Bioorganic & Medicinal Chemistry Letters, Vol.3, No.10, pp. 2105-2112, 1993
Printed in Great Britain

0960-894X/93 $6.00 + .00

EVALUATION OF COMPLEXATION ENERGIES BETWEEN ASPARTIC PROTEASE OF HIV-1 AND ITS HEXAPEPTIDE INHIBITORS INCLUDING THE EFFECT OF SOLVATION

Stanislav MIERTUS

International Institute for Pure and Applied Chemistry ICS UNIDO
Area di Ricerca, Padriciano 99, TRIESTE, Italy

(*Received in USA* 2 *April* 1993; *accepted* 13 *August* 1993)

Abstract: The complexation energy and their individual contributions between the Aspartic Protease enzyme (AP) of human immunodeficiency virus and hexapeptide inhibitors(I) with CH_2NH reduced peptide bond have been first evaluated using the method of molecular mechanics. Then the evaluation of solvation energies for individual species AP:I, AP and I has been done by polarizable continuum method. As to the complexation itself, the dominant role for the electrostatic interaction is found. Various replacements of P3.....P3' inhibitor residues have been modelled with the goal to design inhibitors with the high affinity towards AP enzyme. Solvation energy modifies the order of stability of complexes depending on the solvent polarity. New, hypothetically more active structures have been forwarded for synthesis and testing.

Aspartic protease (AP) is one of virally encoded enzymes necessary for replication of HIV-1 and therefore it is considered as a potential target for the design of anti AIDS drugs. Recently, X-ray structures of AP of HIV-1[1] and of some complexes between HIV-1 protease and reduced oligopeptide inhibitors have been elucidated[2-3], showing the importance of hydrogen bond contacts between the enzyme and the inhibitor[4]. Consequently, attempts have been made to design modified structures of the oligopeptide analogues[5-8] based on the transition state mimetic concept, where various reduced forms (isosters) were taken into consideration. Recently, X-ray structures of AP-inhibitors complexes have been reviewed[9].

Original X-ray structural analysis of AP:I complex[2], with MVT-101 inhibitor (Thr-Ile-Nle-Ψ[CH2-NH]-Nle-Gln-Arg+) revealed close contacts between inhibitor residues and those of AP, which enabled to estimate a type of prevailing interactions (hydrophobic vs. ionic) for some of inhibitor residues. This inhibitor (MVT 101) was also used as a pattern for the design of new structures [6,7]. Thaisrivongs et al.[7] using the molecular modelling (with Amber force field) came to qualitative conclusions, that each inhibitor amid group between P_3' and P_3 positioned to make two H-bonds to the protein backbone, except for the P_2 and P_1' carbonyl oxygens (which are indirectly H-bonded to the protein through buried water molecule). Moreover they found low sensitivity to the choice of residues at P3

due to apparent lack of the specific interactions between this position and the enzyme. However, this study doesn't report any numerical value demonstrating these conclusions.

Recently we have started systematic molecular modelling studies[10] with the focus to analyse in details most important factors which are expected to govern the strength of complexation processes between AP and its inhibitors.

Firstly, the role of overall interactions of individual residues (P3' ...P3) with whole enzyme (i.e. not only those through direct H-B contacts) should be evaluated. In this respect electrostatic long range interactions can be especially important.

Secondly, the stability of the AP:I complex depends not only on the strength of interaction within the complex, but also on the relative stability of all species involved in the complexation process, i.e.

$$AP + I \quad <==> \quad AP:I \qquad (1)$$

This stability is also influenced by the solvent polarity especially because all species involved in complexation can be charged. Consideration of the "crystallographic" water molecules only (i.e. 17 solvent molecules surrounding AP enzyme) didn't represent sufficiently the effect of solvent and gave almost same results as *in vacuo*[10a].

Therefore, in the present paper we present results related to MVT 101 inhibitor (abbreviated as 'Orig') as a reference structure analysing the energetical contributions to AP:inhibitor residues interaction and we have analysed in details both effects i.e. the interaction between AP and I as well as the solvation on complexation equilibrium. Such analysis enables also to design the structural changes of original MVT-101 inhibitor leading to the variation of the complexation and solvation energies and consequently to the expected increase of anti-HIV activity.

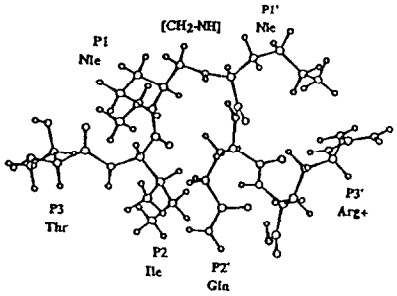

Figure 1 The structure of hexapeptide MVT-1 inhibitor[2]. Its sequence is Thr - Ile - Nle - [CH2-NH] - Nle - Gln - Arg+.

Molecular mechanics with cvff force field[11] of Biosym - Insight II package[12] have been used for the evaluation of total energies E and their components i.e. bond, nonbonding

(dispersion + repulsion) and electrostatic ones of individual systems (AP:I complex, AP enzyme and I inhibitor). The complexation energy E_{compl} related to Eq.1 has been then evaluated as

$$E_{compl} = E_{AP:I} - E_{AP} - E_I \qquad (2)$$

Individual energetic components of E_{compl} (bonding, nonbonding and electrostatic) have been evaluated in the same way. Relative changes of these quantities have been calculated for various inhibitors (abbreviated as Mod) and related AP:I complexes, where the inhibitor structures were varied by replacing amino acid residues P3.....P3' in the reference MVT-1 hexapeptide structure (Fig.1). Thus, the relative changes in the complexation energy for individual modified inhibitors is

$$\Delta E_{compl}(Mod) = E_{compl}(Mod) - E_{compl}(Orig) \qquad (3)$$

The geometries of all inhibitors have been optimized both in the free form as well as in AP:I complex. The X-ray structure of AP:Orig complex[2] was used as a starting geometry. The geometry of enzyme part in the AP:I complex was kept fixed following the docking study of AP:I complexes of Caflisch et al.[13]. Such geometry corresponds to the enzyme structure already adopted to bind an inhibitor. The geometries of inhibitor structures were optimized using the conjugate gradient method[14] starting from different geometries and searching for absolute minima. The solvent effect has been modelled using the Polarizable Continuum Model (PCM), where the solvent is represented by a homogeneous dielectric medium with permitivity ε (see refs. 15,16 for details). The solvation energies have been evaluated for all species considered i.e. AP:I complex as well as separated AP and I and solvation contributions to the complexation process (eq. 1) have been evaluated by analogy to Eqs. 2 and 3.

First, the interaction and complexation energies between known inhibitor Orig and AP enzyme *in vacuo* has been evaluated using eq. 2. The complex AP:Orig is stabilized by 124.0 kcal/mol vs. separated species. The electrostatic term contributes mainly to the stabilization (by 89.1 kcal/mol). The Orig inhibitor in spite of its positive charge Q_{Orig} = +1, due to Arg+ in the Orig structure in position P3' (Fig. 1) is stabilized by AP (with total charge Q_{AP} = +4) due to the fact that the negatively charged residues are placed closer to the active site than positively charged ones, thus creating a negative electrostatic potential within the large part of the enzyme pocket hole and stabilizing positively charged residues of the inhibitors.

As to the detailed structure of AP:I complex, our calculations with cvff force field give the same qualitative picture about hydrogen bonding between amide groups of Orig and

AP obtained by Thaisrivongs et al.[7] using Amber force field. It supports the opinion [11] that both Amber and cvff force fields lead to very similar results especially for proteins. However, to evaluate the role of individual residues (as those above discussed P3, P3') it is worthwhile to calculate electrostatic contributions of interaction energy $E_{elst}(AP:P_k)$ between AP and each residue P_k of the inhibitor in the complex AP:I, namely:

$$E_{elst}(AP:P_k) = \Sigma_{j=1} V_{j(AP)} \cdot q_j(P_k) \qquad (4)$$

where $V_{j(AP)}$ is the electrostatic potential coming from the charge distribution of all residues of AP at each atom j of the inhibitor and $q_j(P_k)$ is the net charge of atom j of the P_k-th residue. The calculated interaction energies (in kcal/mol) for the Orig inhibitor are following: Thr(P3) -3.8 Ile(P2) -10.4 Nle(P1) -5.9 CH_2NH +3.3 Nle(P1') -27.7 Gln(P2') -1.8 Arg^+(P3') -48.4 .

It can be seen from the given values, that all the amino acid residues of inhibitor (P3.....P3') - except the reduced bond itself - contribute to the stability of AP:I complex, however the contributions differ considerably for individual residues. The largest stabilization comes from Arg+(P3') -48.4 kcal/mol and quite stabilizing are also Ile(P2) -10.4 kcal/mol and Nle (P1') -27.7 kcal/mol, the last one partially due to formal positive fractional charge(+0.22, while on CH2NH is -0.22). On the other hand, the stabilization coming from Thr(P3) and Gln (P2') is very small (-3.8 and -1.8 kcal/mol resp.). The contribution of CH2NH group is slightly destabilizing (+3.3 kcal/mol) due to the negative fractional charge (-0.22) on this group. The role of various structures on the place of central reduced bond is analysed in details elsewhere[10b].

The small value E_{elst} for Thr (P3) and Gln (P2') could indicate that hydrophobic inhibitor residues would be preferred at these sites as it was suggested by Thaisrivongs et al.[7]. However, when we have replaced the P3 residue by Arg^+ its stabilization (calculated according to Eq.4) increased to -22.6 $kcal.mol^{-1}$. Surprisingly, the Asp^- replacement on P3 brings even higher stabilization (-44.4 $kcal.mol^{-1}$). This peculiar behaviour is explained when looking at the spatial distribution of the electrostatic potential V inside P3 site and the mobility of P3 residue. While P3 residue (being Asp- or Arg+) is sitting on the end of active site (but not outside of the cleft as it appears for bulky substituents like chinoline[7]) there is enough space so that the side chain of P3, when charged, can be reoriented to positive or negative regions of V created by enzyme. Negatively charged P3 residue (Asp^- or Glu^-) when sitting in the active site pocket of enzyme is reoriented towards Arg^+108, thus giving considerable stabilization. On the other hand, if Arg^+ is on P3 site, it is oriented towards stabilizing Asp^-29 of AP.

This conclusion can be generalized, that certain sites of AP enzyme pocket can accept very different inhibitor residues- even with the opposite charge - if there is enough space for reorientation and opportunities of different stabilization by enzyme residues.

However, the inhibition potency is more tightly related to the total complexation energy (Eq. 2) as well as to that including solvent effect. Therefore we have modelled several replacements of residues (P3,P2,P2',P3') in the hexapeptide inhibitor structure with the particular aim to design replacements modifying the total charge of the inhibitor and thus mainly affecting the electrostatic part of AP:I interactions. In Table 1 we present the relative changes of complexation energy contributions *in vacuo*, evaluated according to Eqs. 3 and 2 for representative series of modified inhibitor structures.

Table 1. Complexation energy E_{compl} *in vacuo* and their components of individual modified inhibitor structures with AP related to those of known inhibitor Orig (according to Eq. 3).

Inhibitor*	$Q^{\$}$	ΔE_{bond}	ΔE_{nb}	ΔE_{elst}	ΔE_{compl}
Mod-III	+3	+16.3	-10.8	-22.6	-17.1
Mod-II	+2	+8.5	-2.0	-21.1	-14.6
Mod-Ia	+2	+16.4	-6.8	-19.1	-9.5
Orig	+1	0.0	0.0	0.0	0.0
Mod-IV	0	-0.6	+0.4	+10.2	+11.2
Mod-Ib	0	-1.4	+1.3	+34.9	+35.8
Mod-V	-1	-3.3	0.0	+35.0	+31.7
Mod-VI	-2	-10.6	-3.2	+46.2	+32.4

$^{\$}$Q is the total charge of the inhibitor.
*Abbreviations of inhibitor structures are following; (X is CH_2NH reduced bond)
Orig = Thr-Ile-Nle-X-Nle-Gln-Arg+(Fig.1); Mod-Ia = Arg+-Ile-Nle-X-Nle-Gln-Arg+
Mod-Ib = Asp^--Ile-Nle-X-Nle-Gln-Arg+ Mod-II = Lys+-Ile-Nle-X-Nle-Gln-Arg+;
Mod-III = Arg+-Arg+-Nle-X-Nle-Gln-Arg+ Mod-IV = Thr-Ile-Nle-X-Nle-Gln-Thr;
Mod-V = Thr-Ile-Nle-X-Nle-Glu^--Thr Mod-VI = Thr-Glu^--Nle-X-Nle-Glu^--Thr

The trend of stability based on total complexation energy *in vacuo* follows the changes of the total charge Q and thus the changes of electrostatic component of interaction. The Mod-III structure with Q=+3 forms the most stable complex. The complexation of Mod-Ia and Mod-II inhibitors by enzyme is still stronger than that of Orig inhibitor. The structural modifications leading to the zero or negative charges (Mod-IV, Mod-Ib, Mod-V, Mod-VI) destabilize the AP:I complex in comparison to Orig structure. It is interesting to note that the destabilization gradually increases with the decrease of positive charge of I.

The Mod-Ib structure is the particular case. There, the interaction between AP and I calculated according to Eq. 4 is the strongest one from all inhibitors considered due to a favourable stabilization of Asp- (P3) mentioned above. On the other hand, the value of

E_{compl} is smallest one (i.e. weakest complexation from all inhibitors considered) due to very high stability of I alone, which comes from strong reorientation of hexapeptide chain (in I free) approaching Asp-(P3) to Arg+(P3') and thus giving a considerable electrostatic stabilization.

The changes of total energy correlate roughly with those of the electrostatic component. Nonbonding components contribute in less extent to the relative stability in case of more positively charged inhibitors (Mod-Ia - Mod-III), while for Mod-IV - Mod-V their contribution is very small. The contributions of bonding energy components are destabilizing for positively charged inhibitors and stabilizing for negatively charged ones. In case of Mod-Ia and Mod-II, where the total charges are the same (Q= +2) and consequently electrostatic contributions are very similar, there are bonding energy changes responsible for the preference of the Mod-II inhibitor (with Lys+ at P3 and P3'sites) over Mod-Ia (with Arg+ at the same sites). Anyway the electrostatic component clearly dominates at the complexation equilibrium *in vacuo* and provides the following order of hypothetical inhibition activity:

Mod-III > Mod-II > Mod-Ia > Orig > Mod-IV > Mod-V > Mod-VI > Mod-Ib.

In the following part we have evaluated solvation energies as described above for all considered species, i.e. AP:I, AP and I respectively. We considered a polar surrounding solvent (water) with $\varepsilon = 80$ which represents approximately real conditions. For the Orig inhibitor the solvation stabilizes free species by 21.3 kcal/mol *vs* complex AP:I. However, considering the complexation energy E^{compl} (including solvation) the complexation is still in favour of the tight complex AP:I ($E_{\varepsilon=80}^{compl} = 102.7$ kcal/mol). Based on solvation energies calculated for all modified inhibitor structures considered above, we have evaluated relative changes in solvation energies $\Delta E_{\varepsilon}^{solv}$ by analogy with Eq. 3. Then, relative changes of solvation energies were considered together with complexation energies *in vacuo* (Table 1) thus obtaining relative changes of complexation energies in solution $\Delta E_{\varepsilon}^{compl}$. All values are reported in Table 2.

It can be seen, that the increase of the charge of inhibitor to +2 and +3 (Mod-Ia, Mod-II, Mod-III) leads to positive values of $\Delta E_{\varepsilon=80}^{solv}$ i.e. the solvation in very polar environment reduces the stability of AP:I complex for these inhibitors because of better solvation of charged species (AP and I) when are separated. On the other hand, due to weaker solvation of Mod-IV, Mod-Ib (neutral) inhibitors and Mod-V (Q = -1) in free forms, the solvation contributes to the stability of complexes of these inhibitors with AP.

Generally, when changing the charge of inhibitors, the relative effects of solvation on complexation equilibrium are opposite than those of complexation *in vacuo*.

Table 2 Solvation contributions to the complexation energy DE_{ε}^{solv} and complexation energies including solvation $\Delta E_{\varepsilon}^{compl}$ for various I structures in AP:I=AP + I reactions

Inhibitor*	Q	$\Delta E_{\varepsilon=80}^{solv}$	$\Delta E_{\varepsilon=80}^{compl}$
Mod-III	+3	+11.1	-6.0
Mod-II	+2	+18.1	+3.5
Mod-Ia	+2	+18.0	+8.5
Orig	+1	0	0
Mod-IV	0	-29.7	-18.5
Mod-Ib	0	-28.6	+6.3
Mod-V	-1	-12.7	+19.0
Mod-VI	-2	+6.5	+38.9

* See also comments in the footnote of Table 1

By considering the solvation and complexation together (last column of Table 2) we can evaluate the order of overall stability of individual complexes, which has to correlate with anti HIV activity of these inhibitors.

The trend is the following:

Mod-IV > Mod-III > Orig > Mod-II > Mod-Ib > Mod-Ia > Mod-V > Mod-VI

It follows that at least two structures namely Mod-III and Mod-IV are expected to be more active in a biological-like environment. The designed structures have been forwarded to the synthesis and testing to verify this conclusion.

It can be also concluded from gained results that at least three factors are important for AP:I complexation and consequently for HIV inhibition, namely: i) interactions between AP and I residues within the complex, which are often dependent not only on the nature of interacting residues, but also on the possibility of conformation reorientation within the complex; ii) relative stability of individual species (AP:I, AP, I), which can vary significantly, mostly due to conformational differences for I in the complex *vs* I free; iii) the effect of solvation shifting considerably the complexation equilibrium especially when inhibitors are charged.

Acknowledgements

The author thanks to ICS UNIDO and especially to N. Rahman for the support of this work. Valuable discussion with V. Frecer is also acknowledged.

References

1. Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J.R., Whittle, P.J., Danley, D.E., Goegheran, K.F.,Hawrylik, S.J., Lee,

E.S., Scheld, K.G., Hobart,P.M. *Nature* **342**, 299-302 (1989).
2. Miller, M., Schneider, J., Sarhyanarayana, B.K., Toth M.V., Marshall G.R., Clawson, L., Selk, L., Kent S.B.H., Wlodawer, A. *Science* **246**, 1149-1152 (1989).
3. Swain, A.L., Miller, M., Green, J., Rich, D.H., Schneider, J., Kent S.B.H., Wlodawer, A. *Proc. Natl. Acad.Sci. USA* **87**, 8805-8809 (1990).
4. Weber, I.T., Miller, M., Jaskolski, M.,Leis, J., Skalka,A.M., Wlodawer, A. *Science***243**, 928-931 (1989).
5. Roberts, N.A., Martin, J.A., Kinchington, D., Broadhurst, Craig, J.C., Duncan, I.B., Galphin, 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.J., Machin, P.J. *Science* **243**, 928-931 (1989).
6. Tomasselli, A.G., Olsen, K.M., Hui, J.O., Staples, D.J., Sawyer, T.K., Heinrikson, R.L., Tomich C.S.C. *Biochemisrty* **29**, 264-269 (1990).
7. Thaisrivongs, S., Tomasselli, A.G., Moon, J.B., Hui,J., McQuade, J.,Turner S.R., Strohbach, J.W., Howe, W.J., Tarpley W.G., Heinricson, R.L., *J. Med. Chem.* **34**,.2344-2356 (1991).
8. Sansom,C.E., Wu, J., Weber, I.T. *Protein Engineering* **5**, 659-667 (1992).
9. Wlodawer, A., Erickon, J.W. *Annu. Rev. Biochem.* **62**, 543-585 (1993).
10. a)Miertus, S. *FEBS Letters*, submitted for publication.
    b) Miertus, S, Bizik, F. in preparation.
11. Dauber-Osguthorpe, P., Roberts, V.A., Osguthorpe, D.J., Wolff, J., Genest,M., Hagler, A.T. *Proteins: Structure, Function & Genetics* **4**, 31-47 (1988).
12. Discover 2.7, Insight II, Biosym Technologies 1990.
13. Calflisch, A., Niederer, P., Anliker, M. *Proteins: Structure, Function & Genetics* **4**, 31-47 (1988).
14. Fletcher, R. *Practical Methods of Optimization* **1**,J. Wiley ,New York (1980).
15. Miertus, S., Scrocco, E., Tomasi, *J. Chem.Phys.* **55**, 117-129 (1981).
16. Frecer, V., Miertus, S. *Int. J. Quant. Chem.*,**42**, 339-347 (1992)
17. Miertus, S., Tomasi, *J. Chem.Phys.* **65**, 237-245 (1982).