Journal of Computer-Aided Molecular Design, 10 (1996) 23–30
ESCOM

J-CAMD 322

Calculation of solvation and binding free energy differences between VX-478 and its analogs by free energy perturbation and AMSOL methods

B.G. Rao*, E.E. Kim and M.A. Murcko

Vertex Pharmaceuticals Inc., 40 Allston Street, Cambridge, MA 02139, U.S.A.

Received 4 August 1995
Accepted 2 November 1995

Keywords: AMBER; HIV-1 protease inhibitor; *N,N*-dialkyl benzene sulfonamides; Force field; Hydrogen bonding; Structure-based drug design; Continuum solvation methods

Summary

VX-478 belongs to a novel class of HIV-1 protease inhibitors that are based on *N,N*-disubstituted benzene sulfonamides. Force field parameters for the *N,N*-dialkyl benzene sulfonamide moiety have been assembled from the literature and from our own ab initio calculations. These parameters were employed to calculate solvation and binding free energy differences between VX-478 and two analogs. The free energy perturbation method has been used to determine these differences using two approaches. In the first approach, intergroup interaction terms only were included in the calculation of free energies (as in most reports of free energy calculations using AMBER). In the second approach, both the inter- and intragroup interaction terms were included. The results obtained with the two approaches are in excellent agreement with each other and are also in close agreement with the experimental results. The solvation free energies of *N,N*-dimethyl benzene sulfonamide derivatives (truncated models of the inhibitors), calculated using continuum solvation (AMSOL) methods, are found to be in qualitative agreement with the experimental and free energy perturbation results. The binding and solvation free energy results are discussed in the context of structure-based drug design to show how physicochemical properties (for example aqueous solubilities and bioavailabilities) of these HIV-1 protease inhibitors were improved, while maintaining their inhibitory potency.

Introduction

VX-478 [(3*S*)-tetrahydro-3-furyl *N*-((1*S*,2*R*)-3-(4-amino-*N*-isobutylbenzenesulfonamido)-1-benzyl-2-hydroxypropyl) carbamate] is a potent, low-molecular-weight and orally bioavailable inhibitor of HIV-1 protease that is currently undergoing Phase I clinical trials for treatment of AIDS [1,2]. VX-478 (compound **I** in Fig. 1) emerged from a focused program of structure-based drug design that was pursued to reduce inhibitor size and enhance solubility, while maintaining subnanomolar in vitro inhibitory potency [2]. The crystal structure of this inhibitor complexed with HIV-1 protease at high resolution has been reported recently [1], illustrating the factors behind its high potency against HIV-1 protease. Furthermore, the structure also showed that the aniline (P2') group of the inhibitor is positioned in the S2' subsite of the enzyme active site. This subsite consists of hydrophobic side chains such as Ile^{50}, Ile^{84}, $Ala^{28'}$, $Ile^{32'}$ and $Val^{47'}$, and hydrophilic groups including the main-chain amides of $Asp^{29'}$ and $Asp^{30'}$ and the side chain of $Asp^{30'}$. This particular positioning of the aromatic ring in the S2' pocket in all the structures of our sulfonamide-based inhibitors with HIV-1 protease [1,3; E.E. Kim, unpublished results], close to both hydrophobic and hydrophilic groups, provided an excellent opportunity for the structure-based drug design approach [4]. Specifically, we wanted to test a variety of modifications of this novel class of inhibitors for enhancing solubility without adding obligate charge, while maintaining inhibitory potency without increasing molecular weight [2]. For this report, we have chosen two analogs of VX-478 (compounds **II** and **III** in Fig. 1),

*To whom correspondence should be addressed.

0920-654X/$ 6.00 + 1.00 © 1996 ESCOM Science Publishers B.V.

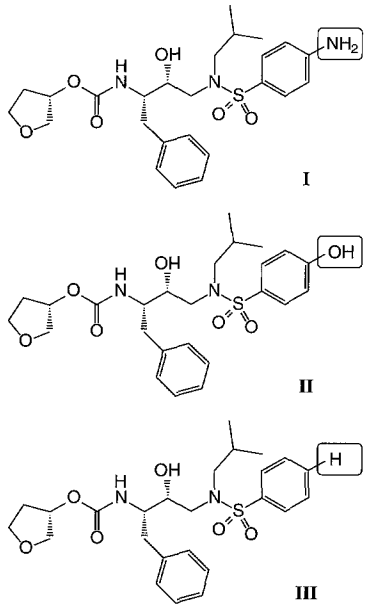


Fig. 1. Structures of VX-478 (**I**) and two of its analogs. The para-substituents on the P2' aromatic ring, which are mutated during free energy perturbation calculations, are boxed.

which illustrate this point. Starting from compound **III** ($K_i = 0.5$) without any substituent on the P2' benzene ring, we obtained the 4-NH_2 analog, **I** ($K_i = 0.6$) and the 4-OH analog, **II** ($K_i = 0.1$) with similar potencies, but with much higher aqueous solubilities [1,2; S. Pazhanisamy et al., unpublished results]. In the crystal structure of **I** (VX-478) [1], the 4-NH_2 substituent on the P2' ring makes a strong hydrogen bond with the side-chain oxygen of $Asp^{30'}$ (2.8 Å). This hydrogen bond does not seem to help binding, as its potency is comparable to that of its analog, **III**. The 4-hydroxy analog, **II**, is about sixfold more potent. This hydroxyl group is likely to make three hydrogen bonds; two with the main-chain NH of $Asp^{29'}$ and $Asp^{30'}$ and one with either the side-chain oxygen or the main-chain carbonyl of $Asp^{30'}$. In a different class of inhibitors, a hydroxyl group attached to the P2' group contributes an increase in binding of about 60-fold [5]. The smaller increase in binding of **II** over **III** appears to suggest a weaker hydrogen-bond interaction between the hydroxyl of **II** and the enzyme active site. The observation of an even smaller contribution to binding by the NH_2 group of **I** appears to suggest that the hydrogen bond it forms with the $Asp^{30'}$ side chain is not as strong as indicated by the crystal structure [1]. Moreover, **I** is much more soluble than most of its analogs [1,2] and its binding is likely to be affected by a higher cost of desolvation. We report here a quantitative understanding of the relative contributions of hydrogen bonding and solvation effects to the binding of the three inhibitors to HIV-1 protease, obtained by computational studies such as molecular dynamics and the free energy perturbation (FEP) [6,7] and continuum solvation [8,9] methods. The FEP method was applied to calculate both solvation free energy differences and binding free energy differences between the three analogs by means of two mutations of the para-substituent of the P2' benzene ring: (1) $NH_2 \rightarrow H$ and (2) $OH \rightarrow H$. The continuum solvation methods were employed to calculate solvation free energies of truncated model compounds of the three inhibitors, which were compared with the experimental and free energy perturbation results.

The FEP method has been applied to a variety of systems [6,7], including inhibitor binding to HIV-1 protease [10]. Most of the studies on HIV-1 protease inhibitors addressed *S* to *R* (or *R* to *S*) inversion of the central hydroxyl group of the inhibitor [10,11], except for a few studies that perturbed main-chain [12] or larger groups [13,14]. Solvation and binding free energy differences ($\Delta\Delta G_{sol}$ and $\Delta\Delta G_{bind}$) are calculated by the FEP method according to the thermodynamic cycles shown in Fig. 2. In most studies, however, only the right-side cycle in Fig. 2 was employed to obtain $\Delta\Delta G_{bind}$. In studies where the left-side cycle was employed to calculate $\Delta\Delta G_{sol}$, no experimental data were provided for comparison. Furthermore, some of these studies included only intermolecular interactions of the perturbed group, while other studies included both inter- and intramolecular interactions for the calculation of free energy differences. In using the first approach, it is assumed that intragroup contributions to free energy are similar in the gas phase and in solution, but this assumption has not been verified [15] by direct comparison of results obtained with the two approaches. In this study, we have addressed these issues by calculating free energy differences using the two approaches and these results are compared with the results obtained by continuum solvation methods and experiments.

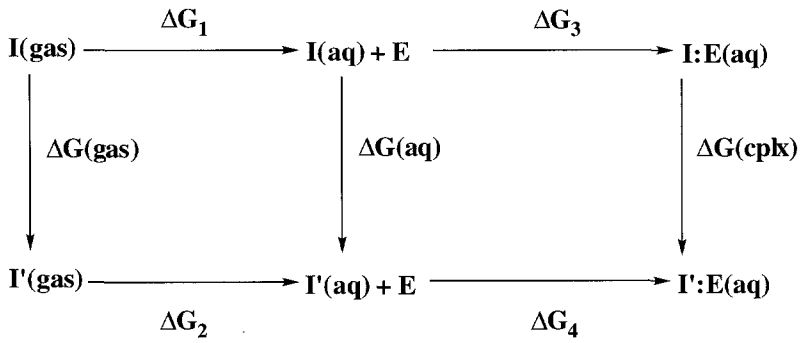
Methods

AMBER (v. 3.3) [16] was used for all molecular mechanics minimizations, molecular dynamics simulations and

TABLE 1
BOND AND ANGLE PARAMETERS FOR THE BENZENE SULFONAMIDE GROUP

Bond / angle	k (kcal/mol)	Length (Å) / angle (°)
CT-NS	337.0	1.47
NS-SO_2	314.0	1.62
CA-SO_2	222.0	1.77
O-SO_2	628.0	1.43
CT-NS-CT	50.0	115.1
O-SO_2-O	118.0	119.7
NS-SO_2-O	78.0	107.0
NS-SO_2-CA	70.0	106.8
CA-SO_2-O	69.0	107.8
CT-NS-SO_2	95.0	117.4
SO_2-CA-CA	85.0	120.0

k is the force constant.



$$\Delta\Delta G_{sol} = \Delta G_2 - \Delta G_1 = \Delta G(aq) - \Delta G(gas)$$
$$\Delta\Delta G_{bind} = \Delta G_4 - \Delta G_2 = \Delta G(cplx) - \Delta G(aq)$$

Fig. 2. The thermodynamic cycles used in this study to compute solvation (left cycle) and binding (right cycle) free energy differences. The vertical arrows indicate mutations carried out by free energy simulations. Inhibitor I is mutated to I' in the gas phase (the first arrow), in a water box (the second arrow) and in a solvated box of enzyme–inhibitor complex (the third arrow) to compute ΔG(gas), ΔG(aq) and ΔG(cplx) values, respectively. $\Delta\Delta G_{sol}$ and $\Delta\Delta G_{bind}$ values are computed according to the two equations given in the lower part.

free energy calculations. A united atom representation of the AMBER force field [17] was used for all enzyme residues. Water molecules were modeled by TIP3P parameters [18]. Inhibitors were modeled by an all-atom representation and the appropriate parameters were determined as follows. Backbone torsion angles for the sulfonamide portion of VX-478 observed in the crystal structure of the complex are close to the average value of 10 related entries without ortho-substituents that were found in the Cambridge Structural Databank [1]. We have used the mean of the bond lengths and angles of these 10 crystal structures to arrive at the valence parameters in Table 1. The force constants for bond stretching were mostly taken from the CHARMm parameter handbook [19]. We have performed extensive ab initio calculations on *N,N*-dimethyl benzene sulfonamide to characterize the conformational behavior of the sulfonamide moiety. These studies (the details are to be published separately) show

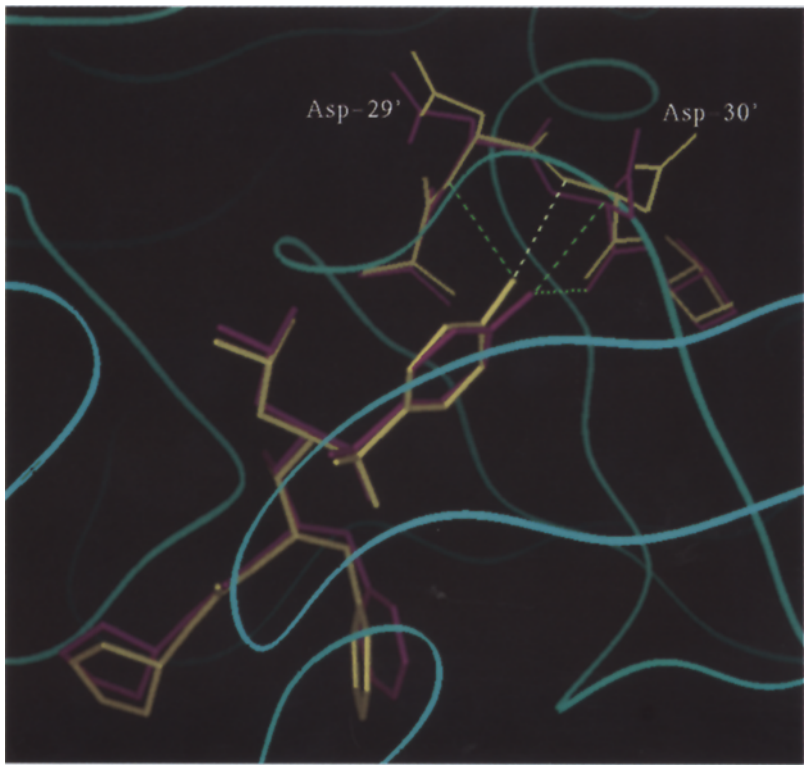


Fig. 3. The overlap of the minimized HIV-1 protease complex structures of compound **I** (VX-478, colored purple) and **II** (the hydroxy analog, colored yellow). The two structures overlap with an rms deviation of 0.77 Å using all enzyme atoms. For clarity reasons, the enzyme is shown as a blue colored ribbon and only heavy atoms of the two inhibitors and two enzyme active-site residues, $Asp^{29'}$ and $Asp^{30'}$, are included. The dashed green lines indicate hydrogen bonds discussed in the text.

TABLE 2
DIHEDRAL PARAMETERS FOR THE BENZENE SULFONAMIDE GROUP

Dihedral angle	Multiplicity	$V_{1/2}$ (kcal/mol)	Theta (°)	Phase (°)
X-CT-NS-X	6	0.9	0.0	3.0
X-CA-SO_2-X	6	1.5	180.0	2.0
X-NS-SO_2-X	5	0.0	0.0	2.0
CT-NS-SO_2-CA	1	3.4	0.0	2.0
SO_2-NS-CT-CT[a]	0	0.0	180.0	2.0

V is the force constant.
[a] SO_2-NS-CT-CT is an improper dihedral angle.

that there are two low-energy conformations of this moiety, which are accessible through (i) S-N bond rotation with a high energy barrier of 8–10 kcal/mol; or (ii) nitrogen inversion with a low barrier of ~2 kcal/mol. We followed the suggestion of Bindal et al. [20] in arriving at the dihedral parameters shown in Table 2. The contributions of all dihedrals involving the S-N bond were set to zero except one specific (CT-N-S-CA) rotation, which is treated as a twofold rotor with $V_{1/2} = 3.4$ kcal/mol. The value for the S-N-CT-CT improper dihedral, which is responsible for maintaining the asymmetry at the nitrogen, was set to zero to allow nitrogen inversion as the lowest energy path for conformational interconversion. The atom-centered partial charges were determined as described previously [11] by fitting the 6-31G* electrostatic potential of the geometry-optimized molecule. These charges, along with the AMBER PREP files, are available on request from the authors.

The crystal structure of the VX-478 complex [1] (PDB code: 1hpv) was used as the starting structure for all simulations of the enzyme–inhibitor complex. In the active site, one aspartic acid residue (Asp^{25} of the first strand) was made neutral based on simple energetic considerations [11]. The hydrogen of the central hydroxyl group of the inhibitor, which is not observed in the crystal structure, is oriented to have a low-energy trans orientation such that the hydroxyl forms the maximum number of hydrogen bonds with the two aspartates. In this configuration, the hydroxyl hydrogen points towards the side chain of $Asp^{25'}$ (of the second strand). Therefore, $Asp^{25'}$ was left charged and Asp^{25} was made neutral. The crystal structure of **II** complexed with HIV-1 protease is not available and has been modeled using the VX-478 complex structure by substituting the NH_2 group of VX-478 with a hydroxy group. The modeled structure is very close to the crystal structure of the 4-OMe analog of VX-478 (E.E. Kim, unpublished results). Each of the complexes was solvated in a rectangular box (64.1 Å × 53.0 Å × 43.1 Å) of about 3200 water molecules in addition to 100 crystallographic water molecules, including the flap water. For mutations in water, each of the inhibitors in their bound conformation was placed at the center of a rectangular water box (37.7 Å × 30.7 Å × 27.2 Å) of about 950 water molecules. All systems were subjected to a standard protocol of energy minimization [11,13]. To maintain coplanarity of the two active-site aspartate side chains observed in the crystal structure, the four inter-oxygen distances of these side chains have been constrained to the values observed in the crystal structure [13]. The minimized structure was equilibrated for 10 ps using a time step of 0.001 ps. For nonbonded interactions, a cutoff distance of 8 Å was used. All simulations were carried out with periodic boundary conditions at constant temperature (300 K) and pressure (1 atm). The bond lengths were constrained using the SHAKE algorithm in the AMBER program. All simulations were carried out with a dielectric constant of 1. The equilibrated structures of the two enzyme–inhibitor complexes were subjected to 100 ps of molecular dynamics with a time step of 0.002 ps, and their trajectories were saved every 1 ps for further analysis. Free energy changes for each mutation were calculated using two approaches: the first included only intermolecular interactions of the perturbed group and the second included both inter- and intramolecular interactions. The entire inhibitor was treated as the perturbed group. To compute the solvation free energies using the second approach, the simulations were repeated in the gas phase. The mutations were achieved over 21 or 51 windows for 51 or 101 ps using a time step of 0.001 or 0.002 ps, starting with the same equilibrated structure in each case. All calculations were carried out on an SGI PowerChallenge computer. Each mutation in water took 12 to 45 h of CPU time, while the same mutations in the enzyme complex took 63 to 238 h of CPU time on a single processor.

The continuum solvation calculations were carried out by the AM1-SM2 and PM3-SM3 methodologies incorporated in AMSOL (v. 4.1) [21]. Since the inhibitors are too large for these calculations, we considered *N,N*-dimethyl benzene sulfonamide and its 4-NH_2 and 4-OH derivatives as truncated models for the inhibitors **III**, **I** and **II**, respectively. Starting with the crystal structure conformations, these models were first fully geometry optimized in vacuo and the optimized geometries were used to evaluate the free energies of solvation. Each of these calculations took less than 5 h of CPU time on the same computer.

Results

The minimized structures of the two complexes (compounds **I** and **II**) are shown in Fig. 3. The 4-NH_2 group of **I** makes hydrogen-bond interactions with the side-chain oxygen and the main-chain carbonyl of $Asp^{30'}$ (3.77 and 3.16 Å, respectively), whereas the 4-OH group of **II** forms hydrogen bonds with the backbone NH groups of $Asp^{29'}$ and $Asp^{30'}$ (3.30 and 3.14 Å, respectively). Variations of these hydrogen-bond distances as a function of

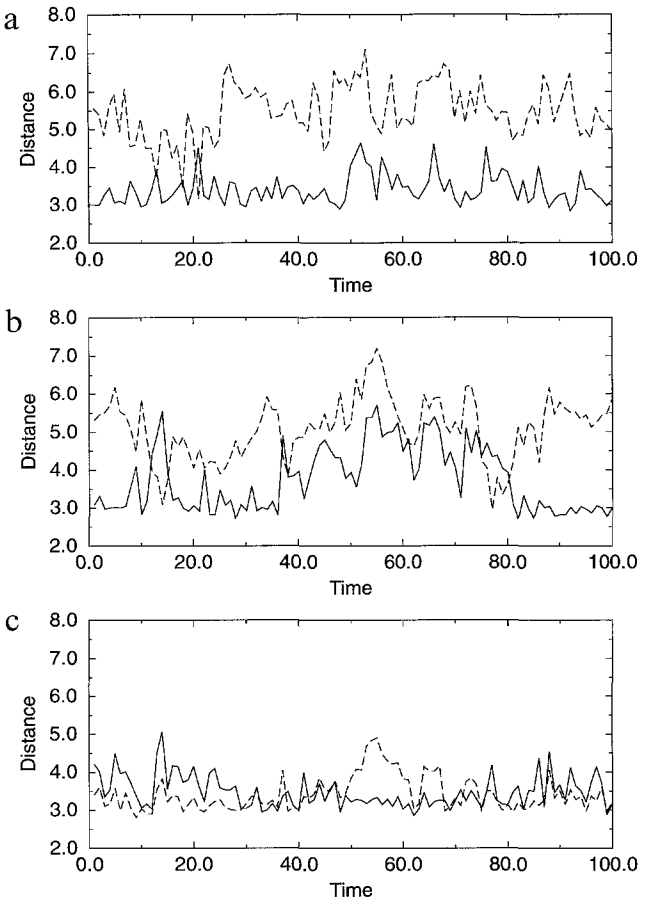


Fig. 4. Variations in hydrogen-bond distances (Å) plotted as a function of simulation time (ps). (a) The distance of NH of **I** to the main-chain carbonyl oxygen of $Asp^{30'}$ (solid line) and to the side-chain oxygen of $Asp^{30'}$ (dashed line); (b) the distance of OH of **II** to the main-chain carbonyl oxygen of $Asp^{30'}$ (solid line) and to the side-chain oxygen of $Asp^{30'}$ (dashed line); and (c) the distances of OH of **II** to the main-chain NH groups of $Asp^{29'}$ (solid line) and $Asp^{30'}$ (dashed line).

simulation time during the 100 ps MD run are plotted in Fig. 4. It is clear from Fig. 4a that the hydrogen-bond interaction of NH_2 with the $Asp^{30'}$ side chain is lost most of the time due to the flexible nature of the side chain, but the hydrogen-bond interaction with the main-chain carbonyl is maintained. The mean distances of the two hydrogen bonds over the entire MD trajectory are calculated to be 5.5 ± 0.7 Å and 3.4 ± 0.4 Å, respectively. The hydrogen-bond distance between the 4-OH group of **II** and the side-chain oxygen of $Asp^{30'}$ (shown in Fig. 4b) is similarly lost most of the time and the interaction of the same hydroxyl group with the main-chain carbonyl of $Asp^{30'}$ is also found to be weaker. The mean distances of these hydrogen bonds are 5.1 ± 0.8 Å and 3.8 ± 0.8 Å, respectively. However, the hydroxyl group of **II** forms two additional hydrogen-bond interactions with the backbone NH groups of $Asp^{29'}$ and $Asp^{30'}$, which remain quite tight throughout the 100 ps of molecular dynamics trajectory. The mean distances of these hydrogen bonds are 3.5 ± 0.4 Å and 3.4 ± 0.5 Å, respectively. Since the 4-OH group of **II** maintains three hydrogen bonds while 4-NH_2 maintains only one hydrogen bond during the 100 ps of MD, it is reasonable to say that **II** makes stronger interactions with the enzyme active site compared to **I**.

The free energy changes for the $NH_2 \rightarrow H$ (**I** → **III**) and $OH \rightarrow H$ (**II** → **III**) mutations in water using intergroup contributions are given in Table 3. The means of five runs give 3.53 ± 0.09 and 3.74 ± 0.18 kcal/mol, respectively, for the two mutations. For the same two mutations in water, the free energy changes obtained using both inter- and intragroup contributions are 3.45 ± 0.24 and 3.80 ± 0.30, respectively (Table 4). These two sets of results are in excellent agreement with each other and are in reasonable agreement with the experimental values. However, these results suggest that **II** is slightly more soluble than **I**, whereas the experimental observation is that **I** has a higher solubility than **II**.

The solvation free energies of the truncated models of the three inhibitors obtained by AMSOL are given in Table 5, which clearly shows that the AMSOL results are model dependent. The AM1-SM2 method gave larger free energy values than the PM3-SM3 method for the $NH_2 \rightarrow H$ and $OH \rightarrow H$ changes, but the free energy values computed for the $NH_2 \rightarrow OH$ change from the two sets of data on the first two changes were almost similar (1.3 and 1.6 kcal/mol, respectively). Although these values are close to the experimental value of 0.78 kcal/mol, the comparison is only qualitative since the AMSOL results

TABLE 3
FREE ENERGY CHANGES (kcal/mol) FOR $NH_2 \rightarrow H$ AND $OH \rightarrow H$ MUTATIONS OF THE P2' GROUP OF VX-478 DERIVATIVES IN WATER (USING INTERGROUP INTERACTIONS ONLY)

Simulation	Time step / total time	$\Delta\Delta G_{sol}$	$\Delta\Delta G$(expt.)
$NH_2 \rightarrow H$			
1	0.002 / 51	3.68	
2	0.001 / 51	3.53	
3	0.002 / 101	3.53	
4[a]	0.002 / 101	3.41	
5[a]	0.001 / 101	3.50	
Average		3.53 ± 0.09	3.82
$OH \rightarrow H$			
6	0.002 / 51	3.77	
7	0.001 / 51	3.95	
8	0.002 / 101	3.52	
9[a]	0.002 / 101	3.87	
10[a]	0.001 / 101	3.61	
Average		3.74 ± 0.18	3.04

The positive value indicates that the compound in front of the arrow is more soluble than the compound behind the arrow. Experimental solvation free energy differences were calculated by the solvation data (190, 51 and 0.3 μg/ml for compounds **I**, **II** and **III**, respectively) [1,2; S. Pazhanisamy et al., unpublished data].

[a] In these simulations 21 windows were used. In all other simulations 51 windows were used.

TABLE 4
FREE ENERGY CHANGES (kcal/mol) FOR $NH_2 \rightarrow H$ AND $OH \rightarrow H$ MUTATIONS OF THE P2' GROUP OF VX-478 DERIVATIVES IN WATER AND GAS PHASES (USING BOTH INTER- AND INTRAGROUP INTERACTION TERMS)

Simulation	Time step / total time	ΔG(aq)	ΔG(gas)	$\Delta\Delta G_{sol}$	ΔΔG(expt.)
$NH_2 \rightarrow H$					
1	0.002 / 51	23.34	20.05	3.29	
2	0.001 / 51	23.40	19.59	3.81	
3	0.002 / 101	23.48	19.88	3.60	
4[a]	0.002 / 101	22.81	19.61	3.20	
5[a]	0.001 / 101	23.19	19.85	3.34	
Average		23.24	19.80	3.45 ± 0.25	3.82
$OH \rightarrow H$					
6	0.002 / 51	20.06	16.21	3.85	
7	0.001 / 51	20.14	16.10	4.04	
8	0.002 / 101	19.85	16.58	3.27	
9[a]	0.002 / 101	20.03	16.09	3.94	
10[a]	0.001 / 101	19.91	16.01	3.90	
Average		20.00	16.20	3.80 ± 0.30	3.04

[a] In these simulations 21 windows were used. In all other simulations 51 windows were used.

were obtained using truncated models of the three inhibitors. Nevertheless, AMSOL results correctly predicted that **I** and **II** are considerably more soluble than **III**, and **I** is more soluble than **II**.

The results of binding free energy differences using the two FEP approaches are given in Tables 6 and 7. The $\Delta\Delta G_{sol}$ and ΔG(aq) values were taken from Tables 3 and 4, respectively, and were subtracted from the ΔG(cplx) values to obtain $\Delta\Delta G_{bind}$ values according to the thermodynamic cycle in Fig. 2. The mean of five runs using intergroup terms gives a ΔG(cplx) equal to 3.94 ± 0.18 kcal/mol for the $NH_2 \rightarrow H$ mutation, which leads to a $\Delta\Delta G_{bind}$ of 0.41 ± 0.22 kcal/mol (Table 6). For the same mutation, the free energy change ΔG(cplx) using both inter- and intragroup contributions is 23.65 ± 0.28 kcal/mol, which leads to a $\Delta\Delta G_{bind}$ of 0.41 ± 0.33 kcal/mol (Table 7). These two values obtained by the two FEP approaches are identical, and are also reasonably close to the experimental value of –0.1 kcal/mol.

For the $OH \rightarrow H$ mutation in the enzyme–inhibitor complex, the mean of five runs using intergroup terms gives a ΔG(cplx) equal to 4.40 ± 0.25 kcal/mol, which leads to a $\Delta\Delta G_{bind}$ of 0.66 ± 0.22 kcal/mol (Table 6). Using both inter- and intragroup contributions, the free energy change ΔG(cplx) was found to be 20.48 ± 0.28 kcal/mol, which leads to a $\Delta\Delta G_{bind}$ of 0.48 ± 0.23 kcal/mol (Table 7). These two values, obtained by the two FEP approaches, are reasonably close to each other. The mean of these two values (0.57 kcal/mol) is reasonably close to the experimental value of 0.95 kcal/mol.

Discussion

The solvation free energy changes for the $NH_2 \rightarrow H$ and $OH \rightarrow H$ mutations are in very good agreement with the experimental values, considering the error of about 0.3 kcal/mol in the FEP calculations and the even larger error (about 0.7 kcal/mol) in the experimental values [2]. The FEP results are also in reasonably good agreement with the results obtained by the PM3-SM3 method, but not with those obtained by the AM1-SM2 method. Jorgensen and Nguyen [22] found excellent agreement between the solvation free energy results on several substituted benzenes obtained by the AM1-SM2 and FEP methods. In the case of sulfonamides, however, neither the AM1 nor the PM3 method is satisfactory [23], but the latter method [24] is considered to be somewhat better than the former. It is, therefore, not surprising that the PM3-SM3 results correlate better with the experimental and FEP results. The FEP results suggest that the OH analog is slightly more soluble than the NH_2 analog, which is in disagreement with the experimental observation. This discrepancy may be due to the fact that the simulation method mimics the dissolution of the gases in water, whereas the experimental solubilities were obtained by dissolution of solids in water. It may be of interest to note that the phenols are generally more soluble than the

TABLE 5
SOLVATION FREE ENERGY DIFFERENCES (kcal/mol) OBTAINED BY AMSOL METHODS FOR *N,N*-DIMETHYL BENZENE SULFONAMIDE AND ITS 4-NH_2 AND 4-OH ANALOGS

$\Delta\Delta G_{sol}$	AM1-SM2	PM3-SM3	Expt.	FEP
$NH_2 \rightarrow H$	6.6	4.9	3.82	3.53
$OH \rightarrow H$	5.3	3.3	3.04	3.74
$NH_2 \rightarrow OH$	1.3	1.6	0.78	–0.21

The data in the last row were computed from the first two rows. To be consistent, the sign convention for the free energy change described in Table 2 is used here.

TABLE 6
FREE ENERGY CHANGES (kcal/mol) FOR $NH_2 \rightarrow H$ AND $OH \rightarrow H$ MUTATIONS OF THE P2' GROUP OF VX-478 DERIVATIVES IN ENZYME–INHIBITOR COMPLEXES AND IN WATER (USING INTERMOLECULAR CONTRIBUTIONS ONLY)

Simulation	Time step / total time	ΔG(cplx)	$\Delta\Delta G_{sol}$	$\Delta\Delta G_{bind}$	ΔΔG(expt.)
$NH_2 \rightarrow H$					
1	0.002 / 51	3.85	3.68	0.17	
2	0.001 / 51	3.80	3.53	0.27	
3	0.002 / 101	4.25	3.53	0.72	
4[a]	0.002 / 101	3.98	3.41	0.57	
5[a]	0.001 / 101	3.86	3.50	0.36	
Average		3.94 ± 0.18	3.53	0.41 ± 0.22	–0.1
$OH \rightarrow H$					
6	0.002 / 51	4.72	3.77	0.95	
7	0.001 / 51	4.34	3.95	0.39	
8	0.002 / 101	4.05	3.52	0.53	
9[a]	0.002 / 101	4.41	3.87	0.54	
10[a]	0.001 / 101	4.41	3.61	0.80	
Average		4.40 ± 0.25	3.74	0.66 ± 0.22	0.95

The positive value indicates that the compound in front of the arrow is more potent than the compound behind the arrow. Experimental binding free energy differences were calculated from the K_i data discussed in the text.
[a] In these simulations 21 windows were used. In all other simulations 51 windows were used.

anilines according to gas phase → solution measurements [25] and these results are reproduced by FEP calculations [22,26]. It is possible, therefore, that the differences in heats of sublimation and vaporization of the three inhibitors may be responsible for the observed small discrepancy between the experimental and FEP results for the $NH_2 \rightarrow OH$ mutation.

The experimental binding free energy differences of the inhibitors are quite small (≤ 1.0 kcal/mol) and very few free energy simulations have been attempted [6,7,27] to reproduce such small free energy values. It is, therefore, encouraging that the results obtained by two different approaches are in reasonably close agreement with the experimental values. The potency of **II** is about 10-fold better than those of **I** or **III**. This is consistent with the fact that the OH group of **II** makes a larger number of stronger interactions with $Asp^{29'}$ and $Asp^{30'}$ of the enzyme than the NH_2 group of **I** with $Asp^{30'}$ (Figs. 3 and 4) and **III** has no substituent on the P2' ring to make hydrogen bonds with the S2' pocket of the enzyme active site. Also, the stronger interaction of the OH analog with the enzyme is reflected by the higher ΔG(cplx) values (in Table 6) for the $OH \rightarrow H$ mutation. As the ΔG(cplx) values in Table 7 include both inter- and intramolecular contributions to free energy, a direct comparison of these data is not meaningful. Assuming that the intramolecular contri-

TABLE 7
FREE ENERGY CHANGES (kcal/mol) FOR $NH_2 \rightarrow H$ AND $OH \rightarrow H$ MUTATIONS OF THE P2' GROUP OF VX-478 DERIVATIVES IN ENZYME–INHIBITOR COMPLEXES AND IN WATER (USING BOTH INTER- AND INTRAGROUP INTERACTION TERMS)

Simulation	Time step / total time	ΔG(cplx)	ΔG(aq)	$\Delta\Delta G_{bind}$	ΔΔG(expt.)
$NH_2 \rightarrow H$					
1	0.002 / 51	23.36	23.34	0.02	
2	0.001 / 51	23.57	23.40	0.17	
3	0.002 / 101	24.09	23.48	0.61	
4[a]	0.002 / 101	23.74	22.81	0.93	
5[a]	0.001 / 101	23.52	23.19	0.33	
Average		23.65 ± 0.28	23.24	0.41 ± 0.33	–0.1
$OH \rightarrow H$					
6	0.002 / 51	20.80	20.06	0.74	
7	0.001 / 51	20.42	20.14	0.28	
8	0.002 / 101	20.05	19.85	0.20	
9[a]	0.002 / 101	20.48	20.03	0.45	
10[a]	0.001 / 101	20.47	19.91	0.56	
Average		20.48 ± 0.28	20.00	0.48 ± 0.23	0.95

[a] In these simulations 21 windows were used. In all other simulations 51 windows were used.

butions to free energy are the same in the gas phase and in solution [6], we can obtain intermolecular contributions by subtracting ΔG(gas) values in Table 4 (column 5) from the ΔG(cplx) values in Table 7, which are 3.8 and 4.3 kcal/mol for the $NH_2 \rightarrow H$ and $OH \rightarrow H$ mutations, respectively. These results also suggest stronger interactions between **II** and the protease. However, larger desolvation costs for **I** and **II** over **III** offset the gain in binding due to the hydrogen bonds formed by the NH_2 and OH groups. Hence, there is no net gain in binding potency of **I** over **III**, and only a very small gain in binding potency of **II** over **III** was observed. Nevertheless, compound **I**, due to its higher solubility, is a much more interesting drug candidate than the equipotent compound **III** or the more potent compound **II**. The higher aqueous solubility of VX-478 added considerably to its oral bioavailability [2].

Conclusions

The free energy simulations have been carried out by two different approaches: the first one included only intergroup interaction terms and the second included both inter- and intragroup interaction terms in the calculation of free energy differences. Both methods quite accurately reproduced the solvation free energy differences between VX-478 and two of its analogs, as well as their differences in binding with HIV-1 protease. To the best of our knowledge, this is the first study to show that the two approaches lead to very similar results. Moreover, the results of this study clearly show that intragroup contributions to free energy changes in the solution and gas phases, indeed, cancel each other and provide a reasonable verification for the original assumption. However, it should be noted that this assumption might not be valid in cases where nonbonded interactions are more tightly coupled with the internal degrees of freedom of the perturbing group. AMSOL methods seem to provide quicker and easier means for estimation of solvation free energies if smaller models of the inhibitors are used. The ability of the FEP approaches to determine both the solvation and binding free energies with reasonable accuracy makes it a very attractive tool for structure-based drug design. However, the computer-intensive nature of the parameter determination process, as well as the long simulation times and unavailability of structural information in a timely manner, seem to limit its widespread use in the drug discovery process. We have found the free energy simulation methods to be most useful towards the end of the inhibitor optimization cycle, when relatively small and focused changes on the inhibitor are attempted.

Acknowledgements

We would like to thank Drs. S. Pazhanisamy and D.J. Livingston for letting us include their unpublished K_i and solubility data as a comparison with the calculated results in this paper, all other members of the HIV project team at Vertex for their contributions, and Drs. R.D. Tung, M.A. Navia, Ajay and D.A. Pearlman for helpful discussions.

References

1 Kim, E.E., Baker, C.T., Dwyer, M.D., Murcko, M.A., Rao, B.G., Tung, R.D. and Navia, M.A., J. Am. Chem. Soc., 117 (1995) 1181.
2 Tung, R.D., Livingston, D.J., Rao, B.G., Kim, E.E., Baker, C.T., Boger, J.S., Chambers, S.P., Deininger, D.D., Dwyer, M.D., Elsayed, L., Funlghum, J., Li, B., Murcko, M.A., Navia, M.A., Novak, P., Pazhanisamy, S., Stuver, C. and Thomson, J.A., manuscript in preparation.
3 Vazquez, M.L., Bryant, M.L., Clare, M., DeCrescenzo, G.A., Doherty, E.M., Freskos, J.N., Getman, D.P., Houseman, K.A., Julien, J.A., Kocan, G.P., Mueller, R.A., Shieh, H.-S., Stallings, W.C., Stegeman, R.A. and Talley, J.J., J. Med. Chem., 38 (1995) 581.
4 Navia, M.A. and Murcko, M.A., Curr. Opin. Struct. Biol., 2 (1992) 202.
5 Lyle, T.A., Wiscount, C.M., Guare, J.P., Thompson, W.J., Anderson, P.S., Darke, P.L., Zugay, Z.A., Emini, E.A., Schleif, W.A., Qunitero, J.C., Emini, E.A. and Huff, J.R., J. Med. Chem., 34 (1991) 1228.
6 Kollman, P.A., Chem. Rev., 93 (1993) 2395.
7 Straatsma, T.P. and McCammon, J.A., Annu. Rev. Phys. Chem., 43 (1992) 407.
8 Cramer, C.J. and Truhlar, D.G., J. Am. Chem. Soc., 113 (1991) 8305.
9 Cramer, C.J. and Truhlar, D.G., J. Comput.-Aided Mol. Design, 6 (1992) 629.
10 McCarrick, M.A. and Kollman, P.A., Methods Enzymol., 241 (1994) 370.
11 Rao, B.G. and Murcko, M.A., J. Comput. Chem., 15 (1994) 1241.
12 Cieplak, P. and Kollman, P.A., J. Comput.-Aided Mol. Design, 7 (1993) 291.
13 Rao, B.G., Tilton, R.F. and Singh, U.C., J. Am. Chem. Soc., 114 (1992) 4447.
14 Reddy, M.R., Varney, M.D., Kalish, V., Viswanadhan, V.N. and Appelt, K., J. Med. Chem., 37 (1994) 1145.
15 Van Gunsteren, W.F., Protein Eng., 2 (1988) 5.
16 AMBER (v. 3.3) is a fully vectorized version of AMBER (v. 3.0) by Singh, U.C., Weiner, P.K., Caldwell, J.W. and Kollman, P.A., University of California, San Francisco, CA, 1986.
17 Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. and Weiner, P., J. Am. Chem. Soc., 196 (1984) 765.
18 Jorgensen, W.L., Chandrasekhar, J. and Madura, J.D., J. Chem. Phys., 79 (1983) 926.
19 QUANTA (v. 3.3) Parameter Handbook, Molecular Simulations, Waltham, MA, 1992.
20 Bindal, R.D., Golab, J.T. and Katzenellenbogen, J.A., J. Am. Chem. Soc., 112 (1990) 7861.
21 AMSOL (v. 4.1), Cramer, C.J., Hawkins, G.D., Lynch, G.C., Truhlar, D.G. and Liotard, D.A., 1994.
22 Jorgensen, W.L. and Nguyen, T.B., J. Comput. Chem., 14 (1993) 195.
23 Dewar, M.J.S., Jei, C. and Yu, J., Tetrahedron, 49 (1993) 5003.
24 Stewart, J.J.P., J. Comput. Chem., 10 (1989) 209.
25 Hine, J. and Mookerjee, P.K., J. Org. Chem., 40 (1975) 292.
26 Rao, B.G. and Singh, U.C., J. Am. Chem. Soc., 111 (1989) 3125.
27 Pearlman, D.A. and Connelly, P.R., J. Mol. Biol., 248 (1995) 696.