a model complex [7], binds in an extended β-conformation with similar interactions to those observed in the structures of complexes of non-viral aspartic proteases with inhibitors [13,16–20]. There are hydrogen bond interactions to the inhibitor provided by residues near the two catalytic triplets, and by residues from the flap (residues 74 to 76 in pepsin).

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

Amino acid side chains of peptide-like inhibitors bind in successive subsites formed by the protease and contribute to the specificity of binding. Large hydrophobic residues are often found at P1 and P1′ in substrates of HIV-1 PR. Positions P2 to P4 and P2′ to P4′ are less selective and can accommodate a wide variety of amino acid side chains. The binding affinity probably depends on the specific combination of side chains, as long as it does not interfere with the formation of hydrogen bonds to the peptide groups of the substrate. We consider only the interactions formed to the NH and C = O groups of the peptide-like inhibitors, since these are independent of the type of residues forming the inhibitor. Table 1 compares these interactions in HIV-1 PR and the fungal aspartic protease, rhizopus pepsin. In all cases, the catalytic aspartic acids (Asp-25 in HIV PR) are near the scissile bond between P1 and P1′, and the C = O of the glycines of the third position in the triplet (Gly 27 in HIV PR) accept hydrogen bonds from the inhibitor NH at P1 and P2′ (Fig. 1a). In addition, the fourth residue following the catalytic Asp interacts with the NH and C = O of P3 in the inhibitor. This residue is Asp-29 in HIV PR, and Thr-222 in rhizopus pepsin. The interactions of the flap differ in viral and non-viral enzymes.

The most significant differences in the arrangement of the ligand binding site in viral and non-viral aspartic proteases are due to the more symmetrical environment in the retroviral enzyme. There are two flaps in the active dimer of HIV-1 PR compared to the single flap in non-viral aspartic proteases. Residues in the flap region form numerous interactions with the inhibitor in both cases, although the flaps are oriented differently in the structures of viral compared to non-viral enzymes. In

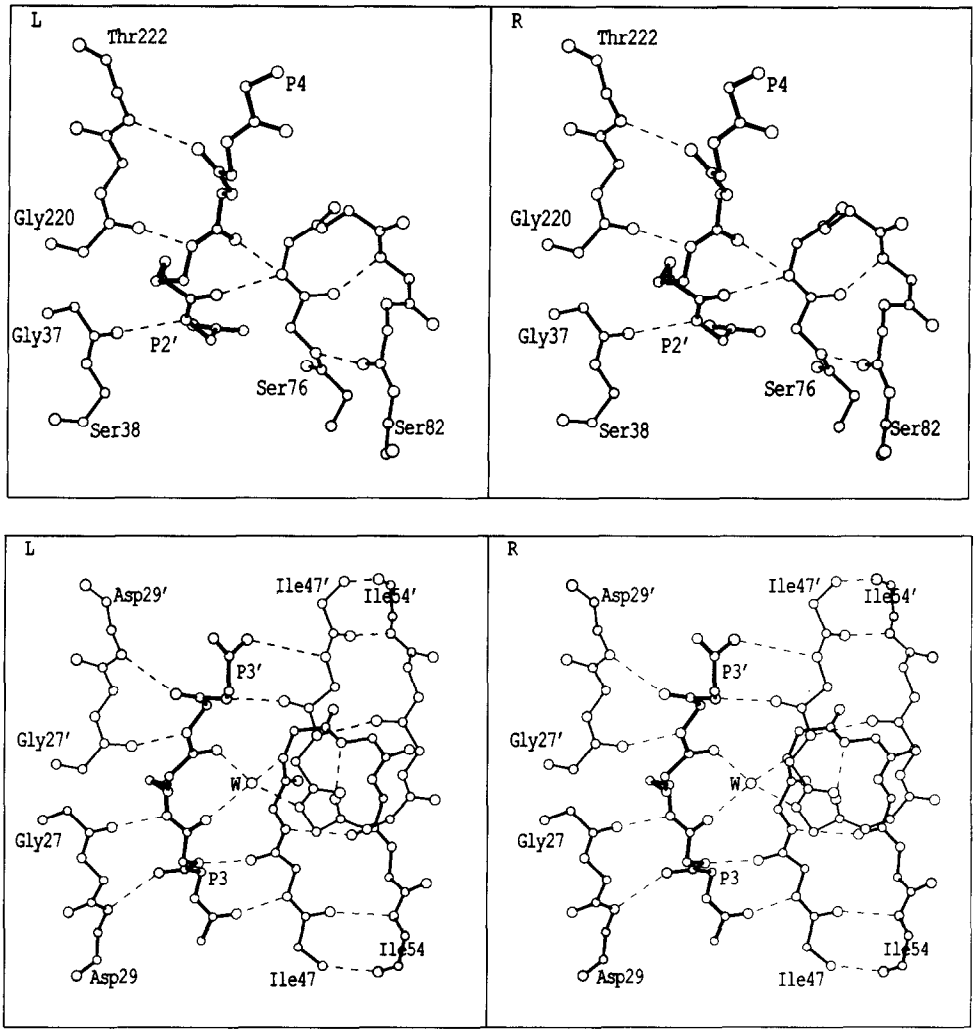

Fig. 1. (a) A stereo view of the structure of rhizopuspepsin complexed with an inhibitor [13]. Only main chain atoms are shown for residues 37–39, 220–222 (equivalent to HIV-1 PR 27 to 29), and residues 76–82 which form the single flap. The inhibitor P4 to P2′ occupying the subsites is indicated by thick lines. (b) A stereo view of HIV-1 PR complexed with an inhibitor [4] in a similar orientation to Fig. 1. Only main chain atoms are shown for clarity. The two β-sheets formed by the two flaps, residues 47–54, the inhibitor occupying the subsites P3 to P3′ (thick lines), and residues 27–29 near the active site are shown. Residues in the second subunit are indicated by a prime and thin lines. Water is W, and hydrogen bond interactions are indicated by dashed lines.

Table I

Potential hydrogen bond interactions between HIV PR or rhizopuspepsin and their inhibitors

Subsite	Substrate atom	HIV PR atom	Rhizopus pepsin atom	
S4	P4 NH			
	P4 C=O	NH Gly 48′		flap
S3	P3 NH	O Asp 29′[b]	OH Thr 222	
	P3 C=O	NH Asp 29′	NH Thr 222	
S2	P2 NH	C=O Gly 48′	O Asp 79+	flap
	P2 C=O	Water 1	NH Gly 78	flap
S1	P1 NH	C=O Gly 27′	C=O Gly 220	*[a]
	P1 C=O	Asp 25	Asp 218	*[a]
	scissile bond			
S1′	P1′ NH	Asp 25′	Asp 35	*[a]
	P1′ C=O	Water 1	NH Gly 78	flap
S2′	P2′ NH	C=O Gly 27	C=O Gly 37	*[a]
	P2′ C=O	NH Asp 29	NE Trp 194	
S3′	P3′ NH	C=O Gly 48		flap
	P3′ C=O	NH Gly 48[c]		flap
S4′	P4′ NH	O Asp 29+		

Interactions with the NH and C = O of the inhibitors [4,5,13,20]. Residues from the two subunits in the HIV PR dimer are distinguished by a prime.
[a] The active site triad, Asp-Thr-Gly.
[b] O atom from the carboxyl side chain of Asp.
[c] This interaction is seen in the structure of Swain et al. [5], and is replaced by a hydrogen bond to a water molecule in the structure of Miller et al. [4].

the dimer of HIV-1 PR, the two flaps lie almost parallel to the inhibitor, and form two short β-sheets, one with half of the inhibitor (P4 to P1, and P1′ to P3′), and a region near the catalytic aspartic acid (residues 27–29) (Table I; Fig. 1b). The two β-sheets are separated near the scissile bond where the two flaps overlap, and there are hydrogen bond interactions with a water molecule. The main-chain-main-chain hydrogen bonds of the β-sheets formed by PR and the inhibitor extend from P4 to P3′. A total of 12 hydrogen bond interactions occur between the enzyme and the inhibitor NH and C=O groups; the flaps are involved in 6 of these, and Gly 27 to Asp 29 from both subunits provide the other 6 interactions.

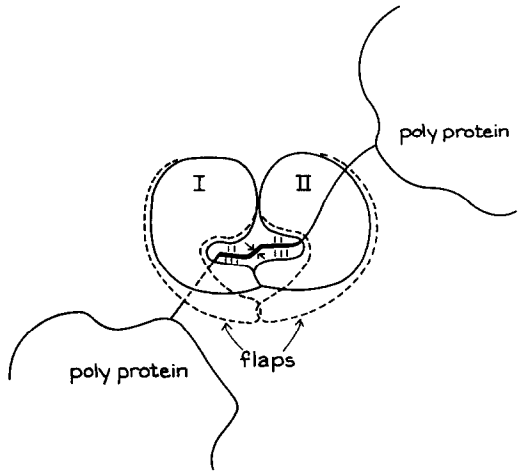

Fig. 2. Schematic diagram of a PR dimer bound to polyprotein substrate. The polyprotein is labeled and shown as two folded regions separated by an extended chain around the scissile bond. The PR subunits are indicated by I and II, with continuous lines indicating the PR dimer structure with substrate bound, and the dashed lines indicating the PR structure in the absence of substrate. The two flaps overlap at one side of the substrate binding region. The scissile bond is indicated by two small arrows, and the dashes indicate hydrogen bond interactions between the substrate peptide and the PR dimer.

In the non-viral aspartic proteases, the single flap is more perpendicular to the bound inhibitor, and forms interactions with P1 to P2′ of the inhibitor. The interactions with the enzyme extend from P3 to P3′ of the inhibitor, and involve 2 to 3 hydrogen bond interactions with residues of the flap and 3 to 4 additional interactions with the residues near the catalytic aspartates (Fig. 1a). No β-sheet can be formed as seen in the HIV-1 PR, and the single flap of pepsin-like proteases contributes fewer interactions with the inhibitor than the two flaps of HIV-1 PR. This difference between HIV and the non-viral proteases is consistent with the experimental measurement of hydrolysis rates with peptide of different lengths. In HIV PR, when the substrate, SQNYPIV, which extends from P4 to P3′, is shortened by one residue on either side, the rate of hydrolysis decreases greatly [21]. This indicates that 7 residues is the minimal substrate size. Measurements on penicillopepsin show a large decrease in catalytic rate on removing residues at P3 or P2′ which suggests a preferred substrate size of at least 5 residues [22].

The viral substrate of HIV-1 PR is the multi-domain polyprotein which consists of folded domains representing the individual protein products separated by linking regions containing the cleavage sites. Comparison of the crystal structures of the HIV-PR-inhibitor complex and the native enzyme suggests that binding of the inhibitor changes the relative orientation of the two subunits in the dimer by a rotation of about 1.7°, and the ends of the flaps move by approximately 7 Å [4].

The single flap of pepsin-like proteases does not require such a large movement for substrate to enter the binding site, and changes of only about 2 Å have been observed in the structures with and without an inhibitor.

In the native HIV-1 PR dimer the ends of the two flaps overlap and form an intersubunit hydrogen bond. The overlapping flaps would prevent entry of the polyprotein substrate into the active site. The two flaps also overlap in the HIV-1 PR complexes with inhibitors, thus 'pinning' the inhibitor in the binding cleft. Therefore both the binding of substrate or inhibitor to the enzyme, and the release of products require a substantial movement of the flaps. Preliminary modeling suggests that the flaps must move by about 15 Å from their position in the inhibitor complex in order to allow the polyprotein to enter the active site.

The substantial conformational changes in the enzyme allow more interactions between PR and the substrate (Fig. 1b), and suggest that the PR dimer is quite flexible in solution. The numerous ligand-enzyme interactions impose a restricted conformation on the linker region in the polyprotein precursor and immobilize the dimeric protease (Fig. 2), providing productive binding. After enzymatic cleavage, each half of the substrate is bound only to one monomer, and dimer flexibility can be restored. The large adjacent domains in the polyprotein may contribute to the independent motion of the monomers facilitating dissociation of the enzyme-substrate complex with subsequent release of the products.

In summary, the inhibitor in HIV PR is stabilized by additional interactions with the two flexible flaps which form two short β-sheets, as opposed to the single flap of non-viral proteases. The minimum substrate size is 7 residues for HIV PR, compared to 5 for pepsin-like proteases. This, together with the requirement for a substantial motion of the flaps, has consequences for the structure of the PR cleavage sites in the retroviral polyprotein. The linker region must be sufficiently long to enter the PR binding site without undesirable close contacts with the rest of the polyprotein. This requirement for about 8–10 residues in a more extended β-conformation may be a significant component of the apparent specificity of binding that is not directly related to the amino acid sequence around the cleavage site.

Acknowledgements: We thank Drs. Anna Marie Skalka, David Davies, Alexander Wlodawer, Steve Foundling and Robert Harrison for helpful discussions, Dr. Amy Swain for providing us with the coordinates of the protease complexed with the JG-365 inhibitor prior to publication, and Elizabeth Kincaid for preparing Fig. 2. This research was sponsored by the National Cancer Institute, DHHS, under contract No. N01-C01-74101 with ABL. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

REFERENCES

[1] 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. and Springer, J.P. (1989) Nature 337, 615–620.
[2] Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J. and Kent, S. (1989) Science 245, 616–621.
[3] 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. and Hobart, P.M. (1989) Nature, 342, 299–302.
[4] Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Kent, S.B.H. and Wlodawer, A. (1989) Science 246, 1149–1152.
[5] Swain, A.L., Miller, M.M., Green, J., Rich, D.H., Kent, S.B.H. and Wlodawer, A. (1990) Proc. Natl. Acad. Sci. USA (in press).
[6] Miller, M., Jaskolski, M., Mohana Rao, J.K., Leis, J. and Wlodawer, A. (1989) Nature 337, 576–579.
[7] Weber, I.T., Miller, M., Jaskolski, M., Leis, J., Skalka, A.M. and Wlodawer, A. (1989) Science 243, 928–931.
[8] Pechik, I.V., Gustchina, A.E., Andreeva, N.S. and Fedorov, A.A. (1989) FEBS Lett. 247, 118–122.
[9] Andreeva, N.S. (1988) in: Biosynthesis and Structure of Proteins, Pushchino 3, 87–96.
[10] Blundell, T. and Pearl, L. (1989) Nature, 337, 596–597.
[11] Toh, H., Ono, M., Saigo, K. and Miyata, T. (1985) Nature, 315, 691.
[12] James, M.N.G., Sielecki, A., Salituro, F., Rich, D.H. and Hofmann, T. (1982) Proc. Natl. Acad. Sci. USA 79, 6137–6141.
[13] Suguna, K., Padlan, E.A., Smith, C.W., Carlson, W.D. and Davies, D.R. (1987) Proc. Natl. Acad. Sci. USA 84, 7009–7013.
[14] Satow, Y., Cohen, G.H., Padlan, E.A. and Davies, D.R. (1984) J. Mol. Biol. 190, 593–604.
[15] Jones, A.T. (1978) J. Appl. Crystallogr. 11, 268–272.
[16] James, M.N.G. and Sielecki, A.R. (1985) Biochemistry, 24, 3701–3713.
[17] Blundell, T., Cooper, J. and Foundling, S.I. (1987) Biochemistry, 26, 5585–5590.
[18] Foundling, S.I., Cooper, J., Watson, F.E., Cleasby, A., Pearl, L.H., Sibanda, B.L., Hemmings, A., Wood, S.P., Blundell, T.L., Valler, M.J., Norey, C.G., Kay, J., Boger, J., Dunn, B.M., Leckie, B.J., Jones, D.M., Atrash, B., Hallatt, A. and Szelke, M. (1987) Nature, 327, 349–352.
[19] Cooper, J., Foundling, S., Hemmings, A., Blundell, T.L., Jones, D.M., Hallett, A. and Szelke, M. (1987) Eur. J. Biochem. 169, 215–221.
[20] Cooper, J.B., Foundling, S.I., Blundell, T.L., Boger, J., Jupp, R.A. and Kay, J. (1989) Biochemistry 28, 8596–8603.
[21] Darke, P.L., Nutt, R.F., Brady, S.F., Garsky, V.M., Ciccarone, T.M., Leu, C-T, Lumma, P.K., Freidinger, R.M., Veber, D.F. and Sigal, I.S. (1988) Biochem. Bioph. Res. Commun. 156, 297–303.
[22] Hofmann, T., Allen, B., Bendiner, M., Blum, M. and Cunningham, A. (1988) Biochemistry 27, 1140–1146.