Journal of Computer-Aided Molecular Design, 7 (1993) 291–304
ESCOM

J-CAMD 199

Peptide mimetics as enzyme inhibitors: Use of free energy perturbation calculations to evaluate isosteric replacement for amide bonds in a potent HIV protease inhibitor

Piotr Cieplak[a,b] and Peter A. Kollman[a]

[a]*Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, U.S.A.*
[b]*Quantum Chemistry Laboratory, Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland*

Received 2 August 1992
Accepted 25 November 1992

Key words: Molecular dynamics; Solvation; Free energies; Drug design

SUMMARY

We present the application of free energy perturbation theory/molecular dynamics to predict the consequence of replacing each of the seven peptide bonds in the potent HIV protease inhibitor JG365: ACE (acetyl)-Ser-Leu-Asn-HEA (hydroxyethylamine analog of Phe-Pro)-Ile-Val-NME (N-methyl) by ethylene or fluoroethylene isosteres. Replacing two of these bonds may well lead to significantly tighter binding; replacing two others is predicted to significantly diminish the binding affinity. Also, for three of the peptide bonds fluoroethylene replacements could lead to increased binding of free energies of the inhibitors. Our results should be considered as predictive since there are, as yet, no experimental results on such peptide replacements as enzyme inhibitors.

INTRODUCTION

The HIV protease cleaves specific peptide bonds in the precursor gag and pol proteins which are necessary to form mature virus particles. Thus, this protease is one of the promising targets for drug design [1].

Since the HIV protease is an aspartic proteinase, one expects that new inhibitor design can be based on the approach which was used previously in designing inhibitors for other aspartic proteinases such as pepsin, penicillopepsin, renin, cathepsin D, etc. [2,3]. Recently Rich et al. [2] reported inhibitors which bind tightly to the HIV protease. In this work it was suggested that the hydroxyethylamine (HEA) analogues of the ACE-Ser-Leu-Asn-Phe-Pro-Ile-Val-OMe peptide, with a modification of the peptide bond between Phe and Pro groups (Fig. 1), can mimic the tetrahedral intermediate for hydrolysis of that peptide bond by the HIV protease. The HEA analogues were found to bind tightly to HIV-protease. The crystal structure of such a complex with synthetic protease has become available [4].

Free energy perturbation/molecular dynamics simulations have already been applied to ana-

0920-654X/$ 10.00

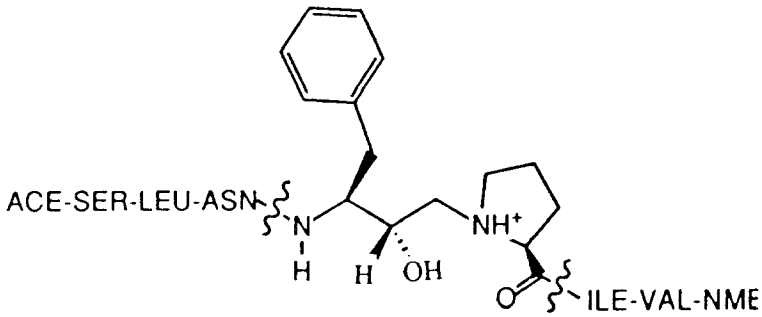


Fig. 1. Amino acid sequence in the inhibitor.

lyze which of the two diastereoisomers, R or S in the HEA -OH group, binds better to the active site of HIV protease. It was calculated that the S stereoisomer binds better by 2–4 kcal/mol than the R isomer [5,6], in good agreement with concurrent experiments which suggest a $\Delta\Delta G$ (R vs. S) of comparable magnitude [7].

In this paper we present another possible modification of the above HIV protease inhibitor. We find that replacing some peptide bonds in the inhibitor by the isosteric *trans*-ethylene or fluoroethylene units could yield an improvement in binding to the enzyme compared to the original inhibitor. Using those and other isosteric groups to replace the peptide bond has attracted some attention in theoretical and experimental studies of protein folding problems and conformational analysis [8–11]. Some experimental work has been done on the *trans*-olefinic unit incorporation into enkephalin analogues and the influence of such a substitution on biological activity has been studied [12–15]. This replacement is found to prevent premature degradation of pseudopeptide bonds in enkephalins by amino-peptidase and the increased lipophilicity of such a peptide should facilitate its passage through the blood–brain barrier and, possibly, through membranes. It was also found that *trans*-olefinic dipeptide isosteres incorporated into enkephalin and substance P analogues can possess increased biological potencies in comparison to the parent polypeptides [14]. This has motivated us to suggest a similar approach to the Rich et al. inhibitors of HIV protease. Our predictions are based on the free energy perturbation technique and molecular dynamics simulations.

METHODS

All our simulations were done using the AMBER program [17] with its all-atom force field parameters [18].

The initial structure taken for simulations was the crystal structure of the HIV-1 protease with the JG365 inhibitor bound to it [2,19]. The X-ray protein structure contains as residue 64 an ABA amino acid, which we have replaced with Cys, because Cys is the amino acid in the natural protein and the O-methyl terminal group was replaced by an N-methyl group. This was because the molecular mechanical parameters for Asn, Leu and NME are already available and, based on the structure–activity relationships [2], one expects comparable binding for these as for the Norleu and OMe analogs. In the inhibitor Norleu used experimentally was in the place of Asn and Ile in the substrate sequence presented in Ref. 19. In our case, the inhibitor sequence was based on the substrate ACE-Ser-Leu-Asn-Phe-Pro-Ile-Val-NME, and the Phe-Pro peptide bond was replaced

by the hydroxyethylamine (HEA) group, which mimics the tetrahedral transition state for the peptide bond cleavage reaction. There are seven peptide bonds in this molecule. In the following discussion we will use the standard notation for the inhibitor amino acids, i.e. P4 for Ser, P3-Leu, P2-Asn, P1-Phe, P1′-Pro, P2′-Ile, P3′-Val. The inhibitor is shown in Fig. 1. All calculations were performed for the S isomer on the carbon atom in the hydroxyethylamine group, since this isomer was found to bind better to the enzyme than the R.

We used an all-atom model to represent the protein and inhibitor. The HIV-1 protease is built of two identical units each containing 99 amino acids. Around the four charged amino acids at the surface of the protease, four additional counterions were placed. The whole system also contained 95 crystallographic water molecules and 207 H_2O molecules added by the EDIT program of AMBER within 20 Å from a point located between Asp^{25} and Asp^{125}. The added water molecules created a cap of hydration shell around the ends of inhibitor extending to about 8 Å. The simulated system contained 4176 atoms in total. The protonated form of the Pro group in the inhibitor causes the whole simulated system to have a charge of +1. At the initial stage of our calculations we used a protonated form of catalytic Asp^{25} in the enzyme, since the protonation state had not been well determined. Subsequently, the studies in Ref. 5 showed Asp^{125} protonation was the more likely state of the JG365 inhibitor-bound enzyme. Thus, for the most promising cases we also calculated the free energy of binding of the Asp^{125} protonation state. For the protonated aspartic acid form, we used the charge model described by Bash [20]. For the counterions we used a model of Cl^- anion with unit negative charge and van der Waals parameters $R^* = 2.1$ Å and $\varepsilon = 0.129$ kcal/mol.

We started our calculations with minimizing first the protease–inhibitor complex crystal structure using 500 cycles of conjugate gradient method followed by 10 ps of molecular dynamics equilibration. The inhibitor in the enzyme simulations shifted ca. 1 Å from the position in the starting X-ray structure during the equilibration stage and remained in this position during subsequent simulations [5]. This structure was used as an initial coordinate set for further molecular dynamics equilibration and/or free energy perturbation simulations. The same protocol was used in the simulations of the inhibitor in water. During the free energy perturbation molecular dynamics simulations we mutated a given peptide bond into the ethylene or fluoroethylene unit, changing gradually the appropriate atomic charges, van der Waals parameters and geometrical parameters. We noted that the net charge change in these perturbations is zero, as is clear from the data in Table 1.

We performed two sets of free energy perturbation simulations for the inhibitor–enzyme complex. A first set started from the structure prepared by the above described procedure and consisted of a separate perturbation of each peptide bond in the inhibitor into the ethylene group. Free energy perturbation molecular dynamics simulations started from the same initial structure and the standard windowing approach was applied. During 21 windows with 500 time steps of equilibration and 500 steps of data collection (averaging) using time steps equal to 0.002 ps a given peptide bond was transformed into a *trans*-ethylene group. The total time of each perturbation spanned 42 ps. We decided to calculate the free energy differences for any of the peptide bond substitutions within the inhibitor; thus, we carried out a series of seven perturbation runs for the HIV protease–inhibitor complex. In these cases we kept rigid all residues further than 8 Å away from any atom of the inhibitor. Those simulations gave us a preliminary analysis on which bond perturbations were the most promising candidates for binding improvement.

The second set of slightly longer simulations focused on those peptide bonds which were deemed most promising for the peptide bond perturbation based on the results of the initial simulations. Here the same initial structure was further equilibrated for another 40 ps. Free energy perturbation was done employing the slow growth approach and using 30 000 of 0.0015 ps steps yielding 45 ps as the total time for each run. In this second set of simulations we applied a variable type of constraint for the enzyme atoms which seemed to work satisfactorily in the previous studies [5]. The residues with all atoms further than 9 Å from any atom of the inhibitor were kept rigid. All enzyme atoms closer than this cutoff were harmonically restrained to their crystallographic position and the harmonic constraint constants increased with the distance from the inhibitor molecule. The constraint on backbone atoms of amino acid residues in each monomer chain of enzyme was as follows: 7–10, 21–23, 33–34, 45–58, 74–76, 88–90: 2.0 kcal/mol Å^2; 24–26, 31–32, 77–79, 85–87: 1.0 kcal/mol Å^2; 27–28, 30, 80–84:0.5 kcal/mol Å^2. Using the above described constraints for the enzyme atoms speeds up calculations and prevents unphysical distortion of the enzyme structure from its crystallographic form. This last approach was used for perturbing separately the third (between P3–P2) and fifth (P1′–P2′) peptide bonds into ethylene isosters and the third, fourth (between P2–P1) and fifth peptide bonds separately into a fluoroethylene bond.

The same perturbation of the peptide bonds of the inhibitor was also performed in aqueous solution in order to determine the solvent effects on $\Delta\Delta G_{bind}$ due to the various peptide bond substitutions. Also in this case, we carried out appropriate short and longer sets of simulations for the inhibitor perturbation into its isosteric forms. In the short simulations, we carried out two sets of perturbation calculations. Since we do not know the appropriate inhibitor conformation in water, we decided to start with two different initial geometries. In the first case we started with the 'native' geometry, i.e., the inhibitor conformation observed in the active site of the protease. In the second case we prepared a fully 'extended' structure with all ϕ and ψ angles equal to 180° and minimized it before adding around it a solvation shell of water molecules. The inhibitor was then placed in a periodic box of Monte Carlo equilibrated water molecules, which were not further than 9 Å from any atom of the centrally located solute. For the 'native' form of the inhibitor the box contained 873 water molecules, whereas for the 'extended' form there were 909 water mole-

TABLE 1
CHARGES ON PEPTIDE, ETHYLENE AND FLUOROETHYLENE GROUPS

Amide		Ethylene		Fluoroethylene	
Atom	Charge[a]	Atom	Charge[a]	Atom	Charge[a]
C	0.616	C1	−0.261	C1	0.273
O	−0.504	H1	0.115	F	−0.133
N	−0.463	C2	0.126	C2	−0.201
H	0.252	H2	0.029	H2	0.094
C-$^{\alpha}$	0.035	C-$^{\alpha}$	−0.073	C-$^{\alpha}$	−0.096
C if in end methyl group	0.067	C if in end methyl group	−0.041	C if in end methyl group	−0.064

[a] Charges in atomic units.

cules. After minimizing these structures, they were equilibrated by molecular dynamics simulation for 20 ps. Perturbation simulations were done employing a standard windowing method with 21 windows, 500 md steps for equilibration and 500 steps for data collection (averaging) using a 0.0015-ps time step. This yields 31.5 ps of the total time for each short perturbation of each peptide bond in water.

The longer set of simulations were only performed using the 'extended' form of the inhibitor. The above structure was further equilibrated for an additional 120 ps. The perturbation starting from such a structure lasted for 84 ps using 21 windows, 1000 md steps for equilibration, and 1000 for data collection with 0.002-ps time steps.

In all perturbation calculations, the whole inhibitor was chosen to be the perturbed group. Such an approach allows us to take into account only the dependence of free energies on intermolecular interactions and avoids the difficulties of calculating large and complicating intramolecular free energies. This approach is not ideal, but has been shown to work well in a large number of free energy calculations [5,16,20] and others.

In both cases (water and enzyme simulations) we used a dielectric constant equal to 1 and for the 1–4 van der Waals and 1–4 electrostatic intramolecular interaction energies we applied the usual scale factor of 2 [18]. The SHAKE procedure [21] was applied to constrain bond lengths to their appropriate equilibrium values for any bonds containing hydrogen atoms.

All the simulations were performed at 300 K using the TIP3P [22] water model. In the case of the inhibitor in water, periodic boundary conditions for water were used, with an 8 Å non-bonded cutoff for the intermolecular interactions.

Our initial simulations for the inhibitor–enzyme complex were done using an 8 Å cutoff, but the free energies were unexpectedly large (more than 20 kcal/mol) for modest changes in molecular structure and electrostatics. This was especially observed in the cases of the fourth (P2–P1) and fifth (P1′–P2′) bonds perturbations, which are closest to the active site. By carrying out free energy calculations using other cutoffs (8, 12, 20 Å) (not reported here), we realized that the free energies were cutoff dependent. We found that using a short cutoff, which is based on residue distance criteria, can produce an unbalanced picture of interaction energies between the charged enzyme groups and the peptide. Thus, we decided to present here the results for enzyme calculations obtained using a variable cutoff [17], i.e. 99 Å for the interaction energy of the perturbed group with any other residue and 8 Å for all other residue–residue interactions. These values seem to be the more consistent and reliable ones.

What kind of electrostatic model for the isosteric ethylene and fluoroethylene group should we use in our simulations? We performed a series of ab initio calculations using the QUEST (modified Gaussian 80) program [23] and an STO-3G basis set [24] for each pair of neighbouring amino acids present in the inhibitor connected by a *trans*-ethylene or fluoroethylene group instead of a peptide bond. Each molecule considered in the ab initio calculations contained a given pair of amino acids linked by an isosteric bond and ACE and NME standard groups at the ends. Such systems were initially optimized before the electrostatic potential in the grid points around them were calculated to fit the appropriate atomic charges [25]. We found that in the *trans*-ethylene group the charges were approximately the same at all seven positions, even on the isosteric bond close to the charged Pro residue. The largest deviation (approx. 0.1 e) from the mean value over all seven bonds is for the C_1 atom charge. For the systems with the fluoroethylene group, the largest deviation of approx. 0.2 e from the mean value was observed on the C_2 atom and around

0.1 e deviation on the C_1 atom. Thus, we decided to use the averaged charges obtained from the seven ab initio calculations performed for a given type of isostere. Since the largest inductive effect takes place on the C_α atom bonded to the N atom of the peptide group we decided to include also the charge on that atom in the set of perturbed electrostatic parameters. In all the ab initio calculations we obtained an asymmetric charge distribution on any isosteric bond, and this is reflected in the set used in our simulations. Charges for the peptide bonds and for the *trans*-ethylene and fluoroethylene groups are presented in Table 1.

RESULTS

We carried out perturbations on each of the seven peptide bonds into ethylene isosteric both in water and in the HIV protease active site. A difficulty in the aqueous simulations was what conformations to choose. As noted above, it is not feasible at present to reliably *predict* the aqueous solution conformation of a complex peptide, assuming a well-defined structure actually exists. Thus, we began the trajectory with the peptide in both extended (ϕ, ψ = 180°) and X-ray bound conformations, and after equilibration, carried out the perturbation. Table 2 contains the results from those two types of calculations as well as the arithmetic mean values for each bond.

We also carried out a model calculation on the alanyl dipeptide in an extended form and found that mutating either of its peptide bonds into ethylene in water led to a ΔG (solvation) of ~8 kcal/mol, in qualitative agreement with gas phase solvation data which suggests a ΔG (solvation) of ~11 kcal/mol [26]. In a dipeptide, peptide bonds are well exposed to the water and, therefore, the above value could be treated as a reference for comparison with the results of perturbation in other peptides. The smaller the value of ΔG (solvation), the more shielded from water environment a given peptide bond in the starting structure could be. The range of the free energies indicates that the nature of solvation of the peptide bond is quite sensitive to its position in the chain and its environment. In one case, the fourth bond perturbation (between P2–P1), we can see that the solvation free energy change could even have a negative value. Of course, this result may be an artifact of our choice of the entire peptide as the perturbed group and our limited sampling of conformational space. In addition we must consider whether a 'hydrophobic' substitution in a conformationally complex molecule may not necessarily yield a more hydrophobic molecule. The

TABLE 2
RESULTS OF THE FREE ENERGY DIFFERENCES (ΔG_i) FOR THE PERTURBATION OF ONE PEPTIDE BOND INTO THE ISOSTERIC ETHYLENE GROUP IN THE INHIBITOR IN WATER

Bond no.	Between residues	'Native'	'Extended'	Mean
1	ACE–P4	5.8 ± 0.1	4.3 ± 0.1	5.1
2	P4–P3	1.9 ± 0.1	−0.5 ± 0.1	0.7
3	P3–P2	2.5 ± 0.2	0.6 ± 0.2	1.6
4	P2–P1	−4.3 ± 0.1	−0.6 ± 0.2	−2.5
5	P1′–P2′	11.4 ± 0.3	9.0 ± 0.1	10.2
6	P2′–P3′	4.1 ± 0.1	5.4 ± 0.2	4.8
7	P3′–NME	6.6 ± 0.3	4.8 ± 0.1	5.7

Short simulations. Results in kcal/mol.

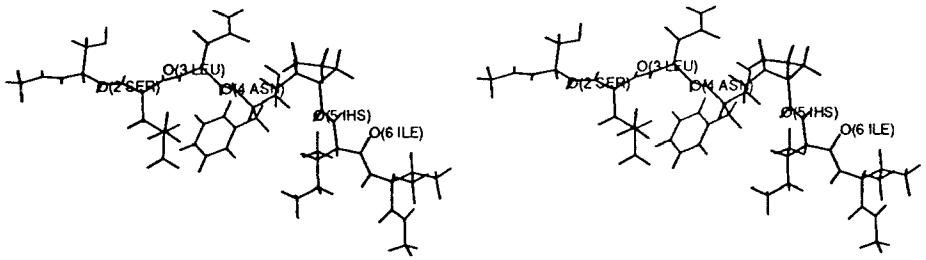


Fig. 2. Stereoview of the minimized 'extended' inhibitor structure taken for molecular dynamics simulation in water.

second (P4–P3), third (P3–P2) and fourth (P2–P1) peptide bonds are the most shielded against solvent access, mostly because of the proximity of the hydrophobic Leu and Phe side chains. That effect can be easily observed in Figs. 2 and 3, which show initial and short equilibrated 'extended' forms of inhibitor in water, respectively.

Examining Table 2, we can state that although the numbers differ quantitatively depending on conformation, it is encouraging that the trend of relative solvation free energy is generally similar for both 'native' and 'extended' conformations.

A general and unexpected result here is that changing a peptide group to an ethylene isostere, if the side chains on either side are hydrophobic, could lead to a much smaller decrease in water solubility than expected from the additivity of this effect. This is because the two side chains and the ethylene form a 'hydrophobic core' and reduce solvent exposure. When the hydrophobic side chains straddle the more polar peptide bond, this cannot happen.

The results of free energy perturbation/molecular dynamics simulations on the inhibitor–protein complex are collected in Table 3. The errors given in Tables 2 and 3 are the so-called hysteresis values for 'forward' and 'backward' statistics. The weakest interactions of the peptide group with enzyme is observed for the third bond (between P3–P2), which is reflected in the smallest value of the free energy changes for removing such a bond and replacing it by the more hydrophobic ethylene group. As noted above, a contribution to that effect is likely due to the shielding effect of the Leu and Phe side chains. This was also confirmed by the initial analysis of the distances between the peptide bonds in the inhibitor and the rest of the enzyme in the minimized structure, which suggested that the weakest attractions of the inhibitor with the protein were likely to be at the third (P3–P2) and fifth (P1′–P2′) bonds. Thus, not only the third but also the fifth peptide bond was expected to be a good candidate for the replacements.

To fully interpret our results with respect to the differences in binding effects we must use the

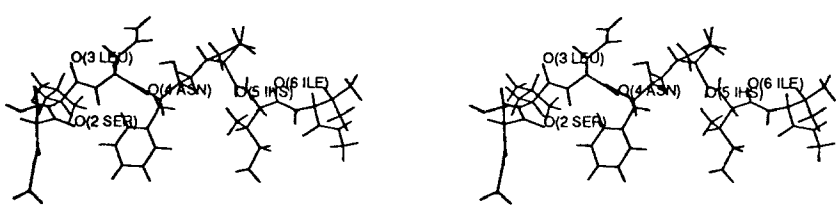
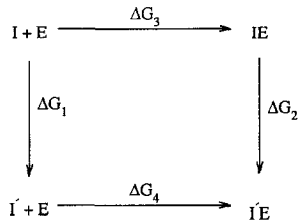


Fig. 3. Stereoview showing inhibitor structure after short equilibration in water.

following thermodynamic cycle:



The ΔΔGs which we can obtain from the simulations are ΔG_1 and ΔG_2. The most interesting is the free energy difference of binding between inhibitors I and I′, where I is the peptide and I′ the ethylene derivative isostere. $\Delta\Delta G_{bind} = \Delta G_4 - \Delta G_3 = \Delta G_2 - \Delta G_1$, which determines the relative binding constant of the two inhibitors to the HIV-1 protease, with a negative $\Delta\Delta G_{bind}$ indicating that the ethylene isostere should bind more strongly to the enzyme. The results based on the shorter simulations are presented in Table 4. The errors given there are calculated as the differences between results from the appropriate lowest and highest values of ΔG_1 from Table 2.

Table 4 shows that replacing the third (P3–P2) or fifth (P1′–P2′) peptide bond by the *trans*-ethylene group in the inhibitor may substantially increase the binding free energy to the enzyme. Furthermore, one can expect, based on the results from Table 4, that the mutation of peptide bonds 2 (P4–P3) and 4 (P2–P1) will be substantially unfavorable.

We found that to properly assess the above results from our simulations we had to look closer into the role of the crystallographic water buried in the enzyme. One of the most important is the tetrahedrally ligated water molecule, which H-bonds to the N–H bonds of residues 50 and 150 of the flaps and the C=O groups of the inhibitor (amide bonds #4 - P2–P1 and #5 - P1′–P2′). Figure 4 shows a view of this water and its environment after 10 ps of molecular dynamics equilibration. As one can see, the qualitative features of the crystal structure have remained, but the hydrogen bond with the C=O of the fifth bond has been significantly weakened. This could be a key factor as to why the $\Delta\Delta G_{bind}$ for bond 5 (P1′- P2′) is so favorable for amide bond replacement, and the strong H-bond still formed between the water and the C=O of bond 4 (P2–P1) could be a reason why the $\Delta\Delta G_{bind}$ for bond 4 is so positive.

TABLE 3
RESULTS OF THE FREE ENERGY DIFFERENCES (ΔG_2) FOR THE PERTURBATION OF ONE PEPTIDE BOND INTO THE ISOSTERIC ETHYLENE GROUP IN THE INHIBITOR IN THE COMPLEX WITH PROTEASE

Bond no.	Between residues	ΔG_2
1	ACE–P4	4.0 ± 0.1
2	P4–P3	4.5 ± 0.1
3	P3–P2	−0.5 ± 0.02
4	P2–P1	7.9 ± 0.2
5	P1′–P2′	4.3 ± 0.1
6	P2′–P3′	4.2 ± 0.5
7	P3′–NME	5.8 ± 0.6

Infinite cutoff for inhibitor–enzyme interactions. Short simulations. ASP25 protonated. Results in kcal/mol.

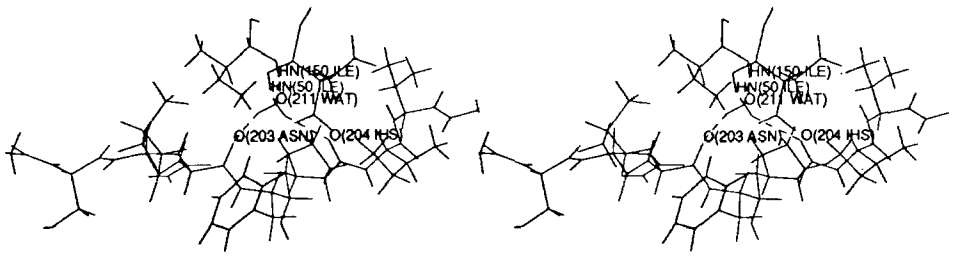


Fig. 4. Stereoview showing the key water (residue 211) in the HIV structure that hydrogen bonds to both residues 50 and 150 of the protease and to the C=O bonds at each end of the HEA residue. This is the structure after 10 ps molecular dynamics equilibration.

Attempting to see how accurate the $\Delta\Delta G_{bind}$ calculated with the short simulations were, we carried out additional, longer simulations in water and in the enzyme after additional equilibration, as described in the Methods section. We also investigated the influence of the enzyme protonation state in the active site. This latter protocol of simulation was also considered for fluoroethylene substitution for the third (P3–P2), fourth (P2–P1) and fifth (P1′-P2′) bonds.

For the fifth bond (P1′–P2′) mutation we applied several additional distance constraints for the four H-bonds which the above mentioned crystallographic water forms with the enzyme and the peptide bond atoms of the inhibitor. In such a case we used constraint force constants of 0 or 20 kcal/mol Å for the two H-bonds between the water oxygen and the N–H atoms of residues 50(Ile) and 150(Ile), and between the C=O groups of the inhibitor (amide bonds #4 -P2–P1 and #5 - P1′–P2′) and the respective water hydrogens. The equilibrium distances for constraints were kept at 1.99, 2.68, 2.04 and 2.30 Å, respectively and were assumed based on an initially minimized structure. After an additional equilibration of the structure with all four hydrogen bonds constrained, we performed three types of perturbation simulations in which (a) all four hydrogen bonds were kept constrained, (b) the hydrogen bond between one water hydrogen and the amide C=O group of bond #5 - P1′–P2′ was unconstrained, and (c) all four bonds were unconstrained. The results are collected in Tables 5 and 6 and the position of the crystallographic water after constrained equilibration is shown in Fig. 5.

Inspection of Table 5 reveals that the calculated free energies for P1′-P2′ (the fifth peptide

TABLE 4
CALCULATED BINDING FREE ENERGIES DIFFERENCES $\Delta\Delta G_{bind} = \Delta G_2 - \Delta G_1$ BETWEEN ETHYLENE ISOSTERIC GROUP AND NORMAL PEPTIDE BOND CONTAINING INHIBITORS

Bond no.	Between residues	$\Delta\Delta G_{bind}$ relative to mean value of ΔG_1
1	ACE–P4	−1.1 ± 0.8
2	P4–P3	3.8 ± 1.2
3	P3–P2	−2.1 ± 1.0
4	P2–P1	10.4 ± 1.9
5	P1′–P2′	−5.9 ± 1.2
6	P2′–P3′	−0.6 ± 0.7
7	P3′–NME	0.1 ± 0.9

ASP25 protonated. Results in kcal/mol.

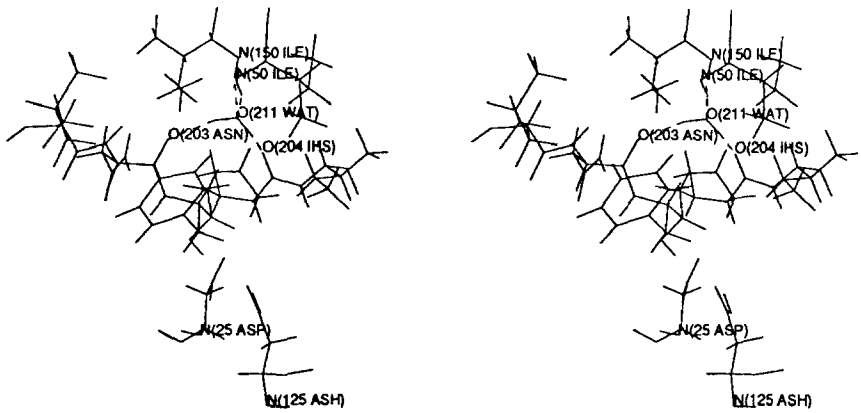


Fig. 5. Stereoview showing the water (residue 211) in the HIV structure that hydrogen bonds to both residues 50 and 150 of the protease and to the C=O bonds at each end of the HEA residue. This is the structure after molecular dynamics equilibration with additional distance constraints.

bond) substitution by the ethylene isostere unit could substantially depend on the protocol of controlling the H-bonds between the crystallographic water and the inhibitor, but in any case the final results for the $\Delta\Delta G_{bind}$ are negative, which means that substituting this peptide bond for its ethylene mimetics could lead to an improved inhibition effect of the enzyme. We did not include in Table 5 (and 6) the values of errors for ΔG_2, since in any case the hysteresis was less than 0.1

TABLE 5
RESULTS OF THE FREE ENERGY SIMULATIONS FOR THE PERTURBATION OF ONE PEPTIDE BOND INTO THE ISOSTERIC ETHYLENE, –CH=CH–, GROUP IN THE INHIBITOR IN WATER, IN PROTEASE AND BINDING FREE ENERGIES

			Simulation in protease, variable constraints		Difference in binding
Bond no.	Between residues	Simulation in water ΔG_1	Protonation state	ΔG_2	$\Delta\Delta G_{bind} = \Delta G_2 - \Delta G_1$
3	P3–P2	0.56 ± 0.07	Asp^{25}	−0.83	−1.4
			Asp^{125}	0.56	0.0
5	P1′–P2′	11.84 ± 0.05	Asp^{25a}	10.6	−1.2
			Asp^{25b}	6.8	−5.0
			Asp^{25c}	6.9	−4.9
			Asp^{125a}	9.5	−2.3
			Asp^{125b}	9.0	−2.8
			Asp^{125c}	11.4	−0.4

Long simulations (see text). Results in kcal/mol.
[a] Four constraints on H-bonds, crystallographic water applied.
[b] Three H-bonds constrained (see text).
[c] No constraints on H-bonds used.

kcal/mol. Taking into account the H-bonds constraints led to results which are not qualitatively different from those obtained in shorter simulations. Another interesting observation which could be made here is that the result obtained for the free energy differences of hydration, (ΔG_1), between peptide and isosteric P3–P2 bond from longer simulations converged to the result obtained from a previous shorter run for the 'extended' form. On the other hand the longer simulations for the fifth bond (P1′–P2′) in water gave a closer value to that obtained for the 'native' structure in short calculations.

In Table 6 the results for free energy simulations for the perturbation of the peptide bond into the fluoroethylene group are shown. The best candidate for this type of replacement is clearly the third (P3–P2) bond, independent of the protonation state of the enzyme.

Our shorter (Table 2) and longer (Table 5) simulations in water show that the inhibitor in water can have quite a different structure than in the enzyme. Also after longer equilibration and subsequent perturbation of the peptide bond into its isostere the inhibitor undergoes further conformational changes. The stereoview of the inhibitor after the long 120 ps equilibration in water is shown in Fig. 6. The biggest conformational changes between the short and long equilibrated structures in water are observed for the ϕ angle between Leu-Asn (P3–P2) ($\Delta\phi \sim 175°$) and the ψ angle between P1′–P2′ ($\Delta\psi \sim 75°$) and P2′–P3′ ($\Delta\psi \sim 140°$) residues. The greatest changes in the χ_1 and χ_2 angles are observed for Leu (res. P3) ($\Delta\chi_1 \sim 32°$, $\Delta\chi_2 \sim 130°$) and Asn (res. P2) ($\Delta\chi_2 \sim 120°$). During perturbation of any peptide bond, the Leu (P3) and Asn (P2) residues undergo quite strong conformational changes where $\Delta\chi_1$ can change up to 95° and $\Delta\chi_2$ up to 115° in the first case, and $\Delta\chi_2$ can vary between 80 and 150° in the second case.

DISCUSSION AND CONCLUSIONS

We have presented free energy calculations on an HIV protease inhibitor, in which the peptide bonds of the inhibitor have been mutated into ethylene or fluoroethylene isosteres. By carrying

TABLE 6
RESULTS OF THE FREE ENERGY SIMULATIONS FOR THE PERTURBATION OF ONE PEPTIDE BOND INTO THE ISOSTERIC FLUOROETHYLENE, –CF=CH–, GROUP IN THE INHIBITOR IN WATER, IN PROTEASE AND BINDING FREE ENERGIES.

			Simulation in protease, variable constraints		Difference in binding
Bond no.	Between residues	Simulation in water ΔG_1	Protonation state	ΔG_2	$\Delta\Delta G_{bind} = \Delta G_2 - \Delta G_1$
3	P3–P2	4.15 ± 0.04	Asp^{25}	2.26	−1.9
			Asp^{125}	1.83	−2.3
4	P2–P1	5.50 ± 0.03	Asp^{25}	4.37	−1.1
			Asp^{125}	5.78	0.3
5	P1′–P2′	4.83 ± 0.04	Asp^{25}	4.84	0.01
			Asp^{125}	3.60	−1.2

Long simulations (see text). Results in kcal/mol.

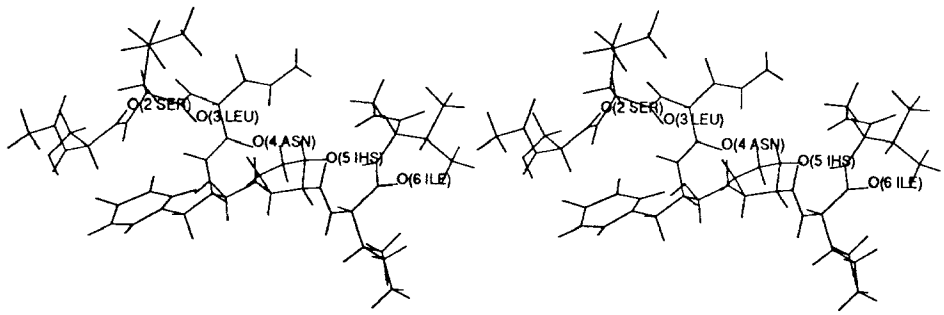


Fig. 6. Stereoview of the inhibitor structure after long equilibration in water.

out the same perturbation in water and in the enzyme active site, we are in a position to suggest peptide bond modifications which are less unfavorable in the enzyme than in water. The conceptual idea is, given the significantly greater polarity of the peptide bond than the alkene double bond, the peptide will have a more favorable interaction free energy in almost any environment. However, the key is whether the environment around a given peptide bond is more favorable in water or in the active site. If the former is the case, the replacement of the peptide bond by an ethylene or fluoroethylene isostere could well lead to a better inhibitor.

We would like to stress that in our study we considered two types of potential modification of peptide inhibitors. One of these is where the peptide bond is replaced by an ethylene unit, which has a small bond dipole moment. In the other, in which the peptide bond is replaced by fluoroethylene, the bond dipole orientation is preserved, but decreased in magnitude. Such an approach could serve as a probe to see in which direction one has to modify a given inhibitor to improve it for a given environment. Our simulations suggest that this concept could be very useful in inhibitor design in many areas.

In our study here, we suggest substitution of at least two positions which could yield better inhibitors of HIV protease. The two bonds, between P3–P2 and P1′–P2′, could be substituted by a fluoroethylene or ethylene unit. The results of the final values of the $\Delta\Delta G_{bind}$ are quite sensitive to the simulation protocol used. Uncertainties in our calculated binding free energies are relatively large, probably greater than 1 kcal/mol, and are caused by a number of factors, which preclude the possibility of more precise estimations. Nevertheless, in the cases of the third (P3–P2) and fifth (P1′–P2′) peptide bonds, fluoroethylene and ethylene substitutions lead to negative values for $\Delta\Delta G_{bind}$ (Tables 5 and 6) in most simulation conditions. Such substitution can have additive effects since the distance between these bonds is large. As a bonus, such an inhibitor could be potentially more metabolically resistant, and membrane and blood–barrier permeable.

There are a number of caveats in the application of the free energy methods, mainly inaccuracies in molecular mechanical parameters and shortness of simulation time/limited conformational sampling. For example, the use of STO-3G electrostatic potential charges has led to the underestimation of the free energy of amide → ethylene in solution noted above. We expect, however, that there may be some cancellations of these 'molecular mechanical parameter' errors because one is analyzing the difference between ΔG in solution and in the enzyme.

A referee of this manuscript has noted that the simulation times used are 'very short' and we agree, but we argue, not too short to calculate semi-quantitatively free energies that are *electro-*

statically dominated as shown by Bash et al. [16]. The point we wish to emphasize is that we are estimating the $\Delta\Delta G$ for the amide to (fluoro)ethylene perturbation assuming that the conformation in the protein does not change substantially and that either the protein bound or fully extended conformation is *representative* of both the amide and (fluoro)ethylene peptide in solution. We cannot use the elegant approach for analyzing the conformational free energy described by Straatsma and McCammon [27], because it is computationally intractable at present to completely span conformational space for the peptide, the seven fluoroethylene peptides and the seven ethylene peptides in solution. Even for 18-crown-6 itself, without considering a chemical mutation, Straatsma and McCammon found that 1.5 ns was not sufficient to fully sample conformational space of this molecule. Similarly, Hermans et al. [28] have shown how to determine conformational free energies for simple peptides, but these calculations involve at most one or two variable dihedral angles. In our study we have far more conformational degrees of freedom and we have taken the different approach noted above; to semi-quantitatively estimate the $\Delta\Delta G$ based on assumptions about the conformations noted above. A test of the reasonableness of our assumptions must come from experimentally synthesizing these molecules and testing their binding to HIV protease. The technology to do so has been described [12–15].

Another approximation used here is the assumption, which was questioned by referees, that the $\Delta\Delta G_{bind}$ is due only to differences in the intermolecular interactions between the peptide and either the protein or water. As noted above, this approximation has been shown to be a good one in many cases [5,16,20], but most of those cases involve rather small molecules. We expect that this contribution to the free energy will be small when we calculate the $\Delta\Delta G_{bind}$ based on using the 'bound' conformation in both protein and solution, but it will not necessarily be small if we use the 'bound' conformation in the protein and the fully extended one in solution. Future studies will be required to fully sort out its contributions, but we emphasize that we only 'flag' the $\Delta\Delta G_{bind}$ values that are calculated to be negative, independent of the solution conformation.

Although our results still need experimental confirmation, we suggest that our qualitative conclusion, the wide range of $\Delta\Delta G_{bind}$ for peptides $\rightarrow$ ethylene derivatives isosteres, will be true in many protein–inhibitor interactions and thus that calculations such as these will be useful predictions for the design of new, more metabolically stable, tighter binding enzyme inhibitors with better distributional properties than peptides.

We also stress that the methods employed here can be useful in the design of other peptide inhibitor analogues, e.g. α-methyl, retro-inverso, thioamides. The key feature of all these studies is to begin with a known, tight binding peptide inhibitor and to examine mutations that will change hydrophilic groups into more metabolically stable hydrophobic groups where it can be done with minimal cost of favorable protein–inhibitor interactions. Unfortunately, the binding process itself is too complicated to be described in computer simulations, thus we do not have any assurance that the structure proposed as modified inhibitor would bind properly. In conclusion, we feel that peptidomimetics could be very useful in pharmacology and could gain much more interest in the near future in the field of rational drug design [29].

ACKNOWLEDGEMENTS

We are glad to acknowledge research support of the NIH (GM-29072 to P.A.K., GM-39552 to G.L. Kenyon and DARPA MDA-91-J-1013 to F.E. Cohen). P.C. is partially supported by the

Polish Committee for Scientific Research (KBN-CHEM-BST 492/23). Some of the calculations were carried out at the Pittsburgh and San Diego Supercomputer Centers through supercomputer supports provided to P.A.K. We also acknowledge the UCSF Computer Graphics Laboratory (supported by RR-1081 to R. Langridge), Graham Smith and Kate Holloway of Merck and Co. for useful comments, and Randall Radmer for aid in setting up the protein–inhibitor simulation.

REFERENCES

1 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.J. and Machin, P.J., Science, 248 (1990) 358.
2 Rich, D.H., Green, J., Toth, M.V., Marshall, G. and Kent, S.B.H., J. Med. Chem., 33 (1990) 1285.
3 James, M.N.G. and Sielecki, A.R., In Jurnak, F.A and McPherson, A. (Eds.) Biological Macromolecules and Assemblies, Vol. 3, Wiley, New York, 1987, p. 414.
4 a. Miller, M., Swein, A.L., Jaskolski, M., Sathyanarayana, B.K., Marshall, G.R., Rich, D.H., Kent, S.B.H. and Wlodawer, A., In Pearl, L. (Ed.) Retroviral Proteases: Control of Maturation and Morphogenesis, MacMillan Press, New York, 1992, in press.
b. Swein, A.L., Miller, M.M., Green, J., Rich, D.H., Schneider, J., Kent, S. and Wlodawer, A., Proc. Natl. Acad. Sci. U.S.A., 87 (1990) 8805.
c. Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Kent, S.B.H. and Wlodawer, A., Science, 246 (1989) 1149.
5 Ferguson, D.M., Radmer, R.J. and Kollman, P.A., J. Med. Chem., 34 (1991) 2654.
6 Tropsha, A. and Hermans, J., Prot. Eng., 5 (1992) 29.
7 Rich, D.H., Sun, C.Q., Vara Prasad, J.V.N., Pathiasseril, A., Toth, M.V., Marshall, G.R., Clare, M., Mueller, R.A and Houseman, K., J. Med. Chem., 34 (1991) 1222.
8 Aubry, A. and Marraud, M., Biopolymers, 28 (1989) 109.
9 Baginski, M., Piela, L. and Skolnick, J., J. Comput. Chem., 14 (1993) 296 and 314.
10 Roberts, D.C. and Velaccio, F., The Peptides, Vol. 5, Academic Press, New York, 1983 pp. 341–449.
11 Schnolzer, M. and Kent, S.B.H., Science, 256 (1992) 221.
12 Hann, M.M. and Sammes, P.G., J. Chem. Soc. Chem. Commun., (1980) 234.
13 Cox, M.T., Heaton, D.W. and Horbury, J., J. Chem. Soc. Chem. Commun., (1980) 799.
14 Cox, M.T., Gormley, J.J., Hayward, C.F. and Petter, N.N., J. Chem. Soc. Chem. Commun., (1980) 800.
15 Hann, M.M., Sammes, P.G., Kennewell, P.D. and Taylor, J.B., J. Chem. Soc. Perkin I (1982) 307.
16 Bash, P.A., Singh, U.C., Langridge, R. and Kollman, P.A., Science, 236 (1987) 564.
17 AMBER version 4.0 (1991), Pearlman, D.A., Case, D.A., Caldwell, J., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A., Department of Pharmaceutical Chemistry, University of California, San Francisco., modified to allow variable cutoff for particular residues.
18 Weiner, S.J., Kollman, P.A., Nguyen, D.T and Case, D., J. Comput. Chem., 7 (1986) 230.
19 Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J. and Kent, S.B.H., Science, 245 (1989) 616.
20 Bash, P., Ph.D. thesis, University of California San Francisco, 1986.
21 Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C., J. Comput. Phys., 23 (1977) 327.
22 Jorgensen, W., Chandresekhar, J., Madura, J., Impey, R.. and Klein, M. J. Chem. Phys., 79 (1983) 926.
23 Singh, U.C. and Kollman, P., J. Comput. Chem., 7 (1986) 718.
24 Hehre, W., Stewart, R.F. and Pople, J.A., J. Chem. Phys., 51 (1969) 2657.
25 Singh, U.C. and Kollman, P.A., J. Comput. Chem., 5 (1984) 129.
26 Wolfenden, R., Science, 222 (1983) 1087.
27 Straatsma T.P. and McCammon, J.A., J. Chem. Phys., 91 (1989) 3631.
28 Hermans, J., Anderson, A.G. and Yun, R.H., Biochemistry, 31 (1992) 5646.
29 Roques, B.P., Biopolymers, 32 (1992) 407.