Chem Biol Drug Des 2007; **69:** 413–422

Research Article

doi: 10.1111/j.1747-0285.2007.00519.x

Compensating Enthalpic and Entropic Changes Hinder Binding Affinity Optimization

Virginie Lafont[1], Anthony A. Armstrong[2], Hiroyasu Ohtaka[1], Yoshiaki Kiso[3], L. Mario Amzel[2] and Ernesto Freire[1,2,*]

[1]Department of Biology, Johns Hopkins University, Baltimore, MD 21218, USA
[2]Department of Biophysics and Biophysical Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[3]Department of Medicinal Chemistry, Center for Frontier Research in Medicinal Science, Kyoto Pharmaceutical University, Yamashina-ku, Kyoto 607-8412, Japan
**Corresponding author: Ernesto Freire, ef@jhu.edu*

A common strategy to improve the potency of drug candidates is to introduce chemical functionalities, like hydrogen bond donors or acceptors, at positions where they are able to establish strong interactions with the target. However, it is often observed that the added functionalities do not necessarily improve potency even if they form strong hydrogen bonds. Here, we explore the thermodynamic and structural basis for those observations. KNI-10033 is a potent experimental HIV-1 protease inhibitor with picomolar affinity against the wild-type enzyme (K_d = 13 pM). The potency of the inhibitor is the result of favorable enthalpic (ΔH = –8.2 kcal/mol) and entropic ($-T\Delta S$ = –6.7 kcal/mol) interactions. The replacement of the thioether group in KNI-10033 by a sulfonyl group (KNI-10075) results in a strong hydrogen bond with the amide of Asp 30B of the HIV-1 protease. This additional hydrogen bond improves the binding enthalpy by 3.9 kcal/mol; however, the enthalpy gain is completely compensated by an entropy loss, resulting in no affinity change. Crystallographic and thermodynamic analysis of the inhibitor/protease complexes indicates that the entropy losses are due to a combination of conformational and solvation effects. These results provide a set of practical guidelines aimed at overcoming enthalpy/entropy compensation and improve binding potency.

Key words: HIV protease, isothermal titration calorimetry, ITC, lead optimization, protease inhibitors, thermodynamics crystallography

Received 18 April 2007, revised 8 May 2007 and accepted for publication 10 May 2007

The binding affinity is determined by the Gibbs energy ($K_a = e^{-\frac{\Delta G}{RT}}$) which in turn is a function of the enthalpy (ΔH) and entropy (ΔS) changes associated with binding ($\Delta G = \Delta H - T\Delta S$). Improving the binding potency of a compound from 10 μM to 1 nM, as is typical in drug development, requires chemical modifications that will favorably increase the Gibbs energy of binding by 5.5 kcal/mol. Necessarily, this improvement needs to be made by making the enthalpy or entropy changes more favorable. It has been realized by many that the easiest way to improve affinity is by increasing the hydrophobicity of a compound. At the thermodynamic level this strategy leads to entropically optimized compounds (1). This optimization strategy, however, leads to important roadblocks that prevent the achievement of extremely high affinity. Highly hydrophobic compounds lack the solubility necessary for proper formulation and bioavailablity.

It has been shown before that extremely high affinity (picomolar range) requires a combination of favorable enthalpic and entropic interactions (2). The simultaneous optimization of enthalpic and entropic interactions effectively implies overcoming the so-called 'enthalpy/entropy compensation' phenomenon. If enthalpy/entropy compensation were inevitable, affinity optimization would be impossible to achieve. Nevertheless, the introduction of different chemical functionalities during affinity optimization, often carry unwanted enthalpic or entropic penalties. Knowledge of the origin of these penalties is critical for a fast and effective optimization of binding affinity. While entropic optimization by increasing hydrophobicity does not usually carry serious enthalpic penalties and is relatively easy to achieve, other situations are not as straightforward. For example, enthalpic penalties linked to optimization through conformational constraints have been identified by isothermal titration calorimetry (ITC; 3). More importantly, the enthalpic optimization of the binding affinity has been shown to be extremely difficult due to the frequent occurrence of significant entropic penalties (4). For example, it took about 10 additional years of research before enthalpically optimized HIV-1 protease inhibitors could be developed (2,5–7).

Enthalpic optimization of the binding affinity is difficult to achieve. First, a significant enthalpy gain is hard to obtain because the desolvation of polar groups carries a large enthalpic penalty (8). This penalty needs to be overcome by the enthalpy of interaction, which is optimal only when hydrogen bond donor or acceptor functionalities are positioned within the correct distance and angle from the corresponding partner atoms in the protein. Even if these conditions are accomplished and a significantly favorable enthalpy is attained, the binding affinity may not improve due to compensatory entropic

effects. Thus, enthalpic and entropic penalties hinder the enthalpic optimization of binding affinity. In this paper, we show by using a combination of microcalorimetric and crystallographic techniques, that these entropic penalties are caused by conformational entropy losses due to the structuring of the inhibitor and adjacent regions in the protein, and by solvent entropy losses due to a suboptimal burial of non-polar groups.

Methods and Materials

Protein purification

HIV-1 protease was prepared as previously described (6,9). The plasmid encoding the wild-type HIV-1 protease, containing the mutations Q7K/L33I/L63I designed to remove three hypersensitive autolytic sites (10), was expressed as inclusions bodies in *Escherichia coli* BL21(DE3) cells. The HIV-1 protease containing these three mutations is referred to as HIV-1 protease pseudo-wild type (pWT) and has been shown to exhibit the same catalytic behavior as the wild type (6,9,10).

Cells pelleted after growth were resuspended in extraction buffer [20 mM Tris, 1 mM EDTA, and 10 mM β-mercaptoethanol (BME), pH 7.5], and were broken with three passes though a French pressure cell (≥1600 psi). Cell debris and protease-containing inclusion bodies were collected by centrifugation (20 000 × *g* for 20 min at 4 °C). Inclusion bodies were subjected to three cycles of resuspension (glass homogenizer) followed by centrifugation (20 000 × *g* for 20 min at 4 °C). In a fourth cycle, inclusion bodies were solubilized. In each cycle a different buffer was used for resuspension/solubilization: buffer 1 (25 mM Tris, 2.5 mM EDTA, 0.5 M NaCl, 1 mM Gly-Gly, 50 mM BME, pH 7.0); buffer 2 (25 mM Tris, 2.5 mM EDTA, 0.5 M NaCl, 1 mM Gly-Gly, 50 mM BME, 1 M urea, pH 7.0); buffer 3 (25 mM Tris, 1 mM EDTA, 1 mM Gly-Gly, 50 mM BME, pH 7.0); buffer 4 (25 mM Tris, 1 mM EDTA, 5 mM NaCl, 1 mM Gly-Gly, 50 mM BME, 9 M urea, pH 9.0). After the last centrifugation (at 25 °C), HIV-1 protease present in the supernatant was purified using a Q-sepharose column (Q-Sepharose, Amersham Biosciences AB, Uppsala, Sweden) previously equilibrated with buffer 4. Flow-through fractions containing the protease were pooled and acidified by the addition of formic acid to a final concentration of 25 mM. Following an overnight incubation, precipitated contaminants were removed by centrifugation, and the protease was then concentrated to a final volume of 2–3 mL.

Refolding was initiated by diluting 20-fold the concentrated protease sample into 10 mM formic acid at 0 °C. The pH was gradually increased to approximately 3.8 with a solution of 0.1 M NaOH, and then the temperature was raised to 25 °C. Sodium acetate buffer (2.5 M, pH 5.5) was added to increase the pH to 5.0, and the protein was concentrated. Folded protease was exchanged into 1 mM sodium acetate, 2 mM NaCl at pH 5.0 using a gel filtration column (PD-10, Pharmacia) and stored at −20 °C (≥2.5 mg/mL).

The quality of the refolded protein was assessed by enzyme activity assays and by active site titration with well-characterized inhibitors (amprenavir, atazanavir, and TMC-114) by ITC (VP-ITC, Microcal LLC, Northampton, MA, USA) as described previously (2,6).

Isothermal titration calorimetry

Isothermal titration calorimetry experiments were performed by using a high precision VP-ITC titration calorimetric system (MicroCal LLC). Inhibitors and the enzyme were dissolved in the same buffer (10 mM sodium acetate, pH 5.0, 2% DMSO v/v). The binding enthalpies were obtained by injecting the inhibitors (50–90 μM) into the calorimetric cell containing the enzyme (final protein concentration 5–10 μM dimer). The solutions were thoroughly degassed under vacuum and each experiment was performed by one injection of 2 μL followed by 29 injections of 10 μL with a 400-seconds interval between each injection. The inhibitor concentration was adjusted from stoichiometry determination with a standardized protease solution. The heat evolved after each injection was obtained from the integration of the calorimetric signal using ORIGIN 5.0, and the data were analyzed as described previously (11). Direct ITC experiments cannot be used to determine the binding affinity for tight binding inhibitors. Therefore, we used the displacement titration method for binding affinity determination (2,5,12). In calorimetric displacement titrations, the protease prebound to a weaker inhibitor (acetyl pepstatin) is titrated by the high-affinity inhibitor (2,13). Data were analyzed by software developed in this laboratory.

For the interaction of HIV-1 protease with KNI-10033 and KNI-10075, direct titrations were performed at three different temperatures (15, 25 and 35 °C). The change in heat capacity upon binding, ΔC_p, is obtained from the temperature dependence of the binding enthalpy at constant pressure [$\Delta C_p = (\partial \Delta H/\partial T)_P$].

Crystallization

HIV-1 protease at a concentration of 6 mg/mL was prepared in 1 mM sodium acetate, 2 mM NaCl, pH 5. A 15 mM stock solution of inhibitor in 100% DMSO was added to a final inhibitor to protease ratio of 2:1. Crystals were grown by the hanging drop vapor diffusion method. Drops were set by combining 3 μL protein/inhibitor solution with 3 μL of reservoir solution. The drops were equilibrated against a volume of 1 mL of reservoir solution.

Optimal crystallization conditions were determined to be 100 mM MES pH 6.5, NaCl 500 mM, 10 mM DTT, 3 mM NaN_3 for the complex HIV-1 protease pWT/KNI-10033 and citrate buffer pH 5.4, NaCl 750 mM, 100 mM DTT, 3 mM NaN_3 for the complex HIV-1 protease pWT/KNI-10075.

Structure determination

Crystals prior to being flash frozen in liquid nitrogen were washed in a cryoprotectant solution of reservoir buffer made 20% in glycerol. All data were collected at 100 K. Data for wild-type protease complexed with KNI-10033 were collected using Cu-Kα radiation from a rotating anode source on an R-AXIS IV image plate detector (Rigaru, The Woodlands, TX, USA). Data for wild-type protease complexed with KNI-10075 were collected on an ADSC Quantum 315 CCD detector at beamline X25 of the National Synchrotron Light Source in Brookhaven, NY, USA. All data were indexed, integrated, and scaled using the HKL software suite (14). Initial structures were obtained by molecular replacement using the program MOLREP (15) as implemented in the CCP4

crystallographic software suite (16) using PDB 1MSM as the search model or by direct rigid body refinement of the co-ordinates of previous structures against the new measured structure factors. All models underwent rounds of rebuilding and refinement using the programs O (17) and REFMAC5 (18). Ligands were built into the model using the program O. Ligand topology and parameter files for use in O were generated by the program XPLO2D (19), and library files for subsequent refinement were generated by the program SKETCHER in CCP4. Waters were added using the program ARP/WARP (20).

Co-ordinates

The co-ordinates for the structures of KNI-10033 and KNI-10075 with pWT have been deposited in the Protein Data Bank (accession numbers 2PK6 and 2PK5, respectively).

Structural analysis

Hydrogen bonds were identified using the following algorithm. Hydrogen atoms were added to the crystallographic structure and their positions minimized using the program MOE (21). The program CHIMERA (22) was then used to perform hydrogen bond calculations. There is evidence that either Asp 25A or Asp 25B in the protease dimer can be protonated in some of the complexes depending on the inhibitor (23). Without any prior knowledge of the protonation states for these complexes, both aspartic acid residues were considered protonated for purposes of the hydrogen bond calculations. Both were found to serve as proton donors in hydrogen bond interactions, and hydrogen bonds involving these residues were identical for both compounds. The non-bonded contacts were analyzed using Ligplot (24). The surface shape complementarity was calculated with the program SC (25), and changes in accessible surface area (ASA) were calculated using the algorithm of Lee and Richards (26). The changes in solvent ASA upon binding were calculated considering the contribution of water molecules at the protein/ligand interface (27). A parameterization of the binding enthalpy as described by Luque and Freire (27) was used to estimate thermodynamic parameters from structure.

Results and Discussion

Binding energetics of HIV-1 protease to KNI-10033 and KNI-10075

The energetics of association between pWT HIV-1 protease and the compounds KNI-10033 and KNI-10075 (Figure 1) were measured by high sensitivity ITC. Figure 2A shows the results obtained by directly titrating KNI-10033 into HIV-1 protease. The results indicate that KNI-10033 binds to pWT HIV-1 protease with a favorable enthalpy change of −8.2 kcal/mol at 25 °C. The binding affinity was determined by the displacement method using acetyl pepstatin as the weak binding inhibitor (2). Figure 2B shows the titration of the protease prebound to acetyl pepstatin by KNI-10033. From these experiments an association constant (K_a) of 8×10^{10} M^{-1} (K_d = 13 pM) was determined corresponding to a Gibbs free energy of binding of −14.9 kcal/mol at 25 °C. Thus, the entropy contribution ($-T\Delta S$) to the Gibbs energy is also favorable and amounts to −6.7 kcal/mol at 25 °C.

Similar experiments were performed with KNI-10075 and are illustrated in Figure 3. In this case the binding enthalpy is −12.1 kcal/mol, i.e. the presence of the sulfonyl group makes the enthalpy change 3.9 kcal/mol more favorable. However, the large gain in enthalpy is not reflected in an affinity gain. The Gibbs energy of binding of KNI-10075 is −14.6 kcal/mol at 25 °C indicating that the enthalpic gain is completely compensated by an entropic loss resulting in no gain in affinity. The results of the ITC experiments for both compounds are summarized in Table 1.

The binding heat capacity change (ΔC_p) for each compound was determined by performing direct titration experiments at three

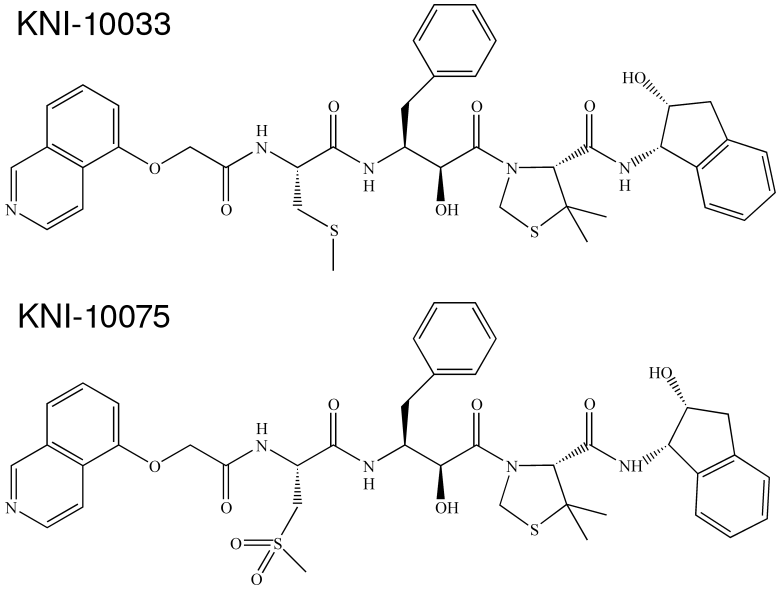

Figure 1: The chemical structure of KNI-10033 and KNI-10075. Both inhibitors are identical except for a single substitution: the $-S-CH_3$ group in KNI-10033 is replaced by a $-SO_2CH_3$ group in KNI-10075.

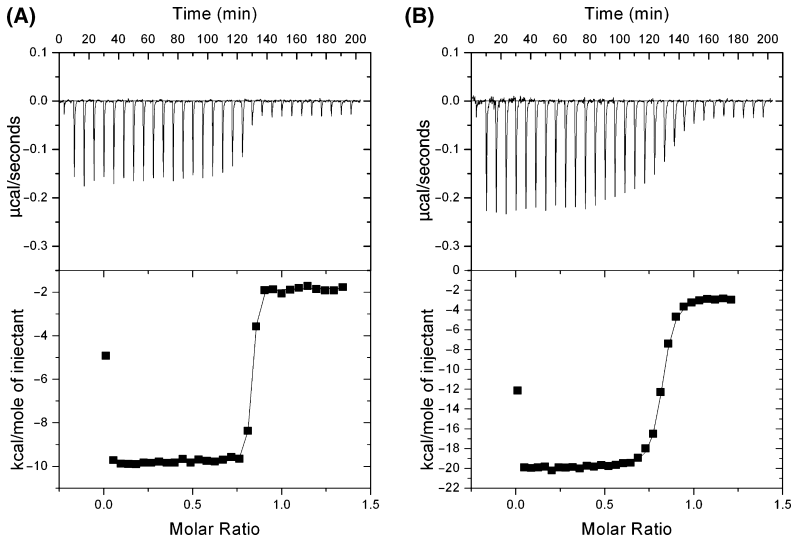

Figure 2: Calorimetric titration of HIV-1 protease with the compound KNI-10033. The experiment on the left panel (A) is a direct titration of KNI-10033 (10 μL per injection of 57 μM solution) into the calorimetric cell (1.4272 mL) containing the HIV protease pseudo-wild type (pWT) at a concentration of 9.6 μM. The experiment was performed at 25 °C in 10 mM acetate buffer, pH 5, 2% DMSO. The experiment on the right panel (B) is a displacement titration experiment in which KNI-10033 (10 μL per injection of 46 μM solution) is injected into the calorimetric cell containing the HIV-1 protease pWT at a concentration of 8.6 μM prebound to acetyl pepstatin at a concentration of 510 μM.

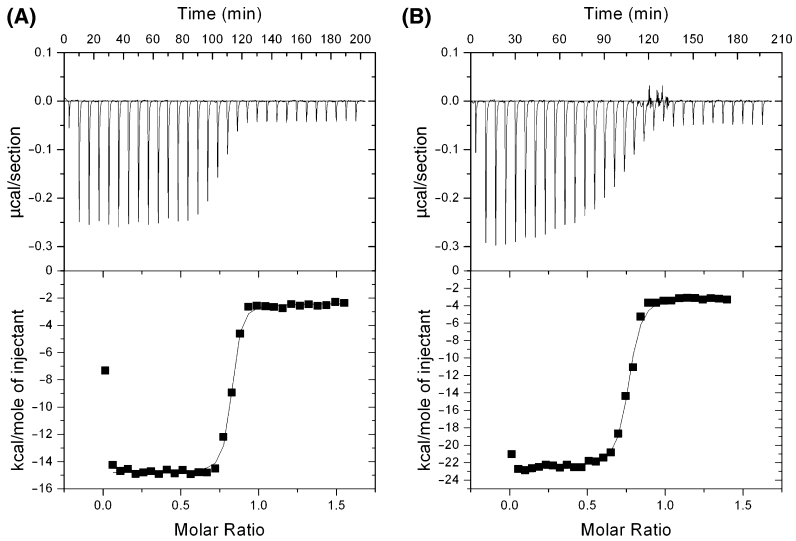

Figure 3: Calorimetric titration of HIV-1 protease with the compound KNI-10075. The experiment on the left panel (A) is a direct titration of KNI-10075 (10 μL per injection of 62 μM solution) into the calorimetric cell (1.4272 mL) containing the HIV protease pseudo-wild type (pWT) at a concentration of 9 μM. The experiment was performed at 25 °C in 10 mM acetate buffer, pH 5, 2% DMSO. The experiment on the right panel (B) is a displacement experiment in which KNI-10075 (10 μL per injection of 55 μM solution) is injected into the calorimetric cell containing the HIV-1 protease pWT at a concentration of 8.8 μM prebound to the acetyl pepstatin at a concentration of 510 μM.

Table 1: The binding thermodynamics of KNI-10033 and KNI-10075 to pseudo-wild type (pWT) HIV-1 protease at 25 °C

	ΔG (cal/mol)	ΔH (cal/mol)	$-T\Delta S$ (cal/mol)	K_d (M)
KNI-10033 → pWT	−14 870 ± 90	−8200 ± 230	−6670 ± 90	$1.3^{-11} \pm 2\ 10^{-12}$
KNI-10075 → pWT	−14 620 ± 190	−12 120 ± 610	−2500 ± 190	$2\ 10^{-11} \pm 8\ 10^{-12}$

The enthalpy changes were measured at 15, 25 and 35 °C to determine the heat capacity change. (ΔC_p) is equal to −461 ± 34 cal/K/mol for KNI-10033 and −305 ± 41 cal/K/mol for KNI-10075.

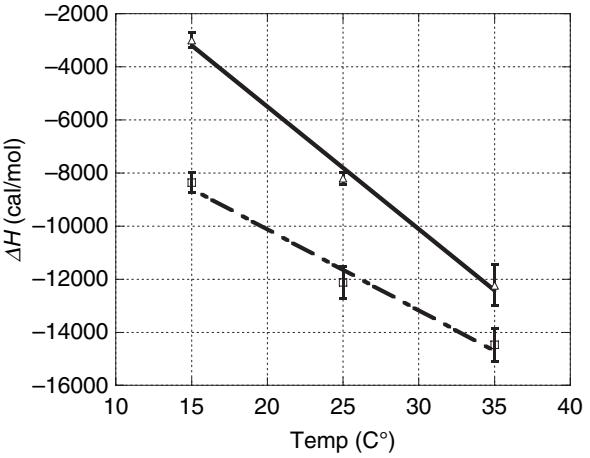

Figure 4: The temperature dependence of the binding enthalpy for KNI-10033 and KNI-10075. The slope of the graph is the change in heat capacity (ΔC_p) upon binding to the HIV-1 protease. The results for KNI-10033 are represented by a solid line and for KNI-10075 by a dashed line.

temperatures (15, 25 and 35 °C). Figure 4 shows the temperature dependence of the binding enthalpy for both inhibitors. These experiments indicate that the binding of both compounds is associated with a negative change in heat capacity. The heat capacity change is equal to −461 ± 34 cal/K/mol for KNI-10033 and −305 ± 41 cal/K/mol for KNI-10075.

Crystallographic structures of KNI-10033 and KNI-10075 with HIV-1 pWT protease

The crystal structures of both compounds in complex with HIV-1 protease were determined under similar conditions. The structures of pWT protease in complex with KNI-10033 and KNI-10075 were refined to 1.45 Å and 1.90 Å, respectively. A sigma A weighted mFo-DFc omit electron density map is shown for the KNI-10033 ligand in Figure 5. The difference omit map shows the electron density for a portion of the model excluded during calculation of the phase angles. Sigma A weighting applied to structure factor amplitudes results in improved density about the omitted portion. Data collection and refinement statistics are summarized in Table 2. In both structures mFo-DFc electron density maps suggested that ligands were bound in two orientations related by the local two-fold axis of the protease. Alternate orientations have been observed in previous crystallographic structures of the HIV-1 protease in complex with clinical inhibitors (28,29). Occupancies were manually adjusted until refined B-factors were similar for both orientations of a given structure. Although both orientations of the ligand interacted similarly with equivalent protease residues on opposite chains, there were slight differences in refined distances and angles in considering hydrogen bonds. Additionally, mFo-DFc electron density maps suggested that the $-SCH_3$ group of KNI-10033 exists in two different conformations which were assigned arbitrarily to different orientations of the ligand. For these reasons, both orientations were considered in analyses of the H-bond patterns and of ASA.

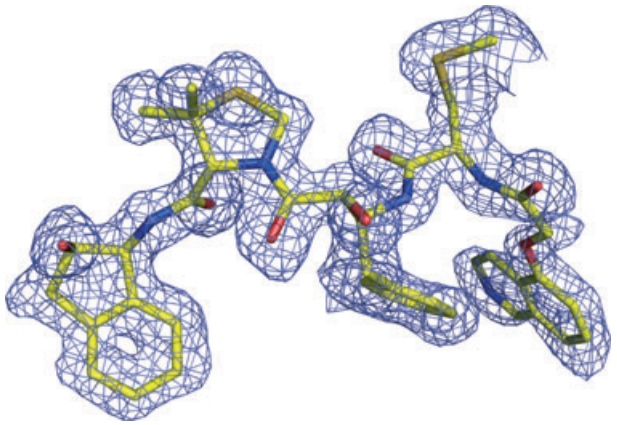

Figure 5: mFo-DFc omit map for the pseudo-wild type/KNI-10033 complex contoured at 2σ. This map was generated by omitting one of the two alternate orientations of the ligand, and fixing the second during subsequent rounds of refinement. The omitted orientation is shown. The map was calculated using CNS (37).

The structures of the two complexes reveal that both compounds bind HIV-1 protease in a similar fashion. There are no major structural differences between the protease molecules. A superposition of the two complexes based on a least-squares fit of protein and compound non-hydrogen atoms reveals negligible structural differences (RMSD = 0.38 Å). The RMSD for protein atoms is 0.38 Å, and the RMSD for atoms common to both compounds (neglecting the substitution) is 0.21 Å. The H-bond interaction, involving the sulfonyl of KNI-10075, pulls this group more toward the backbone of Asp B30 relative to KNI-10033. This, however, results in only negligible conformational differences between common atoms within this half of the molecule, the largest deviation being a slight displacement of one isoquinoline ring relative to the other about an axis perpendicular to their shared plane (Figure 6). The isoquinoline group does not hydrogen bond with the protease; however, and the relative displacement does not result in a difference in van der Waals interactions with the protease.

Both inhibitors bind the protease in an extended conformation and are almost completely buried upon complex formation. Relative to the unbound inhibitor in the bound conformation, binding results in a 95% decrease in total solvent ASA for KNI-10033 (948 $Å^2$ buried) and a 94% decrease for KNI-10075 (918 $Å^2$ buried). Binding of KNI-10033 to pWT protease buries a total of 985 $Å^2$ of non-polar surface area and 682 $Å^2$ of polar surface area. For KNI-10075 the amount of buried non-polar and polar surfaces are 953 $Å^2$ and 704 $Å^2$, respectively. Both compounds bind the protease with high shape complementarity (sc values of 0.811 and 0.805 for KNI-10075 and KNI-10033, respectively). The 2-indanol ring binds in a hydrophobic pocket comprising Ile B50, Val A32, Ala A28, and Ile A84, and the phenyl ring packs against the side chain of Ile A84. The $-SO_2$ group of KNI-10075 and the $-SCH_3$ group of KNI-10033 bind in a pocket formed by the residues Ala 28, Asp 29, Asp 30, Val 32, and Ile 47 of chain B.

Bound KNI-10033 is involved in a total of 12 hydrogen bond interactions (Figure 7). Three water molecules make a total of four

Table 2: Statistics for crystallographic data collection and refinement

	KNI-10033:pWT	KNI-10075:pWT
Space group	$P2_12_12$	$P2_12_12$
a (Å)	58.4	58.8
b (Å)	86.0	86.0
c (Å)	46.6	46.2
Data collection		
Source	Rotating anode	X25
Wavelength	1.54	1.10
Detector	R-AXIS IV	ADSC Q315
Resolution (Å)	33.5–1.45	48.6–1.90
Outermost shell (Å)	1.50–1.45	1.97–1.90
Total number of reflections	153 058	114 205
Number of unique reflections	41 430 (4071)	19 013 (1856)
Completeness (%)	97.4 (96.3)	99.8 (99.9)
Redundancy	3.7 (3.6)	6.0 (5.9)
$\langle I\rangle/\langle\sigma(I)\rangle$	34.7 (1.7)	39.0 (8.6)
R^{a}_{sym}	0.05 (0.81)	0.11 (0.42)
Refinement		
R^{b}_{work}	0.20	0.16
$R^{b,c}_{free}$	0.23	0.20
Stereochemistry		
R.M.S. bond lengths (Å)	0.017	0.015
R.M.S. angles (°)	1.60	1.40
B-factor		
Protein	18.51	21.97
Inhibitor	16.87	21.04
Water	32.86	37.90
Model composition		
Amino acids (atoms)	198 (1516)	198 (1516)
Ligands (atoms)		
Inhibitor	1 (54)[d]	1 (56)[d]
Water	273	237
Glycerol	1 (6)	1 (6)

Data collection statistics given in parentheses are for the highest resolution shell.

[a] $R_{sym} = \Sigma\Sigma_j|I_j - \langle I\rangle|/\Sigma\langle I\rangle$, where I_j is the measured intensity of reflection j and $\langle I\rangle$ is the mean over multiple measurements.

[b] R_{free} is the R value as calculated below for a randomly selected test set of the intensity data not used during refinement. The test set consisted of 10% of the total intensity data.

[c] $R_{work,\ free} = \Sigma||F_o| - |F_c||/|F_o|$.

[d] Two ligand orientations were modeled with occupancies 0.60 and 0.40.

hydrogen bonds to the compound, two to amide carbonyls and one to the hydroxyl substitution on the terminal indane ring. A bound water bridges centrally located amide carbonyls of the inhibitor and backbone nitrogens of residue 50 in both chains A and B of the protease. This tetrahedral co-ordination of a water is a common binding feature among HIV-1 protease inhibitors in clinical use with the exception of Tipranavir (30,31). All hydrogen bonds between inhibitor and protease involve the protein backbone with the exception of those involving the side chains of Asp 25 on chains A and B. Both side chains participate in a hydrogen bond with the central hydroxyl group of the inhibitor. The side chain of Asp A25 additionally forms a hydrogen bond with an amide carbonyl on the indane group of the inhibitor. Structural analysis shows that all the hydrogen bonds found in the complex with KNI-10033 (Figure 7A) are also found in the complex with KNI-10075 (Figure 7B); however, an additional hydrogen bond is found in the latter. This hydrogen bond is formed between one of the oxygen atoms of the $-SO_2$ group and the main chain nitrogen of Asp B30.

Structure-based thermodynamic analysis of the binding enthalpy

The crystallographic structures of the two complexes indicate no significant structural differences. It is apparent that the differences in binding thermodynamics can be attributed to the presence of the sulfonyl group.

Previously, a refined structural parameterization of the enthalpy change associated with the binding of small ligands was presented (27). Four set of parameters were found to be important for a quantitative account of the binding enthalpy in structural terms: (i) the interactions between ligand and protein, reflected in changes in solvent ASAs for ligand and protein; (ii) the conformational change associated with binding; (iii) the presence of water molecules at the protein/ligand interface (a cutoff of 6 Å for completely buried water molecules was reported); and (iv) any effects due to protonation/deprotonation coupled to ligand binding. The resulting parametric equation is:

$$\Delta H(25) = \Delta H_{conf} - 7.35 \times \Delta ASA_{np} + 31.06 \times \Delta ASA_{pol} + \Delta H_{prot} \quad (1)$$

where ΔH_{conf} is the enthalpy change associated with the conformational change and was estimated previously as 5.9 kcal/mol for the HIV-1 protease (27). ΔH_{prot} is the enthalpy associated with any coupled protonation/deprotonation process. ΔASA_{np} and ΔASA_{pol} are the changes in non-polar and polar solvent ASA, respectively. ΔASA_{np} and ΔASA_{pol} are calculated by explicitly including in the structure water molecules that are buried and within 6 Å of the inhibitor. The number of water molecules that satisfy those constraints is seven for both complexes. Protonation effects for these two inhibitors were evaluated by performing microcalorimetric titrations in buffers with different heats of ionization (data not shown; 6). For both inhibitors, experiments in the different buffers yielded similar enthalpies, from which it is concluded that protonation effects do not contribute measurably to the binding of these inhibitors.

Application of eqn 1 to the KNI-10033 complex yields a binding enthalpy of −8.0 kcal/mol, very close to the experimental value of −8.2 kcal/mol. In contrast, application of the same equation to the KNI-10075 complex yields a binding enthalpy of −8.8 kcal/mol, more favorable than that of KNI-10033 but significantly smaller than the experimental value of −12.1 kcal/mol (data are summarized in Table 3). The difference between predicted and experimental values is not unexpected as the sulfonyl moiety was not part of the training set used in the development of the parameterization. Previously, we have noticed that sulfonyl groups that participate in hydrogen bonding interactions and are buried away from the solvent make a strong contribution to the binding enthalpy (32). A similar underestimation of the experimental binding enthalpy is obtained for amprenavir, TMC126 and TMC114, other protease inhibitors with buried sulfonyl groups that participate in hydrogen

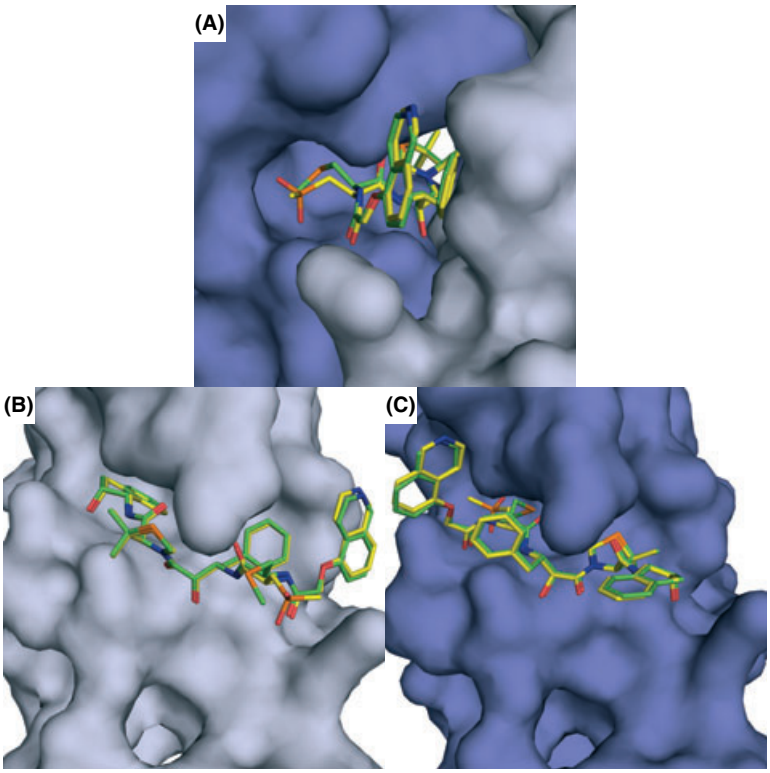

Figure 6: KNI-10033 and KNI-1-0075 bind the HIV-1 protease in a similar conformation and interacting with the same residues in the protease. Shown are surface representations of HIV-1 protease pseudo-wild type bound to KNI-10033 and KNI-10075. (Panel A) The compounds bind in the active site located in a cavity at the dimer interface. The flaps of the protease are located above the inhibitor. For better viewing the interaction of the compounds with chain B (panel B) and chain A (panel C) are shown separately with the second chain omitted. In all panels KNI-10033 and KNI-10075 are superimposed.

bonding interactions. The same situation is observed in other systems such as HMG-CoA reductase, in which the inhibitor rosuvastatin, which contains a hydrogen-bonded, buried sulfonyl moiety, exhibits a much stronger binding enthalpy than inhibitors that lack this group (32). On average, the additional contribution of the sulfonyl group to the binding enthalpy can be estimated as -3.0 ± 1 kcal/mol for the examples above.

Structure-based thermodynamic analysis of the binding entropy

The binding affinity of KNI-10033 and KNI-10075 are similar despite the fact that the binding enthalpy of KNI-10075 is −3.9 kcal/mol more favorable, i.e. if the binding entropies for both compounds were the same, KNI-10075 would have a binding affinity close to three orders of magnitude stronger than that of KNI-10033. However, KNI-10075 suffers entropic losses that totally overcome the enthalpic gains introduced by the presence of the sulfonyl group. The origin of the entropic loses can be investigated by considering the two major entropic contributions to binding, the desolvation entropy, and the conformational entropy changes.

Differences in solvation entropy originate from differences in the magnitude and nature of the surface area that becomes buried from the solvent upon binding.

The burial of non-polar groups upon binding results in a gain in desolvation entropy due to the release of water into the bulk solvent. The magnitude of these changes is proportional to the surface area that becomes buried from the solvent upon binding (33). Based on the crystallographic structures, the complex of KNI-10033 with the HIV-1 protease buries 32 $Å^2$ more non-polar surface area and 22 $Å^2$ less polar surface area than KNI-10075. For KNI-10033, the increase in non-polar surface is consistent with a more favorable desolvation entropy. This structural characteristic is also reflected in a more negative heat capacity change for KNI-10033 as the burial of non-polar groups from the solvent is known to be associated with a negative heat capacity change (33). Using a previous parameterization based upon changes in ΔASA to estimate the desolvation entropy differences (33), a value of $(-T\Delta\Delta S)$ close to 1.5 kcal/mol is obtained relative to an experimental value of 4.2 kcal/mol at 25 °C, suggesting that desolvation accounts only partially for the entropic compensation of the enthalpic gain.

The second major contributor to the entropy change is the conformational entropy. It can be hypothesized that the strong hydrogen bond made by the sulfonyl with the amide of Asp 30 in addition to the larger volume of the sulfonyl group relative to the thioether group in KNI-10033 would result in an increased structuring of the inhibitor as well as nearby protease residues upon binding. Based upon the experimental entropy values and the estimated differences in solvation, the difference in conformational entropy must be on the order of 9 cal/k/mol. Any differences in structuring will be reflected in differential changes in the temperature factors in the crystal structures. Figure 8 shows the temperature factors for both inhibitors within the complex. It is clear that the sulfonyl region in KNI-10075 is the most structured region of the bound inhibitor while the corresponding region in KNI-10033 shows some mobility. Furthermore, the $-SCH_3$ group in KNI-10033 adopts two

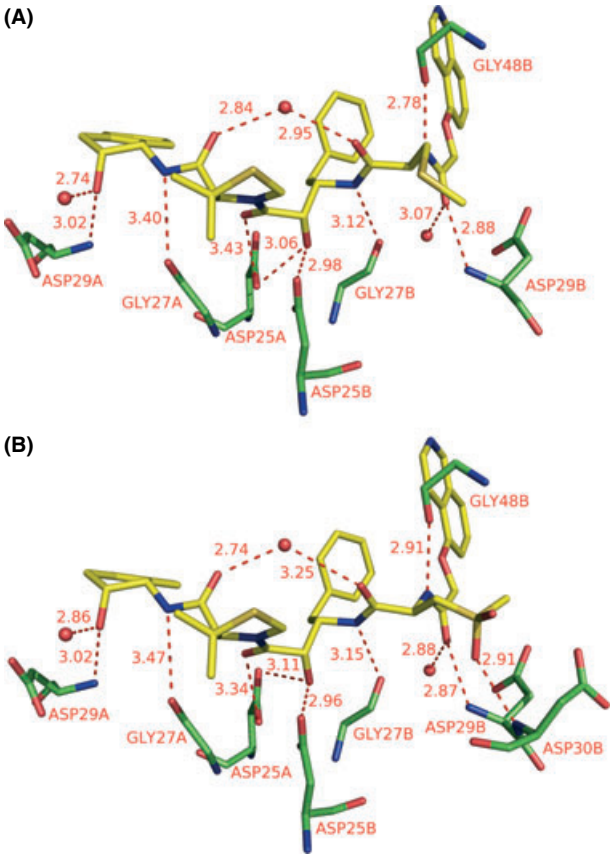

Figure 7: Hydrogen bonds interaction between HIV-1 protease pseudo-wild type and KNI-10033 (A) and KNI-10075 (B). Hydrogen bonds were calculated using the program CHIMERA. Red spheres indicate water molecules. The compound and the protein are colored by atom type with oxygen in red, azote in blue, sulfur in orange, and carbon in yellow for the compound and green for the protein. The hydrogen bonds are indicated in blue dash lines. Twelve hydrogen bonds are formed directly between the compound and the protein or water molecules. A buried water molecule makes two additional hydrogen bonds with the two NH groups of residue Ile 50 present in each flap.

conformations in the crystal structure of the complex. Differences in B-factors are also observed in the protease regions surrounding these moieties. For instance, whereas in the KNI-10075 complex structure the average per atom B-factor for residue Ile B50 is below the average for residues within the core of the protease (5 Å from the inhibitor), the same residue in the KNI-10033 complex structure has an average per atom B-factor which is more than 4 $Å^2$ above that of the core residues in that structure.

Literature estimates of conformational entropy changes indicate that the average conformational entropy loss when burying a side chain ranges from 0 (Gly, Ala) to 7 cal/K/mol (Arg) with the most common values being between 2 and 4 cal/K/mol (34). In addition, the overall change per residue upon protein folding is on the order of 6 cal/K/mol and that for small molecular weight organic molecules binding to proteins, the conformational entropy loss is on the order of 2.5 cal/K/mol per rotatable bond (35,36). The estimated value of 9 cal/K/mol for the conformational entropy differences associated with the binding of KNI-10033 and KNI-10075 indicates that relatively small differences, equivalent to those associated with the structuring of two amino acid residues or freezing the equivalent of three rotatable bonds are sufficient to neutralize the binding affinity gains induced by a strong hydrogen bond.

Conclusions

Affinity optimization is a critical and difficult step in drug development. From a purely thermodynamic point of view, an affinity improvement means an increase in the favorable Gibbs free energy of binding, which can be achieved by enthalpic or entropic improvements as far as they are not compensated by opposite entropic or enthalpic changes. The introduction of a functionality that establishes a strong hydrogen bond does not necessarily result in higher binding affinity as demonstrated here. The engineering of hydrogen bonds for improved binding affinity requires considerations other than donor/acceptor distances or angles. It is apparent that the structuring associated with hydrogen bond formation can significantly compensate for any improvement in binding affinity. Consequently, optimization of the contribution of hydrogen bonds to binding affinity requires minimization of any structuring effect. The minimization of the structuring effect can be achieved by directing the hydrogen bond toward well-structured regions of the protein and/or by placing the acceptor/donor functionalities in regions of the compound that are already conformationally constrained in the unbound state. On the contrary, targeting hydrogen bonds to unstructured regions will carry the maximal structuring penalty.

Another detrimental effect is a lower degree of desolvation originating from the structural constraints imposed by the hydrogen bond that prevents full accommodation of neighboring groups to maximize burial from solvent. According to the results presented here, desolvation accounts for only 35% of the total entropy compensation effect, the majority (65%) being accounted for by conformational entropy losses. Conformational entropy losses arise from both the inhibitor as well as the protein. It follows that a gain in binding potency can only be achieved if these compensatory entropic effects associated with hydrogen bonding are minimized. In order to maximize binding affinity, hydrogen bonds must be targeted against

Table 3: Calculated binding enthalpy using energy parameterization derived from accessible surface area calculations

	Number of water[a]	ΔASA_{pol} ($Å^2$)	ΔASA_{np} ($Å^2$)	ΔH_{int} (cal/mol)	$\Delta H(25) = 5900 + \Delta H_{int}$ (cal/mol)
KNI-10033 pWT	7	−682 ± 17	−985 ± 30	−13 945 ± 750	−8045 ± 750
KNI-10075 pWT	7	−704 ± 2	−953 ± 38	−14 860 ± 350	−8960 ± 350

[a]Water within 6 Å of the inhibitor and fully buried.

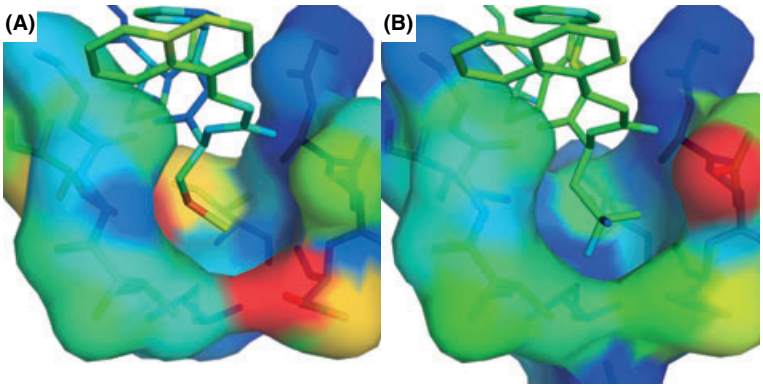

Figure 8: (Panel A) KNI-10033 and surface representation of residues of HIV-1 protease within 5 Å of the group $-SCH_3$. (Panel B) KNI-10075 and surface representation of residues of HIV-1 protease within 5 Å of the group SO_2. In both figures, atoms are colored by ΔB defined as $\Delta B_i = B_i - \langle B \rangle$, where B_i is the crystallographic temperature factor for atom i, and the mean temperature factor, $\langle B \rangle$, is overall protease and inhibitor atoms for a given structure. The scale is set from −5 (blue) to +5 $Å^2$ (red).

structured regions of the protein, thus minimizing conformational entropy losses. Also, hydrogen bond donors or acceptors must be placed in regions of the drug molecule that are conformationally constrained and do not lose significant conformational entropy upon binding.

Acknowledgment

This work was supported by grants from the National Institutes of Health GM 57144 (E.F.) and GM 45540 (L.M.A.).

References

1. Velazquez-Campoy A., Todd M.J., Freire E. (2000) HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. Biochemistry;39:2201–2207.
2. Velazquez-Campoy A., Kiso Y., Freire E. (2001) The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. Arch Biochem Biophys;390:169–175.
3. Davidson J.P., Lubman O., Rose T., Waksman G., Martin S.F. (2002) Calorimetric and structural studies of 1,2,3-trisubstituted cyclopropanes as conformationally constrained peptide inhibitors of Src SH2 domain binding. J Am Chem Soc;124:205–215.
4. Dunitz J.D. (1995) Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. Chem Biol;2:709–712.
5. Ohtaka H., Velazquez-Campoy A., Xie D., Freire E. (2002) Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. Protein Sci;11:1908–1916.
6. Velazquez-Campoy A., Luque I., Todd M.J., Milutinovich M., Kiso Y., Freire E. (2000) Thermodynamic dissection of the binding energetics of KNI-272, a potent HIV-1 protease inhibitor. Protein Sci;9:1801–1809.
7. King N.M., Prabu-Jeyabalan M., Nalivaika E.A., Wigerinck P., de Bethune M.P., Schiffer C.A. (2004) Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor. J Virol;78:12012–12021.
8. Cabani S., Gianni P., Mollica V., Lepori L. (1981) Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous-solution. J Solution Chem;10:563–595.
9. Todd M.J., Semo N., Freire E. (1998) The structural stability of the HIV-1 protease. J Mol Biol;283:475–488.
10. Mildner A.M., Rothrock D.J., Leone J.W., Bannow C.A., Lull J.M., Reardon I.M., Sarcich J.L. et al. (1994) The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties. Biochemistry;33:9405–9413.
11. Leavitt S., Freire E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. Curr Opin Struct Biol;11:560–566.
12. Sigurskjold B.W. (2000) Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. Anal Biochem;277:260–266.
13. Vega S., Kang L.W., Velazquez-Campoy A., Kiso Y., Amzel L.M., Freire E. (2004) A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease. Proteins;55:594–602.
14. Otwinowski Z., Minor W. (1997) Processing of x-ray diffraction data collected in oscillation mode. In: Carter J.C.W., Sweet R.M., editors. Methods in Enzymology. Vol. 276: Academic Press;p. 307–326.
15. Vagin A., Teplyakov A. (1997) MOLREP: an automated program from molecular replacement. J Appl Cryst;30:1022–1025.
16. 4 CCPn (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr;50(Pt 5):760–763.
17. Jones T.A., Zou J.Y., Cowan S.W., Kjeldgaard M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A;47(Pt 2):110–119.

18. Murshudov G.N., Vagin A.A., Dodson E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr;53(Pt 3):240–255.
19. Kleywegt G.J., Jones T.A. (1998) Databases in protein crystallography. Acta Crystallogr D Biol Crystallogr;54(Pt 6 Pt 1):1119–1131.
20. Perrakis A., Morris R., Lamzin V.S. (1999) Automated protein model building combined with iterative structure refinement. Nat Struct Biol;6:458–463.
21. CCG (1996) The Molecular Operating Environment (MOE). Available under license from Chemical Computing Group Inc, Montreal, Canada.
22. Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C. et al. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. J Comput Chem;25:1605–1612.
23. Gustchina A., Sansom C., Prevost M., Richelle J., Wodak S.Y., Wlodawer A. et al. (1994) Energy calculations and analysis of HIV-1 protease-inhibitor crystal structures. Protein Eng;7:309–317.
24. Wallace A.C., Laskowski R.A., Thornton J.M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng;8:127–134.
25. Lawrence M.C., Colman P.M. (1993) Shape complementarity at protein/protein interfaces. J Mol Biol;234:946–950.
26. Lee B., Richards F.M. (1971) The interpretation of protein structures: estimation of static accessibility. J Mol Biol;55:379–400.
27. Luque I., Freire E. (2002) Structural parameterization of the binding enthalpy of small ligands. Proteins;49:181–190.
28. Clemente J.C., Coman R.M., Thiaville M.M., Janka L.K., Jeung J.A., Nukoolkarn S. et al. (2006) Analysis of HIV-1 CRF_01 A/E protease inhibitor resistance: structural determinants for maintaining sensitivity and developing resistance to atazanavir. Biochemistry;45:5468–5477.
29. Kovalevsky A.Y., Tie Y., Liu F., Boross P.I., Wang Y.F., Leshchenko S. et al. (2006) Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. J Med Chem;49:1379–1387.
30. Lam P.Y., Jadhav P.K., Eyermann C.J., Hodge C.N., Ru Y., Bacheler L.T. et al. (1994) Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. Science;263:380–384.
31. Muzammil S., Armstrong A.A., Kang L.W., Jakalian A., Bonneau P.R., Schmelmer V. et al. (2007) Unique thermodynamic response of Tipranavir to HIV-1 protease drug resitance mutations; 81:5144–5154.
32. Carbonell T., Freire E. (2005) Binding thermodynamics of statins to HMG-CoA reductase. Biochemistry;44:11741–11748.
33. Luque I., Freire E. (1998) Structure-based prediction of binding affinities and molecular design of peptide ligands. Methods Enzymol;295:100–127.
34. Lee K.H., Xie D., Freire E., Amzel L.M. (1994) Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. Proteins;20:68–84.
35. D'Aquino J.A., Gomez J., Hilser V.J., Lee K.H., Amzel L.M., Freire E. (1996) The magnitude of the backbone conformational entropy change in protein folding. Proteins;25:143–156.
36. D'Aquino J.A., Freire E., Amzel L.M. (2000) Binding of small organic molecules to macromolecular targets: evaluation of conformational entropy changes. Proteins;4(Suppl. 4):93–107.
37. Brunger A.T., Adams P.D., Clore G.M., DeLano W.L., Gros P., Grosse-Kunstleve R.W. et al. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr;54(Pt 5):905–921.