Simple, Intuitive Calculations of Free Energy of Binding for Protein−Ligand Complexes. 3. The Free Energy Contribution of Structural Water Molecules in HIV-1 Protease Complexes

Micaela Fornabaio,[†,‡] Francesca Spyrakis,[†] Andrea Mozzarelli,[†,‡] Pietro Cozzini,*[,§] Donald J. Abraham,[||] and Glen E. Kellogg*[,||]

Department of Biochemistry and Molecular Biology, National Institute for the Physics of Matter, Molecular Modeling Laboratory, Department of General and Inorganic Chemistry, University of Parma, 43100 Parma, Italy, and Department of Medicinal Chemistry & Institute for Structural Biology and Drug Discovery, School of Pharmacy, Virginia Commonwealth University, Richmond, Virginia 23298-0540

Received November 26, 2003

Structural water molecules within protein active sites are relevant for ligand−protein recognition because they modify the active site geometry and contribute to binding affinity. In this work an analysis of the interactions between 23 ligands and dimeric HIV-1 protease is reported. The X-ray structures of these complexes show the presence of four types of structural water molecules: water 301 (on the symmetry axis), water 313, water 313bis, and peripheral waters. Except for water 301, these are generally complemented with a symmetry-related set. The GRID program was used both for checking water locations and for placing water molecules that appear to be missing from the complexes due to crystallographic uncertainty. Hydropathic analysis of the energetic contributions using HINT indicates a significant improvement of the correlation between HINT scores and the experimentally determined binding constants when the appropriate bridging water molecules are taken into account. In the absence of water $r^2 = 0.30$ with a standard error of $\pm$ 1.30 kcal mol^{-1} and when the energetic contributions of the constrained waters are included $r^2 = 0.61$ with a standard error of $\pm$ 0.98 kcal mol^{-1}. HINT was shown to be able to map quantitatively the contribution of individual structural waters to binding energy. The order of relevance for the various types of water is water 301 > water 313 > water 313bis > peripheral waters. Thus, to obtain the most reliable free energy predictions, the contributions of structural water molecules should be included. However, care must be taken to include the effects of water molecules that add information value and not just noise.

Introduction

"In silico screening" of libraries of chemicals is currently one of the main tools to identify new lead compounds. Key to this approach is the development of computational models that incorporate as much as possible of the physicochemical principles governing protein−ligand encounters. The major molecular events to be considered are (i) recognition between the ligand and protein active site residues mediated by topological and functional complementarity, (ii) release of water molecules from the ligand and the protein, and (iii) conformational changes of ligand and protein induced by the formation of the complex. Each of these is characterized by enthalpic and entropic contributions. Despite intense efforts to model computationally the free energy of ligand binding to proteins, its precise evaluation has yet to be reliably obtained and still represents a major issue in medicinal chemistry and, in general, in biochemistry. Most approaches consider the process as a sequence of steps summing up distinct enthalpic and entropic contributions, whereas, in reality, the ligand−protein recognition is a concerted event and thermodynamic quantities cannot be just simply added up.[1] The enthalpic contributions are estimated by theoretical models, knowledge-based potential functions, or empirical approaches.[2,3] The entropic contributions to binding are much less well-defined and often poorly quantitated or even ignored. The release of water molecules from a protein active site/ligand, resulting in hydrophobic interactions, has been particularly difficult to model computationally.[4] These difficulties in the quantitative evaluation/modeling of entropy are also reflected in the lack of robust prediction methods for protein folding. The energetics of water play a very significant role, directly as in water-mediated hydrogen bonds, and indirectly in ligand and protein desolvation and hydrophobic interactions. All of these effects should be considered in accurately modeling the total free energy of binding. Thus, comprehensive in silico (virtual) screening and free energy scoring methods should, in some manner, take into account water molecules either explicitly or implicitly.

A variety of approaches for evaluating the contribution of water molecules to the free energy of binding have been proposed, and their advantages and limita-

* To whom correspondence should be addressed. P.C.: phone +39 0521 905669, fax +39 0521 905556, e-mail pietro.cozzini@unipr.it; G.K.: phone +01 804 828-6452, fax +01 804 827-3664, e-mail glen.kellogg@vcu.edu.

[†] Department of Biochemistry and Molecular Biology, University of Parma.

[‡] National Institute for the Physics of Matter, University of Parma.

[§] Molecular Modeling Laboratory, Department of General and Inorganic Chemistry, University of Parma.

[||] Virginia Commonwealth University.

10.1021/jm030596b CCC: $27.50
Published on Web 07/29/2004

tions have been reviewed recently.[5] Briefly, the ligand surface area accessible to solvent (SASA) can correlate with the hydrophobic free energy;[6−8] hybrid Free Energy Perturbation calculations[9] with Generalized Born[10] or Poisson−Boltzmann[11,12] electrostatics methods simulate solvent energetics; conserved water molecules modifying the topology of the protein active site[13] can be identified experimentally (crystallography) or computationally (grid-based methods[14]); and, computational methods explicitly considering all water molecules in simulations yield good results but are expensive to perform.[15]

However, a simple empirical method, HINT (Hydropatic INTeractions)[16−18] for estimating the free energy of binding, is based on $LogP_{o/w}$, the partition coefficient between 1-octanol and water. $LogP_{o/w}$ is a thermodynamic parameter related to ΔG for solute transfer between the two solvents. Since water is integral to $LogP_{o/w}$, solvation/desolvation, and the resulting hydrophobic effects, i.e., entropy, are encoded within the HINT constants and HINT yields interaction scores proportional to the free energies of association for ligand binding[19−22] and protein−protein interactions.[23,24]

We have recently undertaken a detailed examination of the issues associated with the calculation of binding free energy for protein−ligand systems using HINT. Our method is able to, with a simple, intuitive, and rapid protocol, predict the free energy of binding for a diverse collection of ligand−protein complexes without structural water molecules bridging ligand and protein.[21] In a second report we detailed the effects of pH on protein and ligand functional group ionization states and suggested a protocol for adapting molecular models for these effects.[22] In the course of both of these projects we consistently encountered situations where the presence of (ordered) water at the ligand binding site was influencing the energetics of binding. The present work is the third contribution in this series. Here we illustrate the energetic contribution of water molecules that are bridging between the protein and ligand and thus influencing binding. These water molecules can be variously termed as "crystallographic", "structural", "constrained" or "conserved", but their role in ligand binding has seldom been investigated thoroughly.[13,25,26] In a previous study of dimer−dimer tetramerization for native and mutant hemoglobins[24] we demonstrated that the energetic contribution of new or displaced water molecules was critical for accurate free energy predictions. Here, we extend this concept to water molecules bridging inhibitors and active site residues of the human immunodeficient virus (HIV-1) protease. HIV-1 protease, a dimer comprised of two identical polypeptide chains, has a single symmetric active site formed by residues from both subunits.[27,28] It is essential for HIV replication and thus represents a crucial therapeutic target in the development of treatments for AIDS.[29,30] Because of its therapeutic relevance and the large volume of experimental research into its structure, inhibitors, and binding, numerous computational studies have been reported.[8,31−37] In the present study, the specific contribution of each water molecule bridging protein and ligand is characterized and applied to binding free energy predictions for 23 HIV-1/inhibitor complexes. While the modeling techniques and scoring protocols we describe here are general and should be applicable to most protein−ligand systems, the high quality of data in the HIV-1 protease system has attracted our interest for testing this model. Likewise, while we are using the HINT model for calculation of binding scores and energies, the basic concepts for including water in free energy estimates should be applicable with other methods.

Results

The key features of HIV protease are the "catalytic" water 300, observed in the unbound form of the enzyme and coordinated to the two catalytic residues Asp25 and Asp125 (Figure 1A), and a conserved water molecule, water 301, located on the HIV-1 protease symmetry axis (bridging the two subunits), hydrogen bonded to protease residues (Ile 50 and Ile150) and specific peptidic inhibitors (Figure 1B). Water 301 is observed in the free form of the enzyme, and except where it has been displaced deliberately, it has been observed in all HIV-1 protease-ligand complexes. The first generation HIV-1 protease inhibitors were designed around the shape of this active site with reference to water 301. In the current study we have analyzed 23 HIV-1 protease-inhibitor complexes of high quality three-dimensional structure and binding affinity (Chart 1 and Table 1). Several classes are represented: hydroxyethylenic ligands (**1**−**12**), peptidomimetic diol derivatives (**13**-**16**), cyclic ureidic derivatives (**18**-**20**), and cyclic sulfamide derivatives (**21**-**23**). The cyclic inhibitors (**18**−**23**) were designed to displace water 301,[38,39] i.e., the ureidic ligands place a carbonyl oxygen in the position usually occupied by the oxygen of water 301 and thus form hydrogen bonds with Ile50 and Ile150.

In some, but not all, HIV-1 protease complexes crystallographic analyses detected two pairs of largely conserved water molecules of interest for this study. Water 313, as named by Jhoti and colleagues,[51] is located by X-ray in several analyzed complexes and appears to be analogous to a water molecule reported in the unbound enzyme (see Figure 1). We have named the water in the pseudo-symmetric position water 313′. Compared to water 301, waters 313/313′ are located in a more peripheral area of the active site near the salt bridge between Asp29 (Asp129) and Arg108 (Arg8) and interacting with both the protein and ligand(s). We have named the other water molecules consistently inside the binding pocket water 313bis and the symmetry-related water 313bis′. These two solvent molecules occupy a more peripheral active site region (Figure 1B) than water 313/water 313′. They are much closer to the protein than to the ligand and interact strongly with residues Arg87, Thr26 (water 313bis), and Arg187, Thr126 (water 313bis′). In some structures a number of other water molecules are reported at the periphery of the active site while still bound both to ligand and protein residues. Comparing the three-dimensional structures of the free and bound forms of the enzyme indicates that some additional water molecules are totally displaced to bulk upon ligand binding (see Figures 1A and 1B). However, water molecules were not always present or reported in the structures of the complexes we analyzed, surprisingly irrespective of the crystallographic resolution. Thus, we applied the software GRID[14] to both analyze the potential sites for

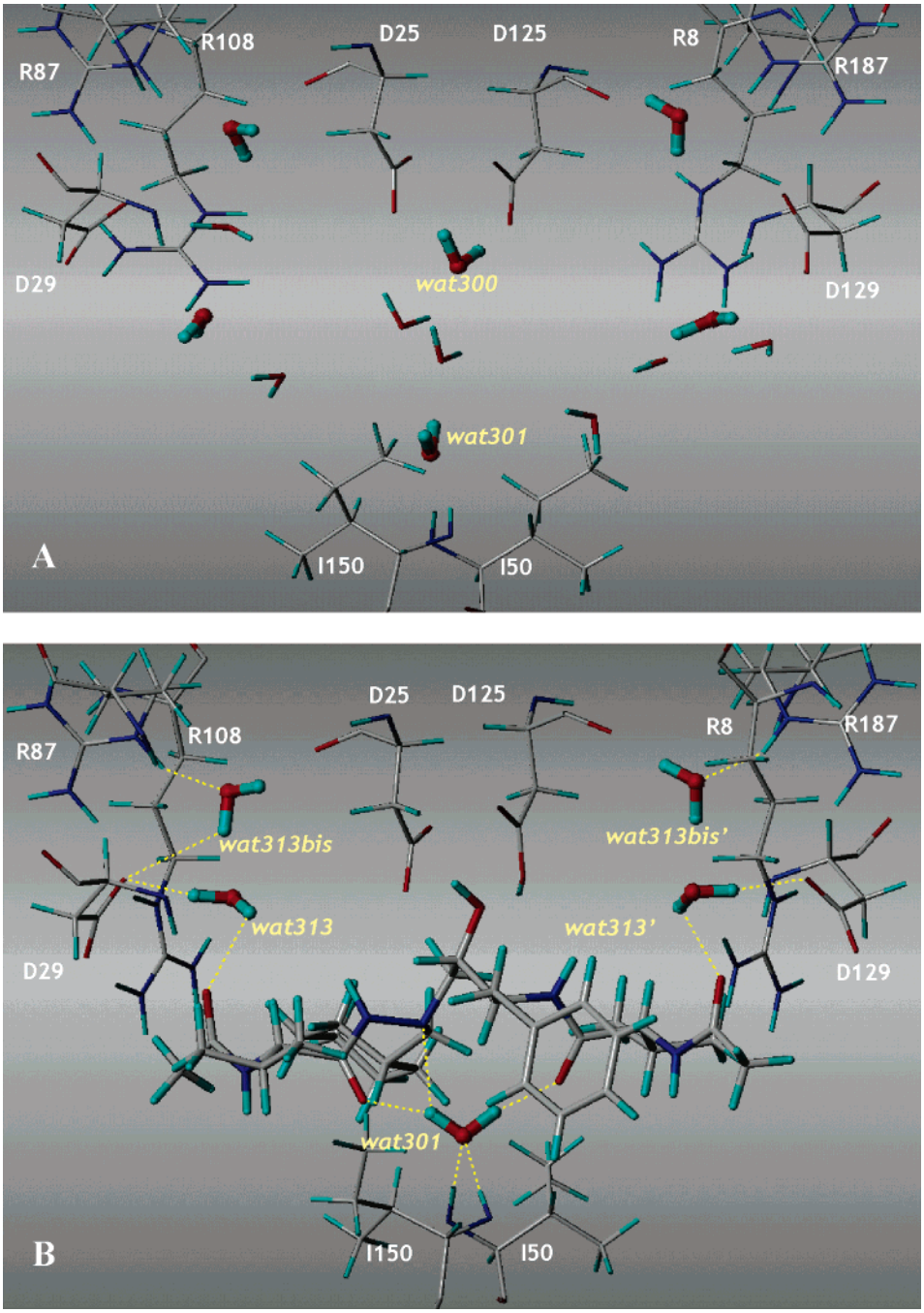
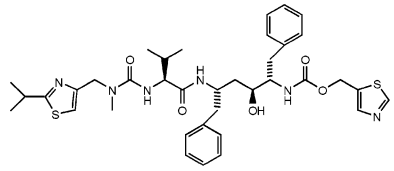


Figure 1. Structure of HIV-1 protease active site. (A) Free form of the enzyme (pdb code 1g6l[71]). The structural water molecules 300 and 301 and the protein residues from the two subunits are highlighted. (B) Enzyme complexed with the pseudosymmetric inhibitor **11.**[62] The structural water molecules at the binding site are highlighted. Yellow dotted lines represent hydrogen bonds between water and ligand, and water and protein.

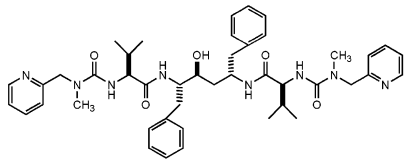
solvent molecules in the free and ligand−bound form of the enzyme and to validate the waters that were experimentally determined. The probability density for water defined by GRID matches the location of nearly all of the structured water molecules, as has been reported.[40−42] Thus, we have a reasonable degree of confidence when placing water molecules using GRID in the HIV-1 active site in cases where the experimental crystallographic structures were incomplete (see Materials and Methods). Modeling of the catalytic (Asp25/Asp125) dyad, based on careful analysis of the aspartates, their possible interactions, their reported pK_a values, and nuclear magnetic resonance studies of HIV inhibitor complexes,[43−44] indicated that the catalytic dyad was generally singly protonated. Each water molecule of interest was optimized with respect to both protein and ligand to maximize hydrogen bonds (vide infra).[45]

The interactions of each ligand with the protein were evaluated by the empirical HINT "force field".[17] The difficult-to-characterize bulk effects of solvent are implicitly encoded in the HINT parameters. However, constrained individual solvent molecules bridging protein and ligands need to be considered explicitly. The total HINT score for the complex interaction is given by the sum of the contributions resulting from protein−

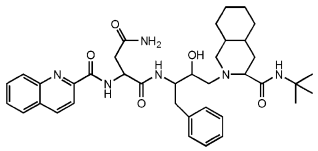
Chart 1. Structures of 23 HIV-1 Inhibitors



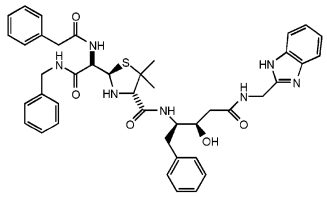
1 (ABT-538)



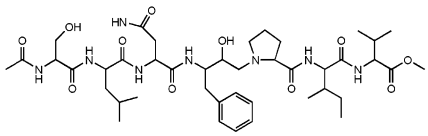
2 (A-78791)



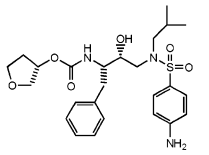
3 (Ro 31-8959)



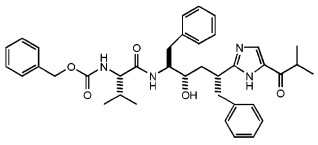
4 (GR137615)



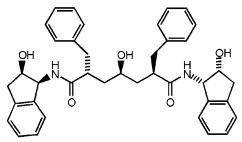
5 (JG-365)



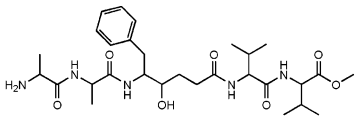
6 (VX-478)



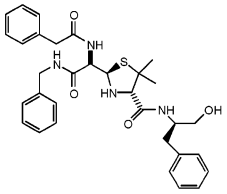
7 (SB 206343)



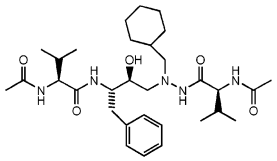
8 (L-700,417)



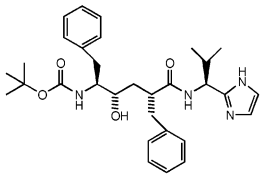
9 (Ala-Ala-Phe-CH(OH)CH_2-Gly-Val-Val-Ome)



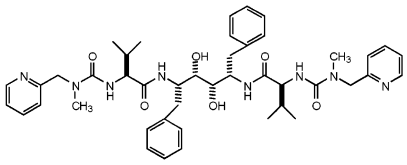
10 (GR126045)



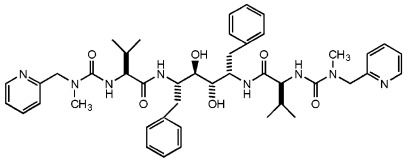
11 (CGP 53820)



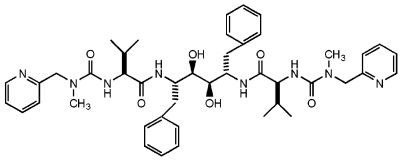
12 (SB 203386)



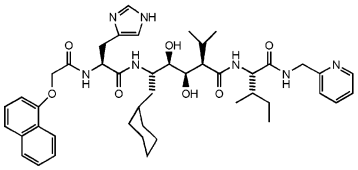
13 (A-76928)



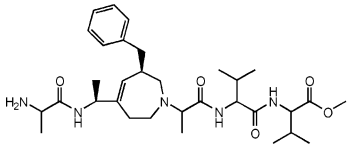
14 (A-77003)



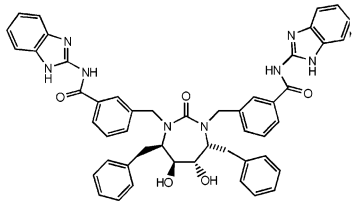
15 (A-76889)



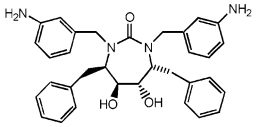
16 (U-75875)



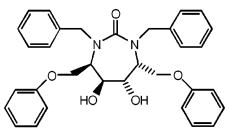
17 (SB203238)



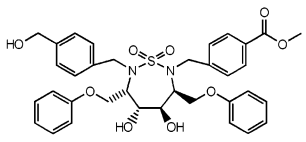
18 (SD146)



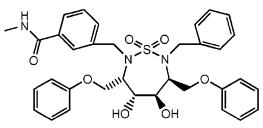
19 (DMP450)



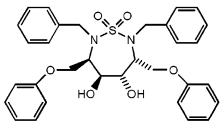
20 (Aha001)



21 (Aha024)



22 (Aha047)



23 (Aha006)

Table 1. HINT Scores for HIV-1 Protease Inhibitor−Water Complexes

PDB (ligand)	$ref_{binding}/ref_{crystal}$	$\Delta G_{binding}$ (kcal mol^{-1})	crystal resolution (Å)	crystal reported waters[a]	$H_{pro-lig}$	water 301[b]	water 313		water 313′		water 313bis		water 313bis′	
						$H_{lig-wat}$	$H_{lig-wat}$	$H_{pro-wat}$	$H_{lig-wat}$	$H_{pro-wat}$	$H_{lig-wat}$	$H_{pro-wat}$	$H_{lig-wat}$	$H_{pro-wat}$
1HXW (**1**)	55	−14.71	1.80	100	3607	1454	578	486	128	598	52	552	15	556
1HVJ (**2**)	56	−14.25	2.00	1	3460	1203	459	509	-	-	41	654	93	520
1HXB (**3**)	57	−13.49	2.30	102	3135	1049	689	587	271	367	51	571	2	656
1HTG (**4**)[c]	51	−13.20[e]	2.00	124	4226	1272	595	406	187	612	51	535	18	490
7HVP (**5**)[d]	49	−13.11	2.40	95	4311	1229	337	387	317	526	37	519	27	689
1HPV (**6**)	58	−12.57	1.90	80	3080	1058	115	424	132	287	4	594	3	515
1HPS (**7**)	59	−12.57	2.30	15	3124	829	715	596	502	274	117	395	40	496
4PHV (**8**)[c]	60	−12.51[f]	2.10	104	3932	789	183	543	195	623	25	569	11	516
1AAQ (**9**)	61	−11.45	2.50	1	3416	633	244	679	575	528	29	531	63	408
1HTF (**10**)[c]	51	−11.04[e]	2.20	94	2641	726	476	895	31	572	34	476	2	538
1HIH (**11**)	62	−10.97	2.20	145	3210	1080	197	603	288	536	21	566	27	609
1SBG (**12**)	46	−10.56	2.30	23	3037	1126	277	282	205	594	164	633	10	551
1HVK (**13**)	56	−13.80	1.80	1	3935	1064	-	-	-	-	77	551	83	566
1HVI (**14**)	56	−13.74	1.80	1	3734	1211	451	466	-	-	46	627	84	532
1HVL (**15**)	56	−12.27	1.80	1	3415	1253	-	-	-	-	61	627	92	625
1HIV (**16**)	63/64	−12.27[f]	2.00	90	3660	1326	528	547	381	556	49	621	41	644
1HBV (**17**)	65	−8.68	2.30	33	2042	777	-	-	526	529	42	665	17	756
1QBT (**18**)	66	−14.44	2.10	0	5170	-	656	376	1090	242	22	677	23	584
1DMP (**19**)	67	−12.99	2.00	0	4988	-	615	349	503	546	9	508	120	650
1AJX (**20**)	68/69	−10.79	2.00	112	3357	-	251	121	234	319	10	620	6	601
1G35 (**21**)	39	−11.06	1.80	143	4198	-	130	360	223[g]	564[g]	9	606	5	579
1G2K (**22**)	39	−10.82	1.95	72	3525	-	210	437	288	483	19	612	5	612
1AJV (**23**)	68/69	−10.52	2.00	75	3916	-	216	146	260	477	19	679	6	643

[a] Number of crystallographic water molecules reported in the pdb files of the HIV-1 protease−ligand complexes. [b] HINT scores for protein−water 301 interactions are not shown. Hydrogen bonds involving protein backbone amide nitrogens are systematically underestimated in the HINT force field. The average N_{amide} (protein)−O (water 301) distance of 3.02 ± 0.14 Å indicates that these interaction scores would be very consistent. [c] Ligand crystallized in two symmetrical orientations, occupying symmetrical sides of the binding pocket; HINT analysis was performed on models for both ligand orientations, but water analysis was carried out only on the model with the highest HINT score for protein−ligand interaction. [d] $pH_{binding}$ for this complex is 6.5; models corresponding to both deprotonation and monoprotonation of the catalytic dyad Asp25/125 were built, and the HINT score mean value is reported. [e] Ref 70. [f] Ref 37. [g] Water 313′ is found in two different crystallographic positions[39] with high B factors for both positions. Our analysis considered both water molecules to have 50% site occupancy; reported HINT scores are the means of resulting scores.

ligand, ligand−water, and protein-water interactions (eq 1)

$$H_{TOTAL} = H_{protein-ligand} + H_{ligand-water} [+ H_{protein-water}] \quad (1)$$

where H_{TOTAL} is the total HINT score and both $H_{ligand-water}$ and $H_{protein-water}$ are calculated for each analyzed water molecule. If we assume that all water molecules located in the active site of ligand−protein complexes are preexisting, i.e., part of the protein, then the latter (bracketed) term can be ignored (see Discussion). First, the $H_{protein-ligand}$ terms of the HINT interaction scores for the 23 HIV-1 protease−inhibitor complexes were calculated, and the correlation between these scores and experimental binding free energies is shown in Figure 2A. The related regression equation is

$$\Delta G = -0.0012 H_{TOTAL} - 7.903 \quad (2)$$

with a relatively poor $r^2 = 0.30$ and standard error of ± 1.30 kcal mol^{-1}. Next, as described in the following sections, the contributions to binding free energy by each of the five conserved and semiconserved water molecules at the active site (wat301, wat313, wat313′, wat313bis, wat313bis′) were evaluated for the complexes with **1**−**23**. Table 1 indicates the individual contributions of these water molecules to the total HINT score for each complex.

Water 301. The function and relevance of water 301 in ligand binding recognition has been investigated thoroughly.[27,46−50] Its highly conserved nature is reflected in low crystallographic B factor values and, in our hands, GRID always identified the region occupied by water 301 as energetically favorable for water placement. Because water 301 is so tightly associated with the protein, we are using the form of eq 1 that ignores the contribution of $H_{protein-water}$. In the Discussion we will debate this issue in more depth because the choice may not be quite so obvious for the other water molecules. The total HINT score for the 23 HIV-1 protease−inhibitor complexes including the contribution of water 301 provides the following regression equation:

$$\Delta G = -0.0017 H_{TOTAL} - 4.789 \quad (3)$$

with $r^2 = 0.63$ and standard error of ± 0.95 kcal mol^{-1} (Figure 2B). *P*, the *t*-test probability that adding the water 301 term to the model is insignificant, is small (<0.0005). Interestingly, the six cyclic urea and sulfamide-based inhibitors (**18**−**23**, shown as open circles in Figure 2B; dashed regression line: $r^2 = 0.82$), where water 301 has been by design displaced deliberately, closely match the experimental free energy−HINT score regression line for the entire data set (solid line). In effect, while these six ligands are tighter binders, the net energy benefit appears to be only that of the removed water−ligand interaction. Also, examination of the correlation of the cyclic urea and sulfamide derived inhibitors nearly perfectly predicts the free energy correction due to water 301.

Waters 313/313′. Not all of the crystallographic structures we examined revealed the presence of waters 313 and/or 313′. The H_2O probe of GRID was used to evaluate the spatial and energetic limits of the binding pockets and indicated whether it was possible to add waters 313 and/or 313′ to complexes when they were not experimentally reported (see Materials and Methods). HINT analysis (Table 1) of each of the putative water 313/313′ molecules evaluates the chemical nature

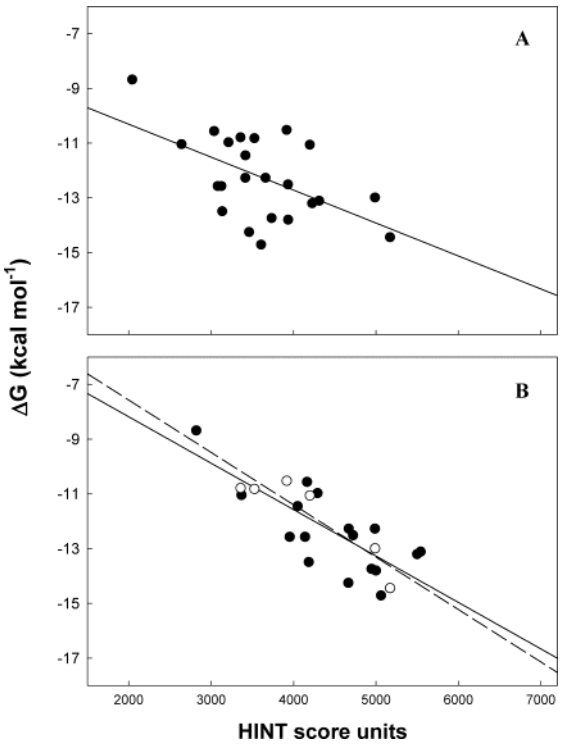


Figure 2. Correlation of HINT score with experimental free energy of binding for 23 HIV-1 protease−inhibitor complexes. (A) Score for inhibitors without considering the contribution of conserved water molecules. The solid line is the least-squares fit by eq 2. (B) Score for 23 inhibitors including the contribution of the highly conserved water 301 (closed circles); six cyclic urea and sulfamide inhibitors designed to displace 301 are indicated (open circles). The solid line is the least-squares fit by eq 3 of all inhibitors (including HINT score contribution of water 301 when appropriate), while the dashed line indicates only the six cyclic urea and sulfamide compounds.

of the interactions and confirmed the GRID-based placements. While water 301 always interacts with the amide hydrogens of Ile50/Ile150 and carbonyl oxygens of the inhibitors, waters 313/313′ interact with the terminal groups of the ligands. Thus, the type and strength of interactions is highly dependent on the structure of the ligand. Waters 313/313′ are highly mobile, occupying different positions depending on the nature of the bound ligand and can assume positions varying by as much as 3.0−3.5 Å. The role of waters 313/313′ can be partially revealed by examination of their HINT scores (Table 1). Both interact more strongly with the protein (313: average score 460 ± 177; 313′: average score 486 ± 124) than with the ligand (313: average score 396 ± 201; 313′ average score 333 ± 235). Also note the smaller variance for the water−protein interactions compared to the water−ligand interactions.

When the contribution of the ligand−water 313 interaction is added to the total HINT score, the expression becomes $H_{TOTAL} = H_{protein-ligand} + H_{ligand-water301} + H_{ligand-water313}$ and the relationship between experimental binding free energies and computational scores for the complexes follows the equation

$$\Delta G = -0.0016H_{TOTAL} - 4.873 \qquad (4)$$

where r^2 is 0.71, the standard error is ± 0.84 kcal mol^{-1}, and $P = 0.139$. Adding the contribution of water 313′ to the total HINT score, $H_{TOTAL} = H_{protein-ligand} + H_{ligand-water301} + H_{ligand-water313} + H_{ligand-water313'}$, yields the correlation between experimental free energy and HINT score:

$$\Delta G = -0.0013H_{TOTAL} - 5.515 \qquad (5)$$

with $r^2 = 0.60$, standard error of ± 0.99 kcal mol^{-1}, and $P = 0.281$.

Waters 313bis/313bis′. Waters 313bis and 313bis′ are both identified crystallographically in nearly all of the 23 complexes except where *only* water 301 or no waters are reported crystallographically, and in **7** where only water 313bis is reported. These molecules are more fixed than water 313/water 313′ with generally smaller *B* factors and more consistent interaction scores with both the ligand and protein (Table 1). They are, in fact, quite strongly hydrogen bonded to the protein (average scores for water 313bis and water 313bis′ are 582 ± 69 and 579 ± 76, respectively), while weakly interacting with all ligands (average scores for water 313bis and water 313bis′ are 43 ± 37 and 34 ± 36, respectively). Water molecules in a similar position are also present in the three-dimensional structure of the free form of HIV-1 (Figure 1A), with a minor shift of only about 0.6−0.8 Å from the unbound to the bound form of the enzyme. Only in the complex between HIV-1 protease and **12**,[46] where water 313 is shifted toward the ligand, does water 313bis occupy a different position (movement of about 2.7 Å). Waters 313 and 313′ are usually able to interact with waters 313bis and 313bis′, respectively, creating a network of two water molecules on each side of the active site bridging protein and ligand.

The HINT scores for the interaction between ligands and waters 313bis/313bis′ in the analyzed complexes are small (Table 1), and their contributions do not significantly impact the total HINT score ($H_{TOTAL} = H_{protein-ligand} + H_{ligand-water301} + H_{ligand-water313} + H_{ligand-water313'} + H_{ligand-water313bis} + H_{ligand-water313bis'}$). The correlation between Δ*G* and HINT score indicates the regression equation:

$$\Delta G = -0.0013H_{TOTAL} - 5.422 \qquad (6)$$

with $r^2 = 0.61$ and standard error of ± 0.98 kcal mol^{-1}, not dissimilar to (eq 5) before waters 313bis and 313bis′ were included. The *t*-test probability that the addition of the water 313bis and water 313bis′ terms is insignificant is correspondingly large ($P = 0.782$). This will be an element of the Discussion.

Other Water Molecules. Water molecules can be incorporated in a binding site in a number of different ways and their mode of binding and their role in protein ligand interaction can change with their positions.[4,13,52] There are a number of "peripheral" water molecules, e.g., at the entrance to the binding pocket, which apparently bridge protein and ligand for HIV-1 protease-ligand complexes. These are usually (1) not confirmed by GRID; (2) have relatively large *B* factors; and (3) have small water−ligand HINT scores. We classify these as casual waters of solvation. In effect they may have taken up their structural positions after the ligand binding event has occurred. The problem of identifying which water molecules *are* important for binding will be addressed below in the Discussion section.

Discussion

The process of ligand binding in a protein active site is driven by (i) the interaction between complementary functional groups (purely enthalpy) and (ii) the release of water molecules from ligand and protein to bulk (mostly entropy). However, active sites may contain other compounds, such as cofactors or coenzymes as well as structural (conserved) water molecules that can simultaneously interact with protein residues and ligand functional groups by bridging between them. Of particular interest to us are structural water molecules that may make favorable hydrogen bonds with ligands, thus significantly contributing to the energetics of binding. Due to the variety of complex roles that (water) solvent plays in biological systems, entropy is especially difficult to estimate computationally in biological systems. Even at the limit of explicit solvation, Newtonian mechanics-based computational methods have not, to date, produced de novo the hydrophobic effect, and quantum mechanics simulations of the magnitude necessary for complex biological systems are still unattainable.

To overcome these limitations we have been using HINT, which, because it is derived from a free energy experiment of solvation, implicitly includes entropic contributions arising from water molecules surrounding the ligand/protein and displaced in complex formation. Our previous successes with using this tool to predict free energy[20−24] have confirmed our assertions that entropy is implicitly included in the HINT model.[17] Up until now we have largely focused on cases where we believed that (ordered) water molecules at the ligand binding site were not substantially influencing the energetics of binding. However, water molecules that bridge ligand and protein need to be accounted for in free energy predictions. The question is *how* to account for them. Above, we have described a simple computational protocol for correcting binding free energy predictions by incorporating the "binding energy" of water molecules.

Are Protein−Water Interactions Significant? As we introduced eq 1 above, we noted that it may be necessary to add terms for both ligand−water and protein−water interactions. In this work we have chosen to consider bridging water molecules as (only) an integral part of the protein surface and ignore the protein−water interactions. This is likely to be the general case and can be justified as follows: first, studies have shown that structural water molecules located in the active site of protein−ligand complexes are almost invariably also present in the free protein forms.[26] This is definitely the case with the HIV-1 protease complexes. Comparisons of free and ligand-bound proteins show that the number of active site water molecules usually decreases with ligand binding, i.e., ligand-coordinated water molecules are generally released. Second, in this set of HIV-1 protease complexes and, likely, in general, the water-to-protein interactions are mostly invariant across the data series, especially for water 301 and waters 313bis/313bis′ (Table 1). The average scores and standard deviations reported above are evidence for this claim. While there is *some* useful information in the protein−water interaction scores, there is also significant noise. The systematic uncertainty in HINT scores, arising from crystallographic effects, $LogP_{o/w}$ measurement errors and numerous other potential computational sources, is around ±50−100 score units.[53] Thus, a potentially significant effect of including the protein−water interaction scores is the propagation of this uncertainty.

Should We Consider only Crystallographic Water Molecules or All Probable Waters? Despite the relatively high resolution of the crystallographic experiments leading to the complex structures analyzed in this report, a highly variable number of water molecules were reported in the PDB files (Table 1). In some cases zero or only one water (301) was reported, even though the structures were determined at a resolution high enough to describe the water network. In other cases 100+ water positions were reported in the crystal structures. Is it most correct to include the contribution to binding of only crystallographic waters, or of all waters including those added through GRID analyses? When the crystallographic analysis does not locate any water molecules (HIV-1 protease complexes with **18** and **19** that displace water 301), or locate only water 301 (complexes with **2**, **9**, **13**-**15**), GRID was instrumental for placing water molecules in the active site. In contrast, complexes with **3** where 102 waters were found, with **6** where 80 waters were found, or with **22** where 72 waters were found (Table 1), were also supplemented with additional water molecules suggested by GRID.

While it is likely that some of these waters were not reported in the crystallography because they are really not present, for consistency we have chosen to add all water molecules that GRID locates. We have used this approach because GRID has been validated, in this study and others,[14,40−42] to reproduce crystallographically determined water positions. If we consider in the HINT analyses *only* the crystallographic water molecules 301, 313, 313′, 313bis, and 313bis′, the correlation between experimental free energy and HINT score exhibits an r^2 of 0.39 and standard error of ±1.22 kcal mol^{-1}. This is clearly inferior to the correlation reported above (eq 6), when *all* crystallographic and GRID water molecules were included in the model.

Role of Waters 301, 313, 313′, 313bis, and 313bis′. From this and other analyses, the relevance of water 301 in the binding process of HIV is extremely evident. In fact, our estimation of its binding energy is about 4−6 kcal mol^{-1}, of a similar order of what Wade et al.[54] reported (10.0 kcal mol^{-1}) for a water molecule tightly bound in a buried cavity of sulfate-binding protein. The energetic contribution of water 301 to binding leads to a better correlation between experimental free energy and HINT score, thus improving binding free energy prediction (Figure 2B). Other waters, located in different regions of the active site, play different roles in the binding process. The contribution of water 313 and water 313′ to the total HINT score is variable (Table 1), depending strongly on the chemical nature and size of the ligand, and on which side of the symmetrical binding pocket the ligand is bound, but is for the most part statistically valid. For large and pseudo-symmetrical ligands that fill the pocket and present functional groups that make similar interactions with the two symmetrical waters, the energetic contribution of these two solvent molecules is also similar. In contrast,

when the inhibitors are smaller and largely occupy only one side of the active site, the contributions from waters 313 and 313′ are also asymmetric. Interestingly, taking into account the contribution of water 313 leads to an improvement in the correlation between experimental and calculated data, but the quality of the correlation slightly degrades when the contribution of water 313′ is also considered. The inclusion of the more peripheral waters 313bis and 313bis′ in H_{TOTAL} does not noticeably improve the correlation and are shown to be statistically insignificant, indicating that these two water molecules really are best considered as being part of the protein active site surface. Thus, it may appear that the only water molecule really impacting the HINT score and the free energy correlation is water 301. Another factor, the growing uncertainty arising from adding more terms to H_{TOTAL}, also must be considered. Because the HINT score uncertainty for an interaction is on the order of 50−100 score units (around 0.1−0.2 kcal mol^{-1}[21]), the propagation of this uncertainty, i.e., noise, will dominate score correlations if enough terms are added to H_{TOTAL}. These cumulative systematic uncertainties from HINT score calculations will eventually dominate over the statistical uncertainty (standard error) of the correlations. Also, we have previously stated[21] that the experimental uncertainties of binding measurements (and crystallography) suggest that predicting ΔG with more confidence than about ±1.0 kcal mol^{-1} is unlikely for data collected in multiple laboratories. This is certainly evident with this set of data where the "best" standard error is ±0.84 kcal mol^{-1} before deteriorating with what appears to be error propagation concomitant with additional terms in H_{TOTAL}.

Conclusions. This study on 23 HIV-1 protease−inhibitor complexes clearly indicates the value of explicitly modeling water at the protein active site. A spectrum of roles for water in the active site from structurally required water molecules, such as water 301, through waters that inhabit the binding pocket and somewhat shape the binding site, such as the water 313/313′ pair and to a lesser extent the 313bis/313bis′ pair, to water molecules that are clearly just opportunistic waters of solvation, has been described. The relative differences between HINT scores and uncertainties provide insight into water roles, and while GRID[14] is also a valuable tool for characterizing the role of water molecules, high quality and complete experimental data for conserved water molecules are paramount to truly understanding the effects of binding site water molecules. In this series of papers we have brought together two very important but often ignored properties of protein−ligand binding: ionization state of acidic or basic protein residues/ligand functional groups[22] and the role of solvent molecules in the active site (in this work), with our empirical HINT force field and model.[21]

Materials and Methods

Model Building. The three-dimensional structures of the 23 HIV protease−ligand complexes analyzed in the study were retrieved from the Protein Data Bank (www.rcsb.org) and read into the modeling program Sybyl (version 6.8, www.tripos.com). These models were processed and optimized as described previously.[21,22] The ionization states of polar groups at the protein binding site and on the ligand were examined, with particular attention on the catalytic Asp25/Asp125 dyad. Nuclear magnetic resonance studies of HIV protease−inhibitor complexes indicate that only one of the two aspartates of the dyad can be protonated in the pH range 2.5−6.5, i.e., pK_{a1} > 6.2 and pK_{a2} < 2.5.[43,44] As both the crystallographic and solution inhibition experiments were typically performed around pH 5−6 (see Table 2, Supporting Information), our models generally singly protonated the catalytic dyad. When the protonation of one aspartate or the other was ambiguous, models corresponding to both situations were built and scored, and the mean HINT score value has been reported.

The GRID program (www.moldiscovery.com) was used to both confirm crystallographically reported water molecules and to propose additional water molecule sites when necessary. The standard water probe was applied over the region of interest, typically a box with dimensions 24 × 24 × 24 Å centered on the ligand, with a grid spacing of 0.33 Å. Table 2 (Supporting Information) indicates the results of the GRID analyses by listing the specific water molecules that were placed as suggested by GRID or for which the site was too constricted for a water to be added. When water molecules were added, the coordinates were, whenever possible, taken from models that exhibit crystallographic waters and translated into the positions suggested by the density contours of GRID. Special considerations were made for the addition of waters 313 and 313′ as the expected binding region for these waters varied considerably as a function of ligand. In four cases, Arg8 and Asp129 assume conformations closer to the ligand than the symmetric Arg108 and Asp29, thus apparently preventing water binding at the 313′ site, while allowing water binding at the 313 site only in two of the four complexes. In general, when present, water 313 and water 313′ are *both* detected by X-ray diffraction studies. There are three exceptions: in two (**12** and **22**) only water 313 was crystallographically present, while in one complex (**6**) only water 313′ was present. [It may be argued that this distinction between 313 and 313′ in these unsymmetrical cases is only due to nomenclature/labeling of the atoms and residues as built during solution of the structure, but we have retained the original numbering convention.]

All water molecules of interest in the binding site were optimized with the HINT "optimize ligand" option.[45] This procedure allows for the optimization of water positions with respect to protein and ligand by maximizing hydrogen bond formation with the protein/ligand and minimizing hydrophobic−polar interactions. Water oxygen atoms were allowed limited (≤0.5 Å) translation during optimization depending on the nature (crystallographic, GRID, etc.) of the optimized water molecules.

Hydropathic Analysis. The software HINT (version 2.35S, www.tripos.com) was applied to these 23 protein−ligand complexes using methodology previously reported.[21,22] Because the focus in this work is on polar (hydrogen bonding) interactions, the "all" option for partitioning the protein and ligand molecules was used. This option assigns a partial Log$P_{o/w}$ (a_i) and solvent accessible surface area (S_i) to each atom, including nonpolar hydrogens. The HINT scores reported in the table are the sums of all atom−atom interactions.[17] Previous work[21,22,24] has shown that about −515 HINT score units corresponds to 1 kcal mol^{-1}.

Acknowledgment. For financial support we thank the Ministry of Instruction, University and Research (COFIN2003 and FIRB to A.M.), C.N.R. (Functional Genomics to A.M.), National Institute for the Physics of Matter (A.M.), University of Parma (P.C.), Virginia Commonwealth University (D.J.A. and G.E.K.), and the National Institutes of Health (D.J.A., R01HL32793-15). Prof. J. N. Scarsdale, Prof. J. P. Rife, and Prof. C. Rizzoli have participated in useful discussions throughout this research project. HINT was developed in the laboratory of G.E.K. and can be obtained from Tripos, Inc.

Supporting Information Available: A table with additional binding, crystallographic, and modeling data for 23

ligand−HIV-1 protease complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dill, K. A. Additivity Principles in Biochemistry. *J. Biol. Chem.* **1997**, *272*, 701−704.

(2) Gohlke, H.; Klebe G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2644−2676.

(3) Lazaridis, T. Binding affinity and specificity from computational studies. *Curr. Org. Chem.* **2002**, *6*, 1319−1332.

(4) Ladbury, J. E. Just add water! The effect of water on the specificity of protein−ligand binding sites and its potential application to drug design. *Chem. Biol.* **1996**, *3*, 973−980.

(5) Cozzini, P.; Fornabaio, M.; Marabotti, A.; Abraham, D. J.; Kellogg, G. E.; Mozzarelli, A. Free energy of ligand binding to protein: evaluation of the contribution of water molecules by computational methods. Review. *Curr. Med. Chem.* **2004**, in press.

(6) Lee, B.; Richards, F. M. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379−400.

(7) Eisenberg, D.; McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **1986**, *319*, 199−203.

(8) Wang, J.; Wang, W.; Huo, S.; Lee, B.; Kollman, P. Solvation Model Based on Weighted Solvent Accessible Surface Area. *J. Phys. Chem. B* **2001**, *105*, 5055−5067.

(9) Kollman, P. Free energy calculations-applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.

(10) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(11) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−897.

(12) Grant, J. A.; Pickup, B. T.; Nicholls, A. A smooth permittivity function for Poisson−Boltzmann solvation methods. *J. Comput. Chem.* **2001**, *22*, 608−640.

(13) Poornima, C. S.; Dean, P. M. Hydration in drug design. 2. Influence of local site surface shape on water binding. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 513−520.

(14) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules *J. Med. Chem.* **1985**, *28*, 849−857.

(15) Reddy, M. R.; Erion, M. D.; Agarwal, A. Free Energy Calculations: Use and Limitations in Predicting Ligand Binding Affinities. *Rev. Comput. Chem.* **2000**, *16*, 217−304.

(16) Wireko, F. C.; Kellogg, G. E.; Abraham, D. J. Allosteric modifiers of hemoglobin. 2. Crystallographically determined binding sites and hydrophobic binding/interaction analysis of novel hemoglobin oxygen effectors. *J. Med. Chem.* **1991**, *34*, 758−767.

(17) Kellogg, G. E.; Abraham, D. J. Hydrophobicity: is $LogP_{o/w}$ more than the Sum of its Parts? *Eur. J. Med. Chem.* **2000**, *35*, 651−661.

(18) Kellogg, G. E.; Burnett, J. C.; Abraham, D. J. Very Empirical Treatment of Solvation and Entropy: a Force Field Derived from Log $P_{o/w}$. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 381−393.

(19) Gussio, R.; Pattabiraman, N.; Zaharevitz, D. W.; Kellogg, G. E.; Topol, I. A.; Rice, W. G.; Schaeffer, C. A.; Erickson, J. W.; Burt, S. K. All-atom models for the non-nucleoside binding site of HIV-1 reverse transcriptase complexed with inhibitors: a 3D QSAR approach. *J. Med. Chem.* **1996**, *39*, 1645−1650.

(20) Gussio, R.; Zaharevitz, D. W.; McGrath, C. F.; Pattabiraman, N.; Kellogg, G. E.; Schultz, C.; Link, A.; Kunick, C.; Leost, M.; Meijer, L.; Sausville, E. A. Structure-Based Design Modifications of the Paullone Molecular Scaffold for Cyclin-Dependent Kinase Inhibition. *Anti-Cancer Drug Des.* **2000**, *15*, 53−66.

(21) Cozzini, P.; Fornabaio, M.; Marabotti, A.; Abraham, D. J.; Kellogg, G. E.; Mozzarelli, A. Simple, intuitive calculations of free energy of binding for protein−ligand complexes. 1. Models without explicit constrained water. *J. Med. Chem.* **2002**, *45*, 2469−2483.

(22) Fornabaio, M.; Cozzini, P.; Mozzarelli, A.; Abraham, D. J.; Kellogg, G. E. Simple, intuitive calculations of free energy of binding for protein−ligand complexes. 2. Computational Titration and pH Effects in Molecular Models of Neuraminidase-Inhibitor Complexes. *J. Med. Chem.* **2003**, *46*, 4487−4500.

(23) Burnett, J. C.; Kellogg, G. E.; Abraham, D. J. Computational Methodology for Estimating Changes in Free Energies of Biomolecular Association upon Mutation. The Importance of Bound Water in Dimer-Tetramer Assembly for β37 Mutant Hemoglobins. *Biochemistry* **2000**, *39*, 1622−1633.

(24) Burnett, J. C.; Botti, P.; Abraham, D. J.; Kellogg, G. E. Accessible Method for estimating Free Energy Changes Resulting from Site-Specific Mutations of Biomolecules: Systematic Model Building and Structural/Hydropathic Analysis of Deoxy and Oxy Hemoglobins. *Proteins: Struct. Funct. Genet.* **2001**, *42*, 355−377.

(25) Poornima, C. S.; Dean, P. M. Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein−ligand interactions. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 500−512.

(26) Poornima, C. S.; Dean, P. M. Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 521−531.

(27) Wlodawer, A.; Vondrasek, J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249−284.

(28) Navia, M. A.; Fitzgerald, P. M.; McKeever, B. M.; Leu, C. T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **1989**, *337*, 615−620.

(29) Abdel-Rahman, H. M.; Al-karamany, G. S.; El-Koussi, N. A.; Youssef, A. F.; Kiso, Y. HIV protease inhibitors: peptidomimetic drugs and future perspectives. *Curr. Med. Chem.* **2002**, *9*, 1905−1922.

(30) Venkatesan, N.; Kim, B. H. Synthesis and enzyme inhibitory activities of novel peptide isosteres. *Curr. Med. Chem.* **2002**, *9*, 2243−2270.

(31) Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M. D.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; Desolms, S. J.; Gaffin, N.; Ghosh, A. K.; Giuliani, E. A.; Graham, S. L.; Guare, J. P.; Hungate, R. W.; Lyle, T. A.; Sanders, W. M.; Tucker, T. J.; Wiggins, M.; Wiscount, C. M.; Woltersdorf, O. W.; Young, S. D.; Darke, P. L.; Zugay, J. A. A-priori Prediction of Activity for HIV-1 Protease Inhibitors Employing Energy Minimization in the Active Site. *J. Med. Chem.* **1995**, *38*, 305−317.

(32) Wallqvist, A.; Jernigan, R. L.; Covell, D. G. A preference-based free-energy parametrization of enzyme−inhibitor binding. Applications to HIV-1-protease inhibitor design. *Protein Sci.* **1995**, *4*, 1881−1903.

(33) Hansson, T.; Aqvist, J. Estimation of binding free energies for HIV proteinase inhibitors by molecular dynamics simulations. *Protein Eng.* **1995**, *8*, 1137−1144.

(34) Debnath, A. K. Three-dimensional quantitative structure−activity relationship study on cyclic urea derivatives as HIV-1 protease inhibitors: Application of comparative molecular field analysis. *J. Med. Chem.* **1999**, *42*, 249−259.

(35) Tawa, G. J.; Topol, I. A.; Burt, S. K.; Erickson, J. W. Calculation of Relative Binding Free Energies of Peptidic Inhibitors to HIV-1 Protease and Its I84V Mutant. *J. Am. Chem. Soc.* **1998**, *120*, 8856−8863.

(36) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; R., H.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(37) Österberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* **2002**, *46*, 34−40.

(38) Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; et al. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **1994**, *263*, 380−384.

(39) Schaal, W.; Karlsson, A.; Ahlsén, G.; Lindberg, J.; Andersson, H. O.; Danielson, U. H.; Classon, B.; Unge, T.; Samuelsson, B.; Hultén, J.; Hallberg, A.; Karlén, A. Synthesis and Comparative Molecular Field Analysis (CoMFA) of Symmetric and Nonsymmetric Cyclic Sulfamide HIV-1 Protease Inhibitors. *J. Med. Chem.* **2001**, *44*, 155−169.

(40) Wade, R. C.; Clark, K. J.; Goodford, P. J. Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 1. Ligand probe groups with the ability to form two hydrogen bonds. *J. Med. Chem.* **1993**, *36*, 140−147.

(41) Wade, R. C.; Goodford, P. J. Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 2. Ligand probe groups with the ability to form more than two hydrogen bonds. *J. Med. Chem.* **1993**, *36*, 148−156.

(42) Pastor, M.; Cruciani, G.; Watson, K. A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure−activity relationship analysis. *J. Med. Chem.* **1997**, *40*, 4089−4102.

(43) Smith, R.; Brerenton, I. M.; Chai, R. Y.; Kent, S. B. H. Ionization states of the catalytic residues in HIV-1 protease. *Nat. Struct. Biol.* **1996**, *3*, 946−950.
(44) Wang, Y. X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. Solution NMR Evidence That the HIV-1 Protease Catalytic Aspartyl Groups have Different Ionization States in the complex Formed with the Asymmetric Drug KNI-272. *Biochemistry* **1996**, *35*, 9945−9950.
(45) Kellogg, G. E.; Chen, D. The Importance of Being Exhaustive. Optimization of Bridging Structural Water Molecules and Water Networks in Models of Biological Systems. *Chem. Biodiversity* **2004**, *1*, 98−105.
(46) Abdel-Meguid, S. S.; Metcalf, B. W.; Carr, T. J.; Demarsh, P.; DesJarlais, R. L.; Fisher, S.; Green, D. W.; Ivanoff, L.; Lambert, D. M.; Murthy, K. H. M.; Petteway, S. R.; Pitts, W. J.; Tomaszek, T. A.; Winborne, E.; Zhao, B.; Dreyer, G. B.; Meek, T. D. An Orally Biovailable HIV-1 protease Inhibitor Containing an Imidazole-Derived Peptide Bond Replacement: Crystallographic and Pharmacokinetic Analysis. *Biochemistry* **1994**, *33*, 11671−11677.
(47) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B.; Wlodawer, A. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **1989**, *246*, 1149−1152.
(48) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M. Design, activity, and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. *Science* **1990**, *249*, 527−533.
(49) Swain, A. L.; Miller, M. M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8805−8809.
(50) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S. B.; Wlodawer, A. Structure at 2.5-Å resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor. *Biochemistry* **1991**, *30*, 1600−1609.
(51) Jhoti, H.; Singh, O. M. P.; Weir, M. P.; Cooke, R.; Murray-Rust, P.; Wonacott, A. X-ray Crystallographic Studies of a Series of Penicillin-Derived Asymmetric Inhibitors of HIV-1 Protease. *Biochemistry* **1994**, *33*, 8417−8427.
(52) Timasheff, S. N. Protein hydration, thermodynamic binding, and preferential hydration. *Biochemistry* **2002**, *41*, 13473−13482.
(53) Kellogg, G. E.; Scarsdale, J. N.; Fornari, F. A. Jr. Identification and Hydropathic Characterization of Structural Features Affecting Sequence Specificity for Doxorubicin Intercalation into DNA Double-Stranded Polynucleotides. *Nucleic Acids Res.* **1998**, *26*, 4721−4732.
(54) Wade, R. C.; Mazor, M. H.; McCammon, J. A.; Quiocho, F. A. A molecular dynamics study of thermodynamic and structural aspects of the hydration of cavities in proteins. *Biopolymers* **1991**, *31*, 919−931.
(55) Kempf, D. J.; Marsh, K. C.; Denissen, J. F.; McDonald, E.; Vasavanonda, S.; Flentge, C. A.; Green, B. E.; Fino, L.; Park, C. H.; Kong, X. P.; Wideburg, N. E.; Saldivar, A.; Ruiz, L.; Kati, W. M.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. ABT-538 is a potent inhibitor of human immunodefiency virus protease and has high oral bioavailabylity in humans. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2484−2488.
(56) Hosur, M. V.; Bhat, T. N.; Kempf, D. J.; Baldwin, E. T.; Liu, B.; Gulnik, S.; Wideburg, N. E.; Norbeck, D. W.; Appelt, K.; Erickson, J. W. Influence of Stereochemistry on Activity and Binding Modes for C_2 Symmetry-Based Diol Inhibitors of HIV-1 Protease. *J. Am. Chem. Soc.* **1994**, *116*, 847−855.
(57) Krohn, A.; Redshaw, S.; Ritchie, J. C.; Graves, B. J.; Hatada, M. H. Novel Binding Mode of Highly Potent HIV−Proteinase Inhibitors Incorporating the (*R*)-Hydroxyethylamine Isostere. *J. Med. Chem.* **1991**, *34*, 3340−3342.
(58) Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. Crystal Structure of HIV-1 Protease in Complex with VX-478, a Potent and Orally Bioavailable Inhibitor of the Enzyme. *J. Am. Chem. Soc.* **1995**, *117*, 1181−1182.
(59) Thompson, S. K.; Murthy, K. H. M.; Zhao, B.; Winborne, E.; Green, D. W.; Fisher, S. M.; DesJarlais, R. L.; Tomaszek, T. A.; Meek, T. D.; Gleason, J. G.; Abdel-Meguid, S. S. Rational Design, Synthesis, and Crystallographic Analysis of a Hydroxyethylene-Based HIV-1 Protease Inhibitor Containing a Heterocyclic P_1'-P_2' Amide Bond isostere. *J. Med. Chem.* **1994**, *37*, 3100−3107.
(60) Bone, R.; Vacca, J. P.; Anderson, P. S.; Holloway, M. K. X-ray Crystal Structure of the HIV Protease Complex with L-700,417, an Inhibitor with Pseudo C_2 Symmetry. *J. Am. Chem. Soc.* **1991**, *113*, 9382−9384.
(61) Dreyer, G. B.; Lambert, D. M.; Meek, T. D.; Carr, T. J.; Tomaszek, T. A.; Fernandez, A. V.; Bartus, H.; Cacciavillani, E.; Hassell, A. M.; Minnich, M.; Petteway, S. R.; Metcalf, B. W. Hydroxyethilene Isostere Inhibitors of Human Immunodeficiency Virus-1 Protease: Structure−Activity Analysis Using Enzyme Kinetics, X-ray Cristallography, and Infected T-Cell Assays. *Biochemistry* **1992**, *31*, 6646−6659.
(62) Priestle, J. P.; Fassler, A.; Rosel, J.; Tiltelnot-Blomley, M.; Strop, P.; Grotter, M. G. Comparative Analysis of the X-ray Structures of HIV-1 and HIV-2 Proteases in Complex with CGP 53820, a novel pseudosymmetric Inhibitor. *Structure* **1995**, *3*, 381−389.
(63) Ashorn, P.; McQuade, T. J.; Thaisrivongs, S.; Tomasselli, A. G.; Tarpley, W. G.; Moss, B. An inhibitor of the protease blocks maturation of human and simian immunodeficiency viruses and spread of infection. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7472−7476.
(64) Thanki, N.; Rao, J. K. M.; Foundling, S. I.; Howe, W. J.; Moon, J. B.; Hui, J. O.; Tomasselli, A. G.; Heinrikson, R. L.; Thaisrivongs, S. Wlodawer, A. Crystal structure of a complex of HIV-1 protease with a dihydroxyethylene-containing inhibitor: Comparisons with molecular modeling. *Protein Sci.* **1992**, *1*, 1061−1072.
(65) Hoog, S. S.; Zhao, B.; Winborne, E.; Fisher, S.; Green, D. W.; DesJarlais, R. L.; Newlander, K. A.; Callahan, J. F.; Moore, M. L.; Huffman, W. F.; Abdel-Meguid, S. S. A Check on Rational Drug Design: Crystal Structure of a Complex of Human Immunodeficiency Virus Type 1 Protease with a Novel γ-Turn Mimetic Inhibitor. *J. Med. Chem.* **1995**, *38*, 3246−3252.
(66) Jadhav, P. K.; Ala, P.; Woerner, F. J.; Chang, C. H.; Garber, S. S.; Anton, E. D.; Bacheler, L. T. Cyclic Urea Amides: HIV-1 protease Inhibitors with Low nanomolar Potency against both Wild-Type and Protease Inhibitors Resistant Mutants of HIV. *J. Med. Chem.* **1997**, *40*, 181−191.
(67) Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C. H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y. S.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. S.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. Improved cyclic urea inhibitors of the HIV-1 protease: synthesis, potency, resistance profile, human pharmacokinetics and X-ray crystal structure of DMP 450. *Chem. Biol.* **1996**, *3*, 301−314.
(68) Hulten, J.; Bonham, N. M.; Nillroth, U.; Hansson, T.; Zuccarello, G.; Bouzide, A.; Aqvist, J.; Classon, B.; Danielson, U. H.; Karlen, A.; Kvarnstrom, I.; Samuelsson, B.; Hallberg, A. Cyclic HIV-1 protease inhibitors derived from mannitol: synthesis, inhibitory potencies, and computational predictions of binding affinities. *J. Med. Chem.* **1997**, *40*, 885−897.
(69) Bäckbro, K.; Löwgren, S.; Österlund, K.; Atepo, J.; Unge, T. Unexpected Binding Mode of a Cyclic Sulfamide HIV-1 Protease Inhibitor. *J. Med. Chem.* **1997**, *40*, 898−902.
(70) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical Scoring Functions: I. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.
(71) Pillai, B.; Kannan, K. K.; Hosur, M. V. 1.9 A X-ray study shows closed flap conformation in crystals of tethered HIV-1 PR. *Proteins* **2001**, *43*, 57−64.

JM030596B