Thermodynamics of Buried Water Clusters at a Protein−Ligand Binding Interface

Zheng Li and Themis Lazaridis*

Department of Chemistry, City College of New York/CUNY, Convent Ave & 138th Street, New York, New York 10031

Received: October 20, 2005

The structure of the complex of cyclophilin A (CypA) with cyclosporin A (CsA, **1**) shows a cluster of four water molecules buried at the binding interface, which is rearranged when CsA is replaced by (5-hydroxynorvaline)-2-cyclosporin (**2**). The thermodynamic contributions of each bound water molecule in the two complexes are explored with the inhomogeneous fluid solvation theory and molecular dynamics simulations. Water (WTR) 133 in complex **1** contributes little to the binding affinity, while WTR6 and 7 in complex **2** play an essential role in mediating protein−ligand binding with a hydrogen bond network. The calculations reveal that the rearrangement of the water molecules contributes favorably to the binding affinity, even though one of them is displaced going from ligand **1** to **2**. Another favorable contribution comes from the larger protein−ligand interactions of ligand **2**. However, these favorable contributions are not sufficient to overcome the unfavorable desolvation free energy change and the conformational entropy of the hydroxylpropyl group of ligand **2** in the complex, leading to a lower binding affinity of ligand **2**. These physical insights may be useful in the development of improved scoring functions for binding affinity prediction.

Introduction

Interactions at the binding interface of biomolecular complexes are often mediated by bound water molecules.[1−12] Several studies have demonstrated the importance of taking such water molecules into account in ligand design.[8,13−17] The displacement of ordered water molecules is sometimes used as a strategy for designing ligands with higher binding affinity, as in the case of cyclic urea inhibitors of HIV-1 protease.[8] However, in some cases, water displacement seems to lower binding affinity.[18−20]

Interfacial water molecules can be isolated or can form clusters. For example, six water molecules forming two water clusters are found at the interface of the trp-repressor/operator complex, mediating the polar contacts to the base.[3,21] These water molecules are the determinants of specificity in this system.[4] Water clusters are also found at the interface between the mouse major urinary protein and its pheromone ligands.[22] Disruption of the hydrogen bonds between one of the bound water molecules and the protein−ligand complex causes a loss of binding enthalpy and a favorable change of entropy. Crystallographic studies of OppA-peptide complexes revealed a water cluster at the binding interface of the dipeptide complex.[9,10] Release of these bound water molecules to the bulk when the ligand changes from dipeptide to tri- and tetrapeptides is entropically favorable and was suggested to be responsible for the gain of binding affinity.[10]

The thermodynamic contribution of water at protein−ligand binding interfaces is still poorly understood. This causes major problems to current scoring functions for prediction of protein−ligand binding modes and affinities.[23] Attempts of refining the scoring functions taking bound water molecules into account improved the prediction quality in some cases.[24,25]

* Corresponding author. E-mail: tlazaridis@ccny.cuny.edu. Telephone: (212) 650-8364. Fax: (212) 650-6107.

We have previously applied a rigorous statistical thermodynamic analysis on tightly bound water molecules in HIV-1 protease−inhibitor and concanavalin A−trimannoside complexes.[26,27] In both cases, we found that the tightly bound water molecules at the binding interface had a large negative entropy that was outweighed by a favorable interaction between the water molecule and the protein or ligand, leading to a favorable free energy contribution (−15.2 kcal/mol and −17.2 kcal/mol, respectively). We also studied the thermodynamic consequences of the displacement of the water molecule in the concanavalin A−trimannoside complex. We found that the water−protein/ligand interactions eliminated were almost compensated by the direct interactions gained. Other contributions from the difference of the desolvation enthalpy, entropy, and conformational entropy of the ligand were much smaller but significant, compared to the binding free energy difference.[27]

In this work, we extend this analysis to water molecules that are not fully buried or form a cluster. Such a water cluster is observed in complexes of cyclophilin A (CypA) with cyclosporin A (CsA, **1**) and analogues. CsA is an immunosuppressive drug that can prevent graft rejection after transplant surgery by forming a complex with its soluble intracellular receptor protein CypA, which then interacts with calcineurin and inhibits its phosphatase activity.[28] In the complex of CypA−CsA, several bound water molecules were found at the binding interface, mediating the interactions between CypA and the Abu residue (L-α-aminobutyric acid) of CsA.[29] Three of them (WTR5, 6, and 7) are well ordered, and one is less tightly bound (WTR133). (5-Hydroxynorvaline)-2-cyclosporin (**2**) (Figure 1, compound 2) is a derivative of **1**, which was designed to form additional direct interactions with CypA. Displacement of two of the bound water molecules in the Abu pocket was observed going from **1** to **2**.[12] X-ray crystal structures of these two

10.1021/jp056020a CCC: $33.50
Published on Web 12/24/2005

complexes show that the two ligands have identical backbone conformation and the protein binding site has similar geometry.

In vitro measurements of the binding affinities of **1** and **2** gave $\Delta\Delta G^{2-1} = +1.3$ kcal/mol, showing an 8−9-fold lower affinity for **2**.[12] Mikol et al. attributed the different binding affinity to several contributions: the loss of the conformational entropy by the constraints of the longer side chain of ligand **2** in the complex, which is counter-balanced by the gained interactions of **2** in the complex, a desolvation term, which was assumed to be less than +2 kcal/mol, including the considerations of both the different desolvation energy of **1** and **2**, and the free energy cost of displacement of WTR6. The less-ordered water molecule WTR133 was treated as a solvent molecule at the hydration shell, and its contribution to the binding affinity was neglected.

In this work, we calculate the thermodynamic contributions of the water clusters in the Cyp−**1** and Cyp−**2** complexes. The contributions of each bound water molecule to the thermodynamic properties are calculated separately and compared to those of other water molecules, showing the different roles they play at the binding interface. The change in binding affinity by the ligand modification is accounted for by considering the contributions from the water clusters, ligand−protein interactions, ligand desolvation, and ligand conformational entropy loss.

Methods

In the inhomogeneous fluid solvation theory,[30,31] the solvation energy and entropy are decomposed into the solute−solvent terms (E_{sw}, S_{sw}) and the solvent reorganization terms (ΔE_{ww} and ΔS_{ww}). The latter are due to the difference in solvent−solvent interactions and correlations in the bulk and in the complex. All of these components can be expressed as integrals over the solute−solvent correlation function $g_{sw}(\mathbf{r}, \omega)$ and solvent−solvent correlation function $g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega')$.[31] Only two-particle contributions to the entropy are considered.

$$S_{sw} = -k \frac{\rho}{\Omega} \int g_{sw}(\mathbf{r}, \omega) \ln g_{sw}(\mathbf{r}, \omega)\, d\mathbf{r}d\omega \quad (1)$$

$$E_{sw} = \frac{\rho}{\Omega} \int g_{sw}(\mathbf{r}, \omega)\, u_{sw}(\mathbf{r}, \omega)\, d\mathbf{r}d\omega \quad (2)$$

$$\Delta S_{ww} = -\frac{1}{2} k \frac{\rho^2}{\Omega^2} \int g_{sw}(\mathbf{r}, \omega) [g_{sw}(\mathbf{r}', \omega')\{g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') \ln g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') - g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') + 1\} - \{g^{o}_{ww}(R, \omega^{rel}) \ln g^{o}_{ww}(R, \omega^{rel}) - g^{o}_{ww}(R, \omega^{rel}) + 1\}] d\mathbf{r}d\mathbf{r}'\, d\omega d\omega' \quad (3)$$

$$\Delta E_{ww} = -\frac{1}{2} \frac{\rho^2}{\Omega^2} \int g_{sw}(\mathbf{r}, \omega)[g_{sw}(\mathbf{r}', \omega')\, g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') - g^{o}_{ww}(R, \omega^{rel})]\, u_{ww}(R, \omega^{rel})\, d\mathbf{r}d\mathbf{r}'\, d\omega d\omega' \quad (4)$$

where k is Boltzmann's constant, ρ is the density of bulk solvent, $\mathbf{r}$ and $\mathbf{r}'$ denote the position of two water molecules, ω and ω' denote the orientation of these two water molecules with respect to the solute, Ω is the integral over ω ($\Omega = 8\pi^2$), R is the distance between two water molecules, ω^{rel} are the five angles that describe the relative orientation of two water molecules, and $\Omega^{rel} = \int d\omega^{rel} = 32\pi^3$; $g^{o}_{ww}(R, \omega^{rel})$ and $g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega')$ are the solvent−solvent correlation function in the pure solvent and in the complex, respectively; $u_{sw}(\mathbf{r}, \omega)$ and $u_{ww}(R, \omega^{rel})$ are water−solute and water−water potentials, respectively. The energy terms (E_{sw} and ΔE_{ww}) can also be more easily evaluated directly from a simulation.

The inhomogeneous fluid solvent theory has been applied to several systems: the excess entropy in pure liquid water,[32] solvation thermodynamics in simple Lennard−Jones and hard-sphere fluids,[30] and the solvent reorganization energy and entropy of hydration of methane.[31] In previous work, we applied this approach to isolated water molecules in biomolecular complexes.[26,27] In that case, the solvent−solvent energy and entropy are negligible, and ΔE_{ww} and ΔS_{ww} are equal to the energy and entropy of removing a water molecule from bulk solvent. Here, we apply the same treatment to water molecules that are not fully buried or form a cluster. Now, the solvent−solvent terms in the complex need to be evaluated, and the definition of a water molecule contribution to the energy and entropy is somewhat more complicated.

The integrals in eqs 1 and 3 are over all space V. $g_{sw}(\mathbf{r}, \omega)$ is zero over the regions occupied by the solute; the only contributions come from regions occupied by the solvent. Because any integral over V can be split into a sum of integrals over distinct subregions ($V = \sum_i \nu_i$, $V' = \sum_j \nu_j$), the contributions of specific water molecules can be determined. As a result, eqs 1 and 3 can be written as:

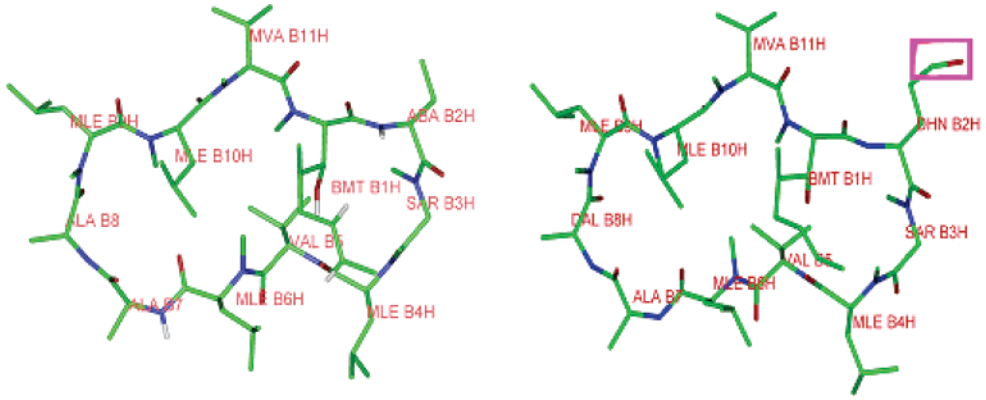
$$S_{sw} = -k \frac{\rho}{\Omega} * \int_V g_{sw}(\mathbf{r}, \omega) \ln g_{sw}(\mathbf{r},\omega)\, d\mathbf{r}d\omega = -k \frac{\rho}{\Omega} \int_{\sum_i \nu_i} g_{sw}(\mathbf{r}, \omega) \ln g_{sw}(\mathbf{r}, \omega)\, d\mathbf{r}d\omega = -k \frac{\rho}{\Omega} \sum_i \int_{\nu_i} g_{sw}(\mathbf{r}, \omega) \ln g_{sw}(\mathbf{r}, \omega)\, d\mathbf{r}d\omega \equiv \sum_i S_{sw(i)} \quad (5)$$

$$\Delta S_{ww} = -\frac{1}{2} k \frac{\rho^2}{\Omega^2} \int_V d\mathbf{r} \int_{V'} d\mathbf{r}'\, g_{sw}(\mathbf{r}, \omega) [g_{sw}(\mathbf{r}', \omega')\{g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') \ln g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') - g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') + 1\} - \{g^{o}_{ww}(R, \omega^{rel}) \ln g^{o}_{ww}(R, \omega^{rel}) - g^{o}_{ww}(R, \omega^{rel}) + 1\}] d\omega d\omega' = -\frac{1}{2} k \frac{\rho^2}{\Omega^2} \sum_i \int_{\nu_i} d\mathbf{r} \int_{\sum_j \nu_j} d\mathbf{r}' g_{sw}(\mathbf{r}, \omega) [g_{sw}(\mathbf{r}', \omega')\{g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') \ln g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') - g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') + 1\} - \{g^{o}_{ww}(R, \omega^{rel}) \ln g^{o}_{ww}(R, \omega^{rel}) - g^{o}_{ww}(R, \omega^{rel}) + 1\}] d\omega d\omega' = -\frac{1}{2} k \frac{\rho^2}{\Omega^2} \sum_i \sum_j \int_{\nu_i} d\mathbf{r} \int_{\nu_j} d\mathbf{r}'\, g_{sw}(\mathbf{r}, \omega) [g_{sw}(\mathbf{r}', \omega')\{g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') \ln g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') - g^{inh}_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega') + 1\} - \{g^{o}_{ww}(R, \omega^{rel}) \ln g^{o}_{ww}(R, \omega^{rel}) - g^{o}_{ww}(R, \omega^{rel}) + 1\}] d\omega d\omega' \equiv \sum_i \sum_j \Delta S_{w(i)w(j)} \quad (6)$$

With the identities

$$g_{sw}(\mathbf{r}, \omega) = g^{tr}_{sw}(\mathbf{r})\, g^{or}_{sw}(\omega|\mathbf{r}) \quad (7)$$

$$g_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega) = g^{r}_{ww}(\mathbf{r}, \mathbf{r}')\, g^{or}_{ww}(\omega, \omega'|\mathbf{r}, \mathbf{r}') \quad (8)$$



1. Cyclosporin A

2. (5-hydroxynorvaline)-2-cyclosporin

Figure 1. Structure of cyclic undecapeptides cyclosporin A (CsA) and (5-hydroxynorvaline)-2-cyclosporin. The only difference between these two molecules is highlighted by the purple rectangle. CsA includes: MeBmt1 (*N*-methyl-(4R)-4-[(*E*)-2-butenyl]-4-methyl-l-threonine), Abu2 (l-α-amino-butyric acid), Sar3 (sarcosine), MeLeu4 (*N*-methyl-leucine), Val5, MeLeu6, Ala7, D-Ala8, MeLeu9, MeLeu10, and MeVal11 (*N*-methylvaline).

the integrals in each subregion (ν_i, ν_j) can be decomposed into a translational and an orientational contribution:

$$S_{\mathrm{sw}(i)} = -k\rho \int_{\nu_i} g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r}) \ln g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r})\, \mathrm{d}\mathbf{r} - k\frac{\rho}{\Omega} \int_{\nu_i} g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r})\, \mathrm{d}\mathbf{r} \int_{\nu_i} g^{\mathrm{or}}_{\mathrm{sw}}(\omega|\mathbf{r}) \ln g^{\mathrm{or}}_{\mathrm{sw}}(\omega|\mathbf{r})\, \mathrm{d}\omega \quad (9)$$

and

$$\Delta S_{\mathrm{w}(i)\mathrm{w}(j)} = -\frac{1}{2}k\rho^2 \int_{\nu_i} \mathrm{d}\mathbf{r}\, g^{\mathrm{r}}_{\mathrm{sw}}(\mathrm{r}) \int_{\nu_j} \mathrm{d}\mathbf{r}' [g^{\mathrm{r}}_{\mathrm{sw}}(\mathbf{r}')\{g^{r,inh}_{ww}(\mathbf{r},\mathbf{r}') \ln g^{\mathrm{r,inh}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') - g^{\mathrm{r,inh}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') + 1\} - \{g^{\mathrm{r,o}}_{\mathrm{ww}}(R) \ln g^{\mathrm{r,o}}_{\mathrm{ww}}(R) - g^{\mathrm{r,o}}_{\mathrm{ww}}(R) + 1\}] - \frac{1}{2}k\rho^2 \int_{\nu_i} \mathrm{d}\mathbf{r}\, g^{\mathrm{r}}_{\mathrm{sw}}(\mathbf{r}) \int_{\nu_j} \mathrm{d}\mathbf{r}' [g^{\mathrm{r}}_{\mathrm{sw}}(\mathbf{r}') g^{\mathrm{r,inh}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') S^{\mathrm{m}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') - g^{\mathrm{r,o}}_{\mathrm{ww}}(R) S^{\mathrm{o}}(R)] \quad (10)$$

with $S^{m}_{ww}(\mathbf{r}, \mathbf{r}')$ and S^{o} defined as

$$S^{\mathrm{m}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') = \frac{1}{\Omega^2} \int g^{\mathrm{or}}_{\mathrm{sw}}(\omega|\mathbf{r}) g^{\mathrm{or}}_{\mathrm{sw}}(\omega'|\mathbf{r}') g^{\mathrm{or,inh}}_{\mathrm{ww}}(\omega, \omega'|\mathbf{r}, \mathbf{r}') \ln g^{\mathrm{or,inh}}_{\mathrm{ww}}(\omega, \omega'|\mathbf{r}, \mathbf{r}')\, \mathrm{d}\omega \mathrm{d}\omega' \quad (11)$$

$$S^{\mathrm{o}}(R) = \frac{1}{\Omega^{\mathrm{rel}}} \int g^{\mathrm{r,o}}_{\mathrm{ww}}(\omega^{\mathrm{rel}}|R) \ln g^{\mathrm{r,o}}_{\mathrm{ww}}(\omega^{\mathrm{rel}}|R)\, \mathrm{d}\omega^{\mathrm{rel}} \quad (12)$$

If we further assume that $g^{\mathrm{or}}_{\mathrm{sw}}(\omega|\mathbf{r})$, $g^{\mathrm{or}}_{\mathrm{sw}}(\omega'|\mathbf{r}')$ are independent of $\mathbf{r}$ or $\mathbf{r}'$ within the small regions ν_i, ν_j, i.e., $g^{\mathrm{or}}_{\mathrm{sw}}(\omega|\mathbf{r}) \approx g^{\mathrm{or}}_{\mathrm{sw}}(\omega)$, $g^{\mathrm{or}}_{\mathrm{sw}}(\omega'|\mathbf{r}') \approx g^{\mathrm{or}}_{\mathrm{sw}}(\omega')$, eqs 9 and 11 become,

$$S_{\mathrm{sw}(i)} = -k\rho \int_{\nu_i} g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r}) \ln g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r})\, \mathrm{d}\mathbf{r} - k\frac{N^{\nu_i}_{\mathrm{wat}(i)}}{\Omega} \int_{\nu_i} g^{\mathrm{or}}_{\mathrm{sw}}(\omega) \ln g^{\mathrm{or}}_{\mathrm{sw}}(\omega)\, \mathrm{d}\omega \quad (13)$$

$$S^{\mathrm{m}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') = \frac{1}{\Omega^2} \int g^{\mathrm{or}}_{\mathrm{sw}}(\omega)\, g^{\mathrm{or}}_{\mathrm{sw}}(\omega')\, g^{\mathrm{or,inh}}_{\mathrm{ww}}(\omega, \omega'|\mathbf{r}, \mathbf{r}') \ln g^{\mathrm{or,inh}}_{\mathrm{ww}}(\omega, \omega'|\mathbf{r}, \mathbf{r}')\, \mathrm{d}\omega \mathrm{d}\omega' \quad (14)$$

where $N^{\nu_i}_{\mathrm{wat}(i)}$ is the number of water molecules in the region of space ν_i, ($N^{\nu_i}_{\mathrm{wat}(i)} = \rho \int_{\nu_i} g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r})\, \mathrm{d}\mathbf{r}$).

The translational correlation function and orientational correlation function were calculated as products of one-dimensional functions, e.g., $g^{\mathrm{tr}}_{\mathrm{sw}}(\mathbf{r}) = g^{\mathrm{tr}}_{\mathrm{sw}}(r)\, g^{\mathrm{tr}}_{\mathrm{sw}}(\theta')\, g^{\mathrm{tr}}_{\mathrm{sw}}(\phi')$ and $g^{\mathrm{or}}_{\mathrm{sw}}(\theta, \phi, \psi) = g^{\mathrm{or}}_{\mathrm{sw}}(\theta)\, g^{\mathrm{or}}_{\mathrm{sw}}(\phi)\, g^{\mathrm{or}}_{\mathrm{sw}}(\psi)$, where r, θ', and ϕ' are spherical coordinates of the water oxygen with respect to its average position; θ, ϕ, and ψ are Euler angles. In this work, the bin sizes for the numerical evaluation of integrals are set to dr = 0.1 Å, $\mathrm{d}\theta' = \pi/10$, $\mathrm{d}\phi' = \pi/10$, and $\mathrm{d}\theta = \pi/10$, $\mathrm{d}\phi = 2\pi/10$, $\mathrm{d}\psi = 2\pi/10$. The translational solvent pair correlation function $g^{\mathrm{r,inh}}_{\mathrm{ww}}(\mathbf{r}, \mathbf{r}')$ and the orientational solvent pair correlation function $g^{\mathrm{or,inh}}_{\mathrm{ww}}(\omega, \omega'|\mathbf{r}, \mathbf{r}')$ are six-dimensional functions, which are very difficult to obtain directly from simulation. The Kirkwood superposition approximation (KSA) assumes that the inhomogeneous pair correlation function is equal to the bulk solvent pair correlation function and depends only on the distance and the relative orientation between two solvent molecules.

$$g^{\mathrm{inh}}_{\mathrm{ww}}(\omega, \omega', \mathbf{r}, \mathbf{r}') = g^{\mathrm{o}}_{\mathrm{ww}}(\omega^{\mathrm{rel}}, R) \quad (15)$$

It has been successfully employed to calculate the solvent reorganization enthalpy and entropy for methane in water.[31]

The KSA can be applied separately to the translational and orientational parts

$$g^{\mathrm{r,inh}}_{\mathrm{ww}}(\mathbf{r}, \mathbf{r}') = g^{\mathrm{r,o}}_{\mathrm{ww}}(R) \quad (16)$$

$$g^{\mathrm{or,inh}}_{\mathrm{ww}}(\omega, \omega'|\mathbf{r}, \mathbf{r}') = g^{\mathrm{or,o}}_{\mathrm{ww}}(\omega^{\mathrm{rel}}|R) \quad (17)$$

Thus, eq 14 becomes

$$S^{\mathrm{m}}_{\mathrm{ww}}(\mathbf{r},\mathbf{r}') = \frac{1}{\Omega^2} \int g^{\mathrm{or}}_{\mathrm{sw}}(\omega)\, g^{\mathrm{or}}_{\mathrm{sw}}(\omega')\, g^{\mathrm{or,o}}_{\mathrm{ww}}(\omega^{\mathrm{rel}}|R) \ln g^{\mathrm{or,o}}_{\mathrm{ww}}(\omega^{\mathrm{rel}}|R)\, \mathrm{d}\omega \mathrm{d}\omega' \quad (18)$$

In our calculations, R is set to the distance of the average positions of the two water molecules. For each ω, ω', and R, we calculate the five relative angles of the two water molecules. As a result, $S^{\mathrm{m}}_{\mathrm{ww}}(\mathbf{r}, \mathbf{r}')$ has a unique value for each water pair. To simplify the calculation of the five-dimensional orientational

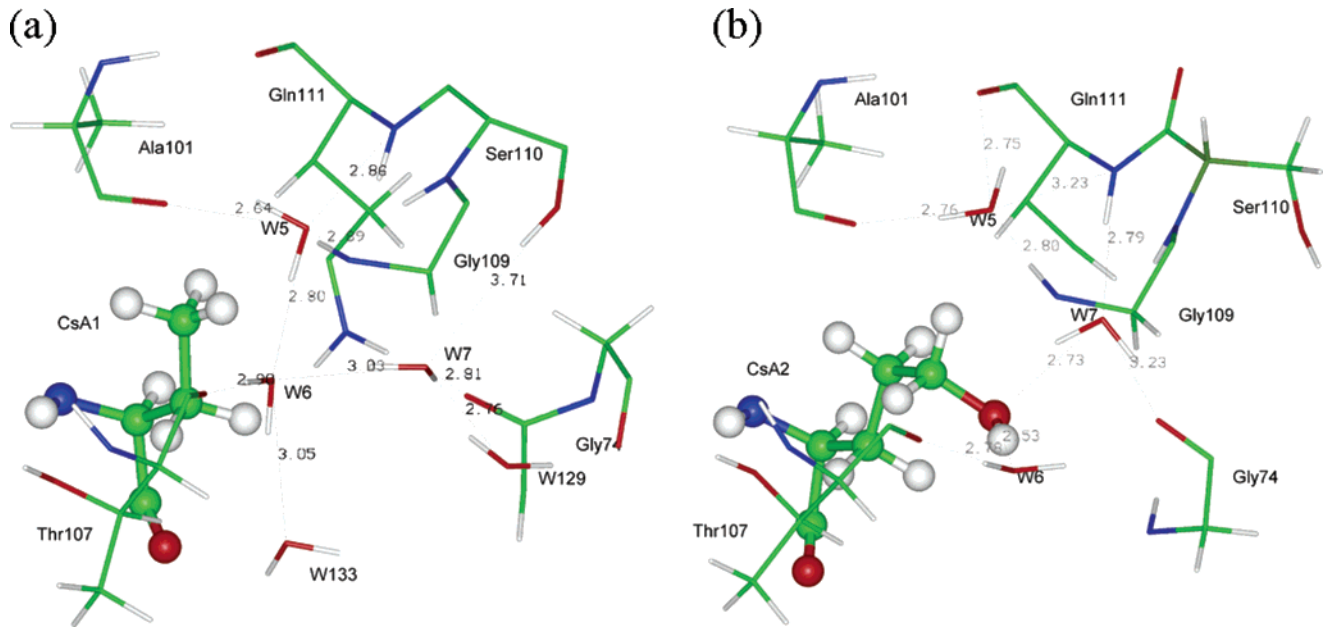


Figure 2. (a) Bound water molecules in the Abu pocket in the CypA−**1** complex. **1** is drawn in balls-and-sticks (only the residue Abu2 is depicted). CypA is drawn in thinner sticks. All potential H bonds are shown in dashed black lines. Oxygen atoms are in red, nitrogen atoms in blue, carbon atoms in green, and hydrogen atoms in gray. (b) Bound water molecules in the Abu pocket in the CypA−**2** complex. **2** is drawn in balls-and-sticks (only the residue Abu2 is depicted). CypA is drawn in thinner sticks. All potential H bonds are shown in dashed black lines. Oxygen atoms are in red, nitrogen atoms in blue, carbon atoms in green, and hydrogen atoms in gray.

function in bulk water $g_{ww}^{or,o}(\omega^{rel}|R)$, a factorization is employed.

$$g_{ww}^{or,o}(\omega^{rel}|R) = g(\theta_1, \theta_2|R)\, g(\theta_1, \chi_2|R)\, g(\theta_2, \chi_1|R)\, g(\chi_1, \chi_2|R)\, g(\phi|R)/[g(\theta_1|R)\, g(\theta_2|R)\, g(\chi_1|R)\, g(\chi_2|R)] \quad (19)$$

This factorization was found to underestimate the magnitude of the entropy in pure water (about 5%).[32] The error this brings to the final result for the solvent−solvent entropy in the complex S_{ww} for any bound water molecule is no more than 0.3 cal/mol K in this study.

Equation 6 is essentially the difference in entropy between a bound water and a water in the bulk. In this work, it has been more convenient to calculate the bound water terms and then subtract the entropy of bulk water. The entropy of bulk water calculated theoretically with similar formulas is −15.2 cal/mol K.[32] The solvent reorganization entropy of one bound water molecule is

$$\Delta S_{w(i)w} = \sum_j S_{w(i)w(j)}^{trans} + \sum_j S_{w(i)w(j)}^{or} + 15.2 \text{ cal/mol K} \quad (20)$$

where the solvent−solvent translational part $S_{w(i)w(j)}^{trans}$ and the solvent−solvent orientational part $S_{w(i)w(j)}^{or}$ of the entropy of w(i) can be calculated by excluding the terms of bulk water in eq 10, i.e.,

$$S_{w(i)w(j)}^{trans} = -\frac{1}{2} k\rho^2 \int_{\nu_i\nu_j} g_{sw(i)}^{r}(\mathbf{r})\, g_{sw(j)}^{r}(\mathbf{r}') \{g_{w(i)w(j)}^{r,inh}(\mathbf{r}, \mathbf{r}') \ln g_{w(i)w(j)}^{r,inh}(\mathbf{r}, \mathbf{r}') - g_{w(i)w(j)}^{r,inh}(\mathbf{r}, \mathbf{r}') + 1\}\, d\mathbf{r}d\mathbf{r}' \quad (21)$$

$$S_{w(i)w(j)}^{or} = -\frac{1}{2} k\rho^2 \int_{\nu_i\nu_j} g_{sw(i)}^{r}(\mathbf{r})\, g_{sw(j)}^{r}(\mathbf{r}')\, g_{w(i)w(j)}^{r,inh}(\mathbf{r}, \mathbf{r}')\, S_{w(i)w(j)}^{m}(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}\, d\mathbf{r}' \quad (22)$$

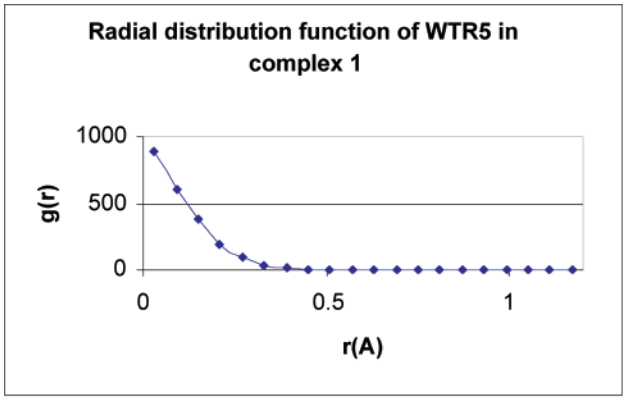
The solvation entropy ΔS_{solv} is the sum of the solute−solvent entropy and solvent reorganization entropy of all bound water molecules.

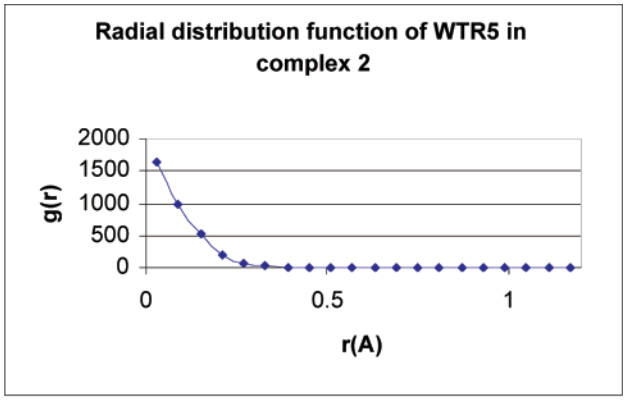
$$\Delta S_{solv} = \sum_i (\Delta S_{w(i)w} + S_{sw(i)}) \quad (23)$$

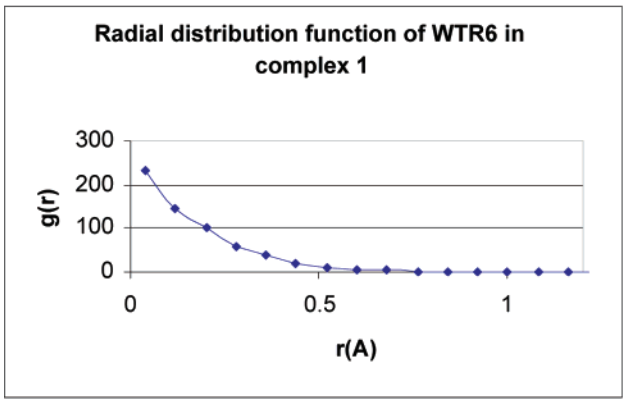
Thus, the contribution of specific water molecules (or regions of space) to the solvation entropy can be determined. The solvation energy is evaluated directly from the simulations.

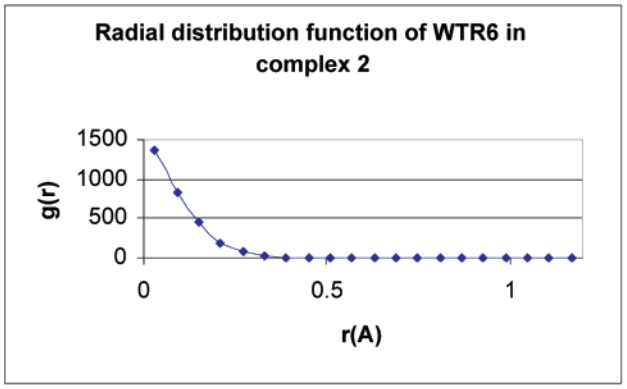
We first split the region occupied by a cluster of bound water molecules into distinct spherical regions (i) of radius 1.2 Å based on the average positions of each bound water molecule obtained from the MD simulation. This cutoff value was selected because, in bulk water, the nearest neighbor's distance is about 2.8 Å for the oxygen−oxygen pair, and half of this value is 1.4 Å. Also, as shown in Figure 3, the density of each bound water molecule decreases to 0 before r reaches 1.2 Å. Next, we calculate the occupancy $O(i)$ of each region. We identify any water molecule located in a region (i) with the corresponding bound water molecule [w(i)], i.e., we allow for exchange between water molecules. The solvent−solvent energy and entropy terms ($\Delta E_{w(i)w(j)}$ and $\Delta S_{w(i)w(j)}$) were calculated separately for each water pair. The water−water pair is denoted by w(i)w(j) and is composed of the bound water molecule in region i and any other water molecule close to it and lying in different subregions (j). The subregions (j) were defined by scanning a spherical space of a bound water molecule [w(i)] within a radius of 5 Å (the magnitude of the interaction energy E_{ww} is no more than 0.2 kcal/mol beyond this distance) and looking for locations of high water density.

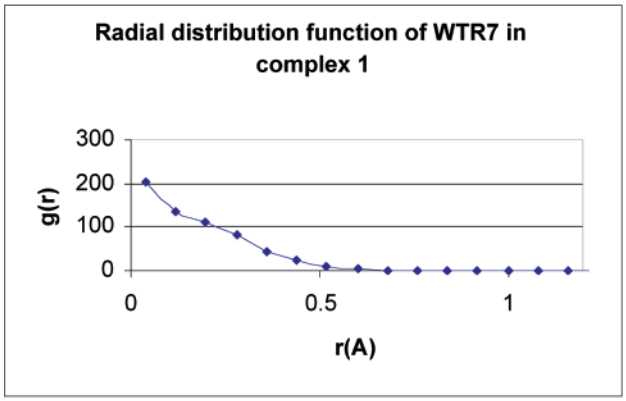
Calculation of the Desolvation Terms. The solvation contributions of polar and apolar groups of the ligands were assumed to be additive and proportional to their solvent accessible surface area (SASA).[33] The analytic surface area method in CHARMM was employed to calculate the polar (oxygen and nitrogen) and apolar (carbon) SASA of each ligand in the bound and free states by using a 1.0 Å water probe radius.

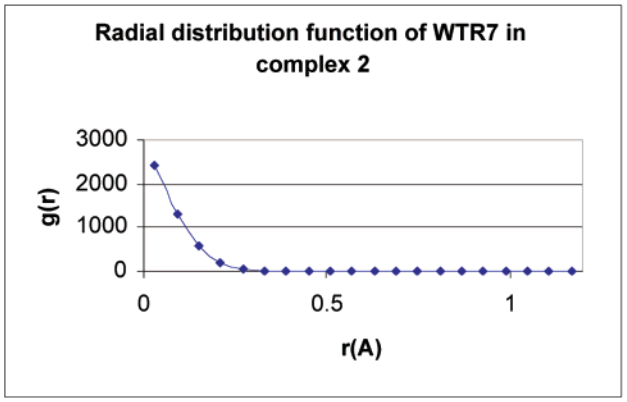












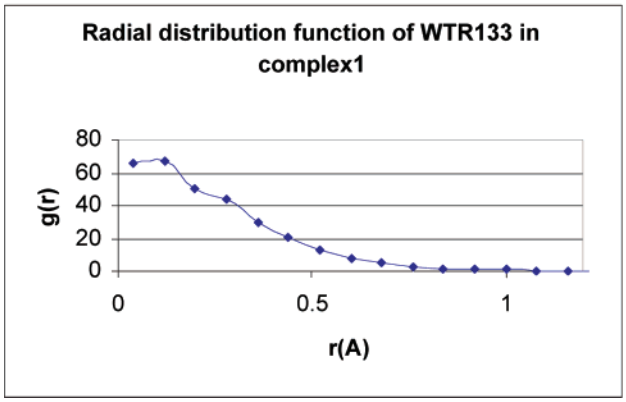


Figure 3. Radial distribution function of different water molecules with respect to their average position in complex **1** and **2** at 300 K.

With this solvent probe radius, the SASA of serine matched the value given by Miller et al.,[34] which was used to derive the dehydration enthalpies, entropies, and free energies per unit of surface area.[33,35] The values for the aliphatic and polar parts of the serine side chain (0.03 and 0.25 kcal/mol Å^2, respectively, for the enthalpy, 0.14 and 0.24 cal/mol K Å^2, respectively, for the entropy, and −0.01 and 0.18 kcal/mol Å^2, respectively, for the free energy[33,35]) were used to calculate the difference in desolvation enthalpy and entropy of the two ligands.

MD Simulations. The initial structures were constructed on the basis of the X-ray structures of the protein CypA complexed with ligand **1** or **2** and all crystallographic water molecules (pdb codes 1CWA and 1MIK, respectively). The major conformation of ligand **2** reported in the crystal structure was used here. The

CHARMM22 force field[36] was used for the protein and ligand **1** and **2**. Partial charges for the ligands were obtained with QUANTA using CHARMM template charges. Molecular dynamics (MD) simulations were carried out with the program CHARMM, version c30a1. A 15 Å sphere of TIP3P water molecules was added around the active site and subjected to spherical stochastic boundary conditions.[37] Solvent molecules overlapping with the protein, ligands, or bound water molecules were deleted. This procedure did not result in the insertion of any additional water molecules at the binding interface. The SHAKE procedure was used to constrain the bonds involving hydrogen. The integration time step of the MD simulations was 2 fs. A cutoff distance of 30 Å was applied for the computation of nonbonded interactions. First, potential energy minimizations were performed for 2000 steps by using the ABNR method. Then, MD simulations at 300 K were performed, starting from the energy-minimized system and lasting for 8 ns with the protein CypA and the ligands kept fixed. The protein and ligands were kept fixed in order to reduce the statistical uncertainties in the computed contributions to the binding affinity and to avoid possible structural shifts due to force field inaccuracies.

Results

The crystal structure of CypA−**1** complex shows four water molecules found in the cavity called "Abu pocket" interacting with each other with hydrogen bonds, i.e., WTR5 hydrogen bonds with Ala101, Gln111, Gly109, and WTR6; WTR6 hydrogen bonds with Thr107 WTR 5, 7, and 133; WTR7 hydrogen bonds with WTR6, WTR129, Gln111 (or SER110), and Gly74; WTR133 hydrogen bonds with WTR6 and other solvent water molecules around it. The crystal structure of the CypA−**2** complex shows that WTR6 moves slightly out of the Abu pocket and presumably induces the displacement of WTR133; the location of the other two bound water molecules (WTR5 and 7) did not change much; no direct hydrogen bond was observed among these water molecules.[12]

In the MD simulations performed on complex **1** and **2**, WTR6 appears less frequently around its location in complex **1** than in complex **2**, where it is observed hydrogen bonding with Thr107, the hydroxyl group of ligand **2**, and some other solvent molecules, while WTR5 and 7 are more stable in both complexes. All these bound water molecules are found to hydrogen bond to the protein and ligand as in the crystal structures[12,29] (see Figure 2). The number of hydrogen bonds formed between each bound water molecule and the protein and ligand observed from the last snapshot of the MD simulation are listed in Tables 1 and 2. In the MD simulation, no hydrogen bond is formed between any of the four bound water molecules and ligand **1** in complex **1** due to the hydrophobic nature of the Abu side chain in ligand **1**, while in complex **2**, either WTR 6 or 7 is found to make one hydrogen bond with ligand **2**.

The region occupied by each bound water molecule is defined as a spherical region of radius 1.2 Å centered at its average position obtained from MD simulation. In this way, we split the region occupied by a water cluster into separate spherical regions and studied the thermodynamic properties of water molecules in each region separately. From the MD simulations performed on CypA−**1**, the calculated occupancy (O) of the four bound water molecules in the Abu pocket decreases as: $O_{WTR5} > O_{WTR7} > O_{WTR6} > O_{WTR133}$ (Table 1). This order correlates with the number of hydrogen bonds formed between these water molecules and the protein mentioned above.

We first calculated the solute−solvent entropy of each water molecule in their corresponding regions. The translational correlation function with respect to the average position ($g_{sw}^{tr}(\mathbf{r})$) and orientational correlation function ($g_{sw}^{or}(\omega)$) were calculated as products of one-dimensional functions (see Figure 4). The results calculated with and without this approximation are similar. The radial distribution function for each water (Figure 3) shows that WTR133 (with the lowest peak) is the least ordered with respect to the solute among four bound water molecules in complex **1**. We also find that WTR7 is much less ordered in complex **1** than in complex **2**, presumably due to the additional hydrogen bond formed between WTR7 and the hydroxyl group of ligand **2** (see Figure 2b).

Integrations over these correlation functions employing eq 9 give the solute−solvent contributions of each bound water molecule to the entropy $S_{sw(i)} = S_{sw(i)}^{trans} + S_{sw(i)}^{or}$ (Tables 1 and 2). In complex **1**, WTR5 exhibits the most negative solute−solvent entropy S_{sw} and solute−solvent energy E_{sw}. E_{sw} and S_{sw} correlate with the occupancy and the number of hydrogen bonds, but the correlation for S_{sw} is not perfect. In complex **2**, all bound water molecules considered have a more negative value of S_{sw} than WTR5 in complex **1**, indicating that the involvement of the

TABLE 1: Contributions to Solvation from the Ordered Water Molecules in CypA−1, Calculated from MD Simulations at 300 K[a]

bound water	WTR5	WTR6	WTR7	WTR133
number of H bonds with the protein[b]	3	1	2	0
occupancy[c]	7999/8000	7747/8000	7983/8000	7481/8000
E_{sw}	−25.6 ± 0.1	−8.3 ± 0.3	−14.1 ± 0.1	−3.7 ± 0.1
S_{sw}^{or}	−8.4 ± 0.2	−6.9 ± 0.4	−5.1 ± 0.01	−5.2 ± 0.2
S_{sw}^{tr}	−9.9 ± 0.2	−6.5 ± 0.3	−7.1 ± 0.1	−4.9 ± 0.4
S_{sw}	−18.3 ± 0.3	−13.4 ± 0.5	−12.2 ± 0.1	−10.1 ± 0.5
E_{ww}	−3.8 ± 0.1	−15.4 ± 0.6	−5.1 ± 0.1	−15.6 ± 0.6
S_{ww}^{or}	−5.1 ± 0.1	−2.9 ± 0.3	−3.5 ± 0.2	−3.2 ± 0.4
S_{ww}^{trans}	−0.4 ± 0.01	−0.5 ± 0.03	−0.3 ± 0.01	−0.3 ± 0.1
S_{ww}	−5.5 ± 0.1	−3.4 ± 0.3	−3.8 ± 0.2	−5.6 ± 0.4
ΔE_{solv}	−17.4 ± 0.1	−5.9 ± 0.4	−6.6 ± 0.05	−1.4 ± 0.3
ΔS_{solv}	−8.7 ± 0.3	−1.6 ± 0.5	−0.7 ± 0.2	+1.5 ± 0.5
ΔG_{solv}	−14.8 ± 0.1	−5.4 ± 0.4	−6.4 ± 0.1	−1.9 ± 0.3

[a] Units for enthalpy and free energy are kcal/mol; Units for entropy are cal/mol K. [b] No H bond is found between the water molecules and the ligand. [c] Probability of finding a water molecule in the specific region, within $r = 1.2$ Å from its average position.

TABLE 2: Contributions to Solvation from the Bound Water Molecules in CypA−2, Calculated from the MD Simulations at 300 K[a]

bound water	WTR5	WTR6	WTR7
number of H bonds with the protein and ligand[b]	3	2	3
occupancy	8000/8000	8000/8000	8000/8000
E_{sw}	−23.0 ± 0.04	−17.7 ± 0.01	−24.3 ± 0.06
E_{ww}	+0.4 ± 0.01	−11.0 ± 0.08	+3.4 ± 0.1
S_{sw}^{or}	−9.6 ± 0.05	−9.9 ± 0.01	−9.1 ± 0.02
S_{sw}^{tr}	−11.3 ± 0.2	−10.6 ± 0.1	−12.2 ± 0.05
S_{sw}	−20.9 ± 0.2	−20.5 ± 0.1	−21.3 ± 0.1
S_{ww}^{or}	+0.2 ± 0.001	−0.6 ± 0.5	−0.2 ± 0.02
S_{ww}^{trans}	−0.01 ± 0.01	−0.6 ± 0.02	−0.04 ± 0.01
S_{ww}	+0.2 ± 0.01	−1.2 ± 0.5	−0.2 ± 0.02
ΔE_{solv}	−12.7 ± 0.04	−13.1 ± 0.08	−12.5 ± 0.1
ΔS_{solv}	−5.5 ± 0.2	−6.5 ± 0.5	−6.3 ± 0.1
ΔG_{solv}	−11.1 ± 0.1	−11.1 ± 0.2	−10.6 ± 0.1

[a] Units for enthalpy and free energy are kcal/mol; Units for entropy are cal/mol K. [b] For either WTR6 or 7, one of the H bonds is formed between the water and the ligand.

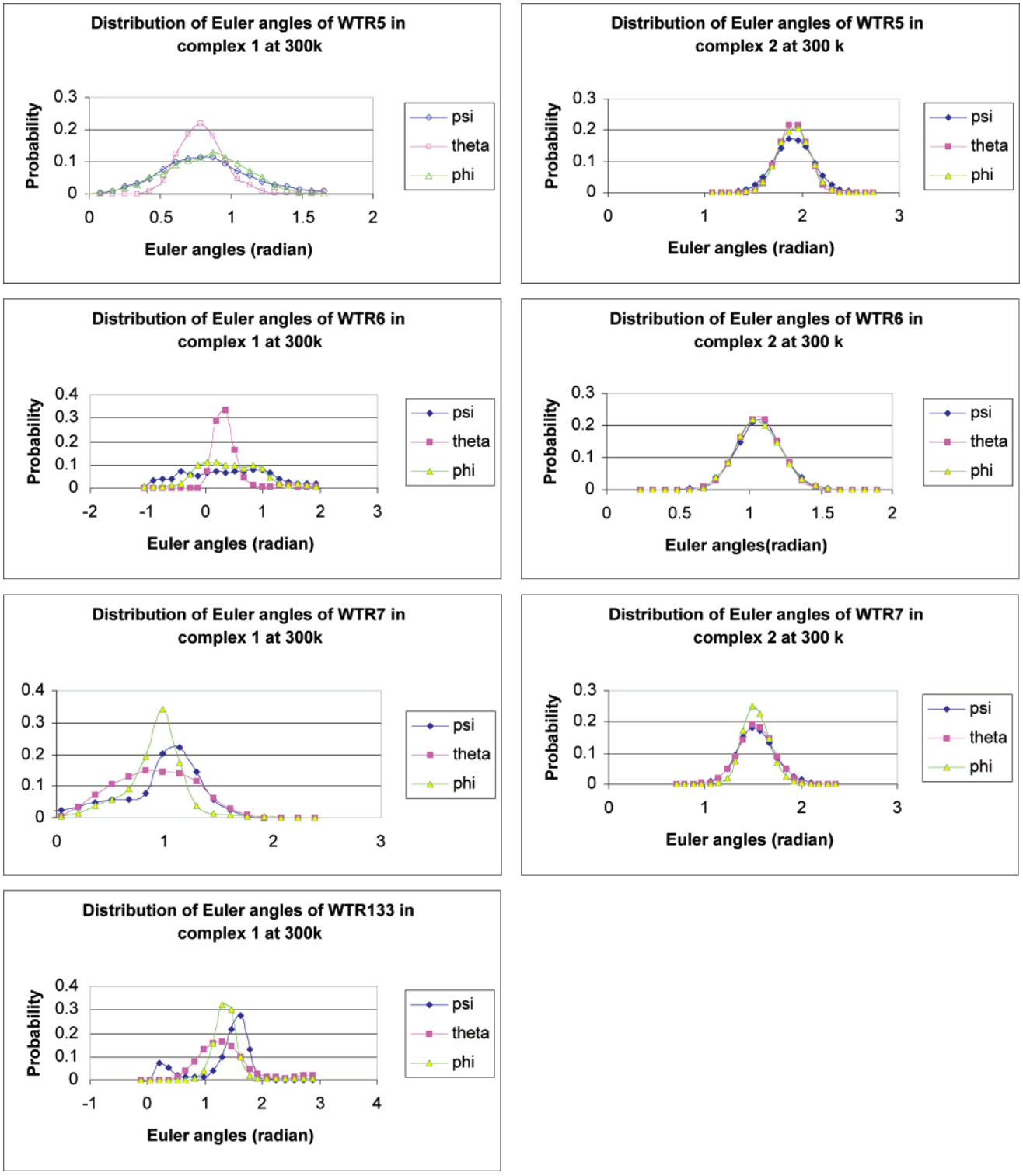


Figure 4. Probability distribution of the Euler angles of different bound water molecules in complex **1** and **2**.

additional hydroxyl group in ligand **2** stabilizes the remaining bound water molecules in the Abu pocket.

To identify water pairs for the solvent−solvent entropy calculations, we determined the distribution of solvent molecules within a 5 Å sphere around each bound water molecule. The sphere was split into several subregions (*j*), each region corresponding to a specific water molecule (Table 3). For instance, the water molecules around WTR7 within 5.0 Å in complex **1** approximately concentrate into 5 subregions whose corresponding water molecules are WTR5, WTR129, WTR6, solv54, and WTR133, respectively (Figure 5). The solvent occupancies in these subregions are 0.984, 0.909, 0.771, 0.711, and 0.302, respectively. Some of the subregions are not fully occupied, especially when the subregion is farther away from the corresponding bound water. Subregions with lower occupancy make a smaller contribution to the solvent−solvent entropy. Every bound water molecule [w(*i*)], together with a corresponding water molecule [w(*j*)] in a specific subregion were treated as a solvent pair and their solvent−solvent correlation functions were calculated.

By applying eqs 16−22, we calculated the solvent−solvent contributions of each bound water molecule to the entropy in

TABLE 3: Regions of Neighboring Water Molecules around Each Bound Water Molecule

	bound water	neighbor regions	R (Å)	θ (radian)	ϕ (radian)	corresponding water[a]
complex **1**	WTR5	1	2.0−4.0	1.1−1.85	0.0−1.6	WTR7
		2	2.0−5.0	1.85−2.35	0.0−1.3	WTR129* [b]
		3	2.0−3.0	2.5−3.14	0.0−1.2	WTR6*
	WTR6	1	2.0−3.0	2.2−2.8	0.9−2.1	WTR133*
		2	2.0−3.0	0.0−0.6	3.1−4.75	WTR5
		3	2.0−5.0	1.2−1.75	0.0−0.9	WTR129*
		4	3.0−4.0	0.4−1.0	0.0−1.1	WTR7
		5	3.0−5.0	1.3−2.2	0.9−1.6	solv54*
	WTR7	1	2.0−4.0	1.3−1.8	3.3−4.3	WTR5
		2	2.0−5.0	2.05−2.70	0−1.1	WTR129*
		3	3.0−4.0	2.2−2.70	3.3−4.3	WTR6*
		4	3.0−5.0	2.2−2.70	1.3−2.3	solv54*
		5	4.0−5.0	2.65−3.0	2.4−3.3	WTR133*
	WTR133	1	2.0−3.0	0.3−0.95	4.0−5.3	WTR6*
		2	2.0−3.0	0.8−1.5	0.8−1.5	solv54*
		3	2.0−5.0	2.05−2.6	1.6−2.5	WTR110*
		4	3.0−5.0	1.55−2.0	0.7−1.9	solv9*
		5	3.0−5.0	1.0−1.6	0.0−0.7	solv59*
		6	3.0−5.0	1.7−2.5	0.0−0.7	WTR91*
complex **2**	WTR5	1	3.0−4.0	0.8−1.2	0.5−1.2	WTR7
	WTR6	1	2.0−3.0	1.3−1.6	5.7−6.28	WTR102*
		2	2.0−5.0	0.85−1.5	1.2−1.9	WTR67*
		3	3.0−4.0	0.3−0.9	0−1.4	WTR114*
		4	3.0−5.0	0.3−0.75	4.4−5.1	WTR7
		5	4.0−5.0	1.5−1.8	1.1−1.9	solv83*
		6	3.0−5.0	1.15−1.45	0.6−0.9	solv98*
	WTR7	1	3.0−4.0	2−2.3	3.6−4.3	WTR5
		2	4.0−5.0	2.15−2.5	0−0.8	WTR102
		3	3.0−5.0	1.6−1.9	0.7−1.5	WTR114*
		4	3.0−5.0	2.4−2.7	1.2−1.9	WTR6

[a] "Corresponding water" means to the water molecule located in that region. "WTR" denotes those water molecules from the crystal structure, "solv" denotes those water molecules in the solvent sphere we added around the binding sites. [b] The water molecule in that region exchanged at least once during the MDs.

complex **1** and **2**. The pure water pair correlation function ($g_{ww}^{or,o}(\omega^{rel}|R)$) was calculated by an 8 ns MD simulation, employing eq 19. This function was used to approximate $g_{ww}^{inh,o}(\omega, \omega'|\mathbf{r}, \mathbf{r}')$ at the corresponding bin of Euler angles ω, ω' and R ($R = |\mathbf{r} - \mathbf{r}'|$). Equation 18 was used to calculate the solvent orientational entropy in the complex $S_{ww}^{m}(\mathbf{r}, \mathbf{r}')$. The solvent pair translational correlation function $g_{ww}^{r,inh}(\mathbf{r}, \mathbf{r}')$ in the complex was approximated by the radial distribution function $g_{ww}^{or,o}(R)$ in the bulk (eq 16). The solvent−solvent entropy was then calculated for each solvent pair by eqs 21 and 22. The results are listed in Table 4. The solvent−solvent entropy of each pair of water molecules varies from −3.78 to +0.15 cal/mol K and reveals how correlated these two water molecules are. The slightly positive values are probably an artifact of the KSA approximation and the numerical integration. Small values of the pair entropy mean that the two water molecules are constrained by the protein or ligand in a conformation that disrupts interactions and correlations between them. S_{ww} in bulk water was calculated as −15.2 cal/mol K[32] and originates mainly from the water−water correlations at the first neighbor shell. For a given water molecule in bulk, there are about four solvent molecules in its first neighbor shell, therefore, the value of S_{ww} for each water−water pair in the bulk is approximately −3.8 kcal/mol K, which is very close to the most negative value of the solvent−solvent entropy of a water−water pair in the complex.

By summing up the pair entropies $S_{w(i)w(j)}$, we obtained the solvent−solvent entropy for each bound water molecule $S_{w(i)w}^{trans}$ and $S_{w(i)w}^{or}$, as listed in Tables 1 and 2. For instance, for WTR7 in complex **1**, $S_{w(7)w}^{or} = \sum_j S_{w(7)w(j)}^{or} = -3.5$ cal/mol K, $S_{w(7)w}^{trans} = \sum_j S_{w(7)w(j)}^{trans} = -0.3$ cal/mol K, and $S_{w(7)w} = S_{w(7)w}^{or} + S_{w(7)w}^{trans} = -3.8$ cal/mol K. The calculated results reveal that, in complex

TABLE 4: The Solvent−Solvent Entropy (cal/mol K) of Each Bound Water Molecule[a]

	bound water [w(i)]	neighbor regions	corresponding water [w(j)]	solvent−solvent entropy $S_{w(i)w(j)}$[b]
complex **1**	WTR5	1	WTR7	−3.50
		2	WTR129*	−0.01
		3	WTR6*	−2.05
	WTR6	1	WTR133*	−1.70
		2	WTR5	−1.60
		3	WTR129*	−0.05
		4	WTR7	−0.10
		5	solv54*	+0.02
	WTR7	1	WTR5	−3.63
		2	WTR129*	−0.03
		3	WTR6*	−0.04
		4	solv54*	−0.02
		5	WTR133*	−0.03
	WTR133	1	WTR6*	−1.51
		2	solv54*	−2.03
		3	WTR110*	−0.02
		4	solv9*	−0.01
		5	solv98*	−0.03
		6	WTR91*	−0.03
complex **2**	WTR5	1	WTR7	+0.15
	WTR6	1	WTR102*	−0.37
		2	WTR67*	−0.17
		3	WTR114*	−0.41
		4	WTR7	−0.26
		5	solv59*	+0.01
		6	solv83*	+0.01
	WTR7	1	WTR5	+0.14
		2	WTR102	−0.12
		3	WTR114*	+0.02
		4	WTR6	−0.26

[a] Units for entropy are cal/mol K. [b] Good convergence of the results of $S_{w(i)w(j)}$ was obtained over two portions of MD trajectory, and the values listed here were obtained over the whole MD trajectory. $S_{w(i)w(j)}$ and $S_{w(j)w(i)}$ give similar values.

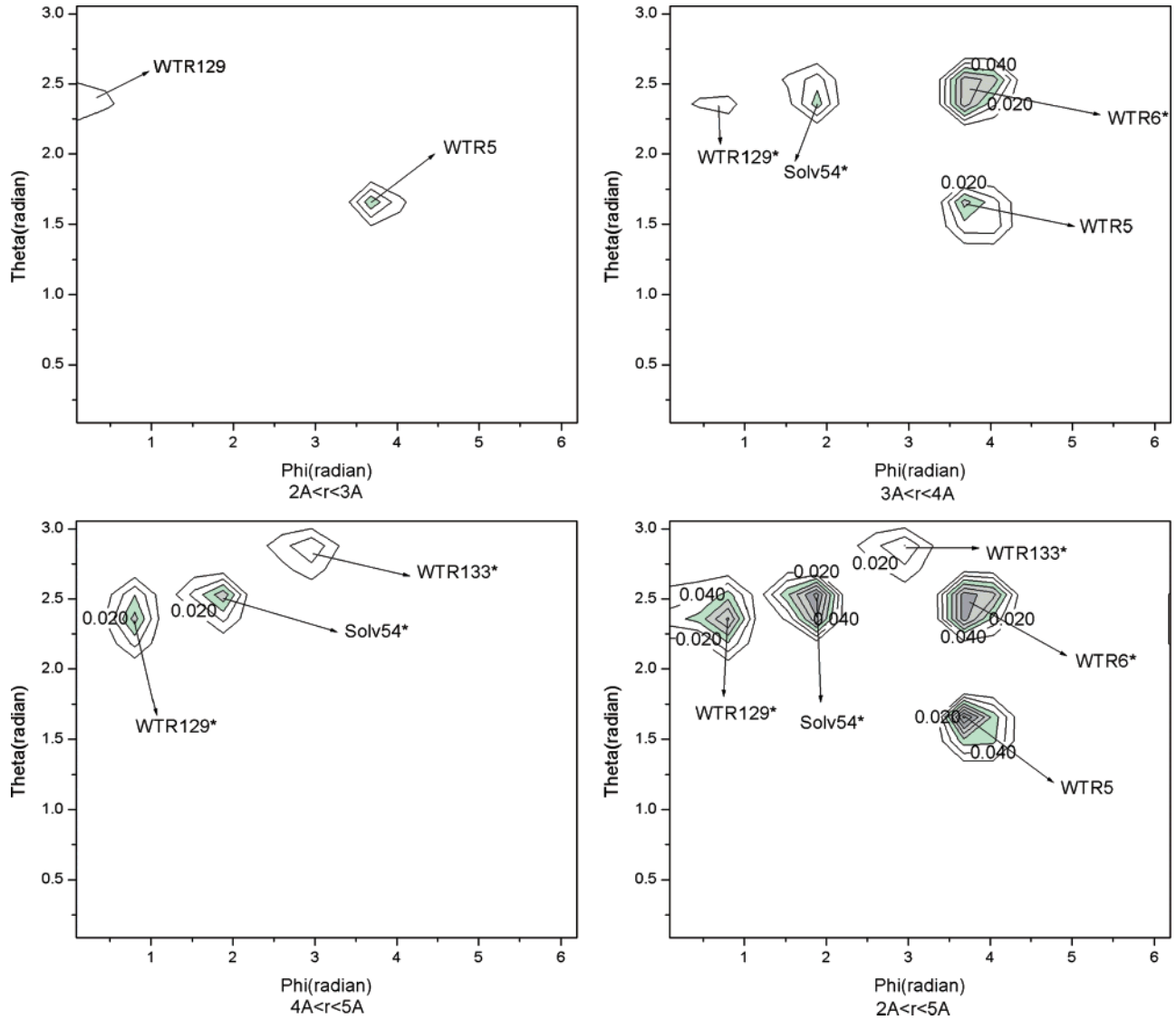


Figure 5. Distribution of probabilities of finding solvent water molecules around the bound water molecule WTR7 in complex **1** during the MDs with different cutoff values of *r*.

1, WTR5 is relatively more correlated to other solvent molecules than WTR6, 7, and 133 and all of them have a more negative value of solvent−solvent entropy than those of the bound water molecules in complex **2**, where the values of solvent−solvent entropy for WTR5 and 7 are close to 0. These results show that WTR 5 and 7 in complex **2** are more isolated than the corresponding water molecules in complex **1**.

The contribution of each bound water molecule to the solvation entropy (ΔS_{solv}, see Tables 1 and 2) indicates that transferring WTR6 and 7 from the bulk to complex **1** is accompanied by a very small entropy penalty compared to that of WTR5. The transfer of WTR133 is even entropically favorable, showing that it is even less ordered in complex **1** than in the bulk, which is related to the fact that it interacts weakly with the protein. In contrast, all three bound water molecules in complex **2** have a large negative contribution to the solvation entropy (ΔS_{solv}), showing that each of them is highly ordered in the complex. Notably, the total entropy of the three water molecules in **2** is more negative than that of the four water molecules in **1**. This is because the excess hydrogen bonds formed between the methyl hydroxyl group of ligand **2** with WTR6 and 7 constrain these water molecules, inducing a more negative entropy for each of them.

The solute−solvent energy $E_{sw(i)}$ and solvent−solvent energy $E_{w(i)w}$ of each bound water calculated directly from the simulations are listed in Tables 1 and 2. Part of the solute−solvent energy $E_{sw(i)}$ is due to long-range interactions (about −4 kcal/mol), and the rest is due to hydrogen bonds formed with the protein and ligand. For example, the H bond energy of WTR5, 6, 7, and 133 is about −20, −5, −10, and 0 kcal/mol, respectively, in complex **1**. These values reflect the number of hydrogen bonds formed between each bound water molecule and the protein−ligand complexes. The solute−solvent energy $E_{sw(i)}$ for WTR 5, 6, 7, and 133 in complex **1** and WTR5 in complex **2** mainly originates from their interaction with the protein; their interaction with the ligand **1** or **2** is relatively weak (−0.1, −0.6, −0.1, −2.4, and −0.7 kcal/mol, respectively), showing that these water molecules essentially stabilize the conformation of the protein residues at the binding site, rather than bridging protein and ligand. WTR6 and 7 in complex **2** were found to interact more strongly with the ligand (−5.4 and −5.5 kcal/mol, respectively) than in complex **1** due to their additional hydrogen bond with the hydroxyl group of ligand **2**.

The solvent−solvent energy of each bound water molecule shows strong interactions of WTR6 and WTR133 in complex **1** (−15.4 and −15.6 kcal/mol, respectively) and WTR6 (−11.0

TABLE 5: Apolar and Polar Accessible Surface Areas (Å^2) of the Ligands at Free and Bound States

	bound		free		difference (ΔASA)	
	polar	apolar	polar	apolar	polar	apolar
CsA (**1**)	95.1	597.0	178.9	1046.2	83.8	449.2
(5-hydroxynorvaline)-2-cyclosporin (**2**)	98.9	597.6	215.1	1063.5	116.2	465.9

kcal/mol) in complex **2**, which are even comparable to the solvent−solvent interactions in bulk water (−20.2 kcal/mol[32]). The solvent−solvent interactions of other bound water molecules are much weaker. The solvent−solvent energy between WTR5 and WTR7 is even positive in complex **2**, showing that these two bound water molecules are tightly fixed by the protein and ligand, leading to unfavorable water−water interaction. The contribution of a bound water molecule [w(i)] to the solvation enthalpy was calculated as

$$\Delta E_{\mathrm{solv}} = E_{\mathrm{sw}(i)} + \frac{1}{2}E_{\mathrm{w}(i)\mathrm{w}} + 10.1 \text{ kcal/mol} \qquad (24)$$

where 10.1 kcal/mol is half of the solvent−solvent interaction energy in bulk water, which is equal to the energy cost of transferring one water molecule from bulk to the gas phase.

The contribution of each bound water molecule to the solvation free energy was calculated as the sum of the entropy and energy terms (the $P\Delta V$ term is negligible):

$$\Delta G_{\mathrm{solv}(i)} = E_{\mathrm{sw}(i)} + \Delta E_{\mathrm{w}(i)\mathrm{w}} - T(S_{\mathrm{sw}(i)} + \Delta S_{\mathrm{w}(i)\mathrm{w}}) \qquad (25)$$

The contribution of WTR5 to the solvation free energy in complex **1** was obtained as −14.8 kcal/mol and originates mainly from its interactions with the protein. The contributions of WTR6, 7, and 133 to the solvation free energy is much smaller, −5.4, −6.4, and −1.9 kcal/mol, respectively. The small magnitude of ΔG_{solv} of WTR133 shows that it does not contribute to the binding between CypA and CsA as favorably as the other three bound water molecules in complex **1**. The contributions of WTR5, 6, and 7 to the solvation free energy in complex **2** are −11.1, −11.1, and −10.6 kcal/mol, respectively, and also originates mainly from their interactions with the protein. The difference of the contributions of the water clusters in complex **1** and **2** is

$$\Delta\Delta G_{\mathrm{solv}}{}^{\mathbf{2-1}} = [\Delta G_{\mathrm{solv}}(\mathrm{WTR5}) + \Delta G_{\mathrm{solv}}(\mathrm{WTR6}) + \Delta G_{\mathrm{solv}}(\mathrm{WTR7})]_{\mathbf{2}} - [\Delta G_{\mathrm{solv}}(\mathrm{WTR5}) + \Delta G_{\mathrm{solv}}(\mathrm{WTR6}) + \Delta G_{\mathrm{solv}}(\mathrm{WTR7}) + \Delta G_{\mathrm{solv}}(\mathrm{WTR133})]_{\mathbf{1}} = -4.3 \text{ kcal/mol} \qquad (26)$$

i.e., there is an *improvement* in the free energy of solvation going from **1** to **2** despite the reduction in the number of bound water molecules from 4 to 3.

In addition to solvation free energy, the different affinity of **2** and **1** is also due to the change in protein−ligand interactions, the different desolvation energy and entropy, and the different conformational restriction entropy of the two ligands.

On the basis of the energy-minimized crystal structures, the direct interactions of the ethyl group of **1** and the hydroxypropyl group of **2** with the protein were −4.1 and −9.1 kcal/mol, respectively, giving $\Delta E_{\mathrm{Abu\_side\_chain-protein}}{}^{\mathbf{2-1}} = -5.0$ kcal/mol. However, the total difference of the direct interactions of ligand **2** and **1** with the protein is $\Delta E_{\mathrm{ligand-protein}}{}^{\mathbf{2-1}} = -1.4$ kcal/mol. The extra CH_2OH of ligand **2** probably pushes it about 0.2 Å farther away from the protein than ligand **1**, and this reduces the interaction of the rest of the ligand with the protein.

TABLE 6: Enthalpic and Entropic Contributions to the Difference in the Binding Affinity between the Two Ligands[a]

	CypA−**1**	CypA− **2**	difference between **1** and **2**
$E_{\mathrm{ligand-protein}}$	−113.6	−115.0	−1.4
ΔE_{solv} (water cluster)	−31.3 ± 0.5	−38.3 ± 0.1	−7.0 ± 0.5
$-T\Delta S_{\mathrm{solv}}$ (water cluster)	+2.8 ± 0.2	+5.5 ± 0.1	+2.7 ± 0.2
$\Delta H_{\mathrm{desolv}}$[b]	+34.4	+43.0	+8.6
$-T\Delta S_{\mathrm{desolv}}$[b]	−25.0	−27.9	−2.9
$-T\Delta S_{\mathrm{config}}$	+0.9	+1.7	+0.8
ΔH (total)[c]	−110.5 ± 0.5	−110.3 ± 0.1	+0.2 ± 0.5
$-T\Delta S$ (total)	−21.3 ± 0.2	−20.7 ± 0.1	+0.6 ± 0.2
ΔG (total)	−131.8 ± 0.5	−131.0 ± 0.1	+0.8 ± 0.5

[a] Units for enthalpy, $T\Delta S$, and free energy are kcal/mol. [b] $\Delta H_{\mathrm{desolv}}$ and $\Delta S_{\mathrm{desolv}}$ are the differences in dehydration enthalpy and entropy. [c] ΔH (total) = $E_{\mathrm{ligand-protein}}$ + ΔE_{solv} (water cluster) + $\Delta H_{\mathrm{desolv}}$, ΔS (total) = $\Delta S_{\mathrm{ligand-protein}}$ + ΔS_{solv} (water cluster) + $\Delta S_{\mathrm{config}}$, and ΔG (total) = ΔE (total) − $T\Delta S$ (total).

Another contribution to the difference in binding affinity is the different desolvation enthalpy and entropy of the two ligands. The polar (oxygen and nitrogen) and apolar (carbon) solvent accessible surface areas in the bound and free states are shown in Table 5. With the dehydration enthalpy and entropy parameters given by Makhatadze and Privalov, the desolvation entropy and enthalpy of the two ligands at 300 K were obtained as $\Delta\Delta H_{\mathrm{desolv}}{}^{\mathbf{2-1}} = +8.6$ kcal/mol and $\Delta\Delta S_{\mathrm{desolv}}{}^{\mathbf{2-1}} = +10.1$cal/mol K, giving $\Delta\Delta G_{\mathrm{desolv}}{}^{\mathbf{2-1}} = +5.7$ kcal/mol.

The entropy of conformational restriction of the hydroxypropyl side chain of ligand **2** and the ethyl side chain of ligand **1** also needs to be considered. Empirical scales give an average value of conformational entropy of about 1.5 cal/mol K per rotatable bond.[38−43] As in our previous work,[27] the conformational entropy change of the ligands between the free and the complex state was estimated by comparing the probability distributions (p) of the dihedral angles in these two states. There are four dihedral angles in the hydroxypropyl group of ligand **2** and two in the ethyl group in ligand **1**. The probabilities (p) (from the Boltzmann expression) of the dihedral angles were obtained with rigid rotations and the integrals of p ln(p) were used to estimate the entropy change, which is $\Delta S_{\mathrm{config}}{}^{\mathbf{2-1}} = -2.7$ cal/mol K, or $-T\Delta S_{\mathrm{config}}{}^{\mathbf{2-1}} = +0.8$ kcal/mol at 300 K.

The contributions to the difference in binding affinity between the two ligands are summarized in Table 6. The calculated change in binding affinity is mentioned above, the difference in binding affinity of the two ligands with the protein CypA is (see Table 6):

$$\Delta\Delta G_{\mathrm{total}}{}^{\mathbf{2-1}} = \Delta\Delta G_{\mathrm{solv}}{}^{\mathbf{2-1}} + \Delta E^{\mathbf{2-1}} + \Delta\Delta G_{\mathrm{desolv}}{}^{\mathbf{2-1}} + (-T\Delta S_{\mathrm{config}}) = +0.8 \pm 0.5 \text{ kcal/mol} \qquad (27)$$

This value is consistent with experiment (+1.3 kcal/mol). No experimental data for the other thermodynamic properties have been reported so far.

Crystallography revealed two different conformations of the ligand in complex **2**.[12] The above results pertain to the major conformer (occupancy 0.62). We repeated the calculations on the structure of complex **2** with ligand **2** in the minor conformation and obtained $\Delta\Delta G_{\mathrm{total}}{}^{\mathbf{2'-1}} = +0.7 \pm 0.5$ kcal/

TABLE 7: Thermodynamic Parameters of the Bound Water Molecules in the HIV-1 Protease−Inhibitor Complex, Con A-Trimannoside Complex, Cyclophilin A−1 Complex, and Cyclophilin A−2 Complex[a]

	bound water	E_{ww}	E_{sw}	ΔE_{solv}	S_{ww}	S_{sw}	$-T\Delta S_{solv}$	ΔG_{solv}
HIV-1 protease	WTR1	0	−28.2	−18.1	0	−25.0	+2.9	−15.2
con A−**1**	WTR1	0	−30.2	−19.2	0	−22.1	+2.0	−17.2
cyclophilin A−**1**	WTR5	−3.8	−25.6	−17.4	−5.5	−18.3	+2.6	−14.8
	WTR6	−15.4	−8.3	−5.9	−3.4	−13.4	+0.5	−5.4
	WTR7	−5.1	−14.1	−6.6	−3.8	−12.2	+0.2	−6.4
	WTR133	−15.6	−3.7	−1.4	−3.5	−10.1	−0.5	−1.9
	cluster	−39.9	−51.7	−31.3	−18.2	−54.0	+2.8	−28.5
cyclophilin A−**2**	WTR5	+0.4	−23.0	−12.7	+0.2	−20.9	+1.6	−11.1
	WTR6	−11.0	−17.7	−13.1	−1.2	−20.5	+2.0	−11.1
	WTR7	+3.4	−24.3	−12.5	−0.2	−21.3	+1.9	−10.6
	cluster	−7.2	−65.0	−38.3	−1.2	−62.7	+5.5	−32.8

[a] Units for the energy are kcal/mol. Units for the entropy are cal/mol K.

mol, quite similar to the result for the major conformer. However, there are significant differences in the contributions, such as those of certain water molecules. WTR5 and 7 in the minor conformer have larger contributions to the solvation free energy than those in the major conformer, while WTR6 shows the opposite result, giving $\Delta\Delta G_{solv}^{2'-1} = -6.4$ kcal/mol.

Discussion

The thermodynamics of isolated bound water molecules at the binding interface of HIV-1 protease−inhibitor complex and concanavalin A-trimannoside complex were explored in our previous work.[26,27] In this work, this approach was extended to the buried water clusters in the CypA−**1** and CypA−**2** complexes. Comparison of all results obtained so far (see Table 7) leads to the following observations.

I. Enthalpic Contributions. The water−water interaction energies in CypA−**1** and CypA−**2** complexes vary in a wide range. The water−water interaction energies of WTR6 and WTR133 in CypA−**1** and WTR6 in CypA−**2** are very negative and even comparable to the water−water interaction energies in bulk water, while they are less negative for WTR5 and WTR7 in CypA−**1**, and even positive for WTR5 and WTR7 in CypA−**2**. Similarly, the interaction energy of a bound water molecule with a protein−ligand complex also varies in a wide range, which approximately corresponds to the number of hydrogen bonds formed between the bound water molecules and the protein−ligand complex. The largest magnitudes of E_{sw} and ΔE_{solv} are observed for the isolated water molecules in the complexes of HIV-1 protease−inhibitor and concanavalin A-trimannoside. We noticed that the bound water molecules with a more negative value of solute−solvent energy usually interact weakly with other water molecules. This is because, when a bound water molecule is strongly constrained by the protein and ligand by several hydrogen bonds, it will be difficult for other solvent molecules to simultaneously approach and interact strongly with it. Nevertheless, E_{sw} and E_{ww} do not fully compensate each other.

II. Entropic Contributions. The solvent−solvent contributions to the solvation entropy in CypA−**1** and CypA−**2** complexes vary in a wide range (from about 0 to −5.5 cal/mol K). Among all of the bound water molecules in these complexes, WTR5 in CypA−**1** has the most negative solvent−solvent entropies, as it is highly correlated with other water molecules (i.e., WTR6 and WTR7, see Table 4) at the binding interface. WTR6, 7, and 133 in CypA−**1** are close, while WTR5 or 7 in CypA−**2** and bound water molecules in HIV-1 protease−inhibitor complex and Con A-trimannoside have water−water entropies close to 0, showing that they are fully buried at the binding interfaces and kept far from other solvent molecules. As with the energy, the largest value for the solute−solvent entropy (S_{sw}) are observed for the isolated water molecules in HIV-1 protease−inhibitor and Con A-trimannoside complexes. From Table 7, we notice that a stronger interaction between the protein and the bound water usually corresponds to a more negative value of the solute−solvent entropy. However, this is not always the case. For instance, in complex **1**, WTR7 interacts more strongly with the protein than WTR6, but the solute−solvent entropies show the opposite result. The broader distributions of the Euler angles, especially the broader distribution of $g(\theta)$ of WTR7 than that of WTR6 in complex **1** (see Figure 5), implies that WTR7 in complex **1** adopts multiple favorable orientations and appears less ordered than WTR6 at the binding interface, although it forms one more H bond with the protein than WTR6.

III. Solvation Free Energy. The values of solvation free energy of the water molecules in CypA−**1** and CypA−**2** reveal that almost all of the water molecules in the water clusters contribute less to protein−ligand binding than those isolated bound water molecules in HIV-1 protease−inhibitor and concanavalin A-trimannoside complexes and that WTR133 has the smallest favorable contribution, which mainly originates from the different water−protein and ligand interactions.

The additional hydroxyl group in complex CypA−**2** forms two more hydrogen bonds with the remaining water molecules in the water cluster (WTR6 and 7) and stabilizes them, which leads to a more favorable contribution to the binding energy (about 7 kcal/mol) even though the longer side chain replaces one less tightly bound water molecule in the water cluster. The additional interactions of **2** with the protein are also favorable for binding. These favorable energy contributions, however, are counterbalanced by the more positive desolvation enthalpy of ligand **2** relative to ligand **1**, leading to a less negative binding energy for ligand **2**. The overall entropy change is also unfavorable, leading to an unfavorable free energy change (see Table 6).

The above analysis differs in some ways from that proposed by Mikol et al.[12] In that work, the contributions of the bound water molecules WTR5 and 7 to the binding affinity were assumed to be the same in the two complexes, the contribution of WTR133 was neglected, and the free energy cost by the displacement of WTR6 together with the desolvation cost of **2** were assumed to be less than 2 kcal/mol. The favorable interaction of CypA−**2** relative to Cyp−**1** was considered as $\Delta E_{ligand-protein}^{2-1} = -0.5$ kcal/mol, which was presumably counterbalanced by contributions due to changes in the translational, rotational, and vibrational degrees of the protein and ligand. Our findings give a different picture. First, the free energy contributions of WTR5 in the two complexes are

different (ca. 3.6 kcal/mol), which presumably originates from the disruption of the hydrogen bond between WTR5 and WTR6 going from **1** to **2**. A similar difference is also found for WTR7, but in the opposite direction due to the formation of a new hydrogen bond between WTR7 and ligand **2**. These two terms compensate each other and add up to $\Delta\Delta G_{\mathrm{solv(w5+w7)}}{}^{\mathbf{2-1}} = -0.6$ kcal/mol. Second, WTR133 does make a small contribution to the free energy in Cyp−**1** ($\Delta G_{\mathrm{solv(w133)}} = -1.9$ kcal/mol). Third, the free energy cost by the displacement of WTR6 was obtained as $\Delta\Delta G_{\mathrm{solv(w6)}}{}^{\mathbf{2-1}} = -5.8$ kcal/mol, combined with the different desolvation free energy $\Delta\Delta G_{\mathrm{desolv}}{}^{\mathbf{2-1}} = +9.6$ kcal/mol, giving +3.8 kcal/mol, which is even larger than the corresponding upper bound proposed by Mikol et al.. The calculated difference in conformational entropy loss of the Abu side chain of **1** and **2** ($-T\Delta S_{\mathrm{config}}{}^{\mathbf{2-1}} = +0.8$ kcal/mol) is similar to the value they assumed. However, the calculated additional interaction energy of ligand **2** with the protein($\Delta E_{\mathrm{ligand-protein}}{}^{\mathbf{2-1}} = -1.4$ kcal/mol) is much larger than they assumed.

The good agreement between the calculated and experimental change in affinity is at least somewhat fortuitous, given the numerous approximations made in the calculations. In the calculation of the contributions of solvent−solvent entropies, we assumed that the inhomogeneous pair correlation function is equal to that in bulk solvent. We also neglected the conformational entropy difference arising from the same parts of the two ligands. The desolvation enthalpy and entropy of protein in the two complexes were assumed to be the same, and the desolvation terms of the ligands were estimated according to their solvent accessible surface area, which is an empirical model. In addition, this work focused exclusively on the three or four bound water molecules at the binding interface and assumed that the contribution of all other water molecules is the same in the two complexes. Applying this theoretical analysis to all water molecules in the system is currently not feasible. All these approximations, together with the standard uncertainties related to force field parameters, make the true uncertainties much larger than the error bars listed in Table 1, 2, and 6. However, the qualitative picture that arises from the calculations should be reliable.

From the results obtained so far, gain of binding free energy from the displacement of water might be achieved under the following conditions: (i) the new ligand must interact strongly with the protein and the remaining water molecules, (ii) the additional group ideally should not disrupt interactions between bound water molecules, and (iii) the conformational entropy loss should be minimized if possible, for example, by use of a carbonyl group instead of a hydroxyl group. This is consistent with the rules introduced by Garcia et al.[44]

This approach can be useful in drug design by helping to distinguish the different role bound water molecules play in the complex and which of them is the most favorable to displace. In addition, the insights obtained from this analysis could help improve empirical scoring functions for prediction of protein−ligand binding affinity by taking the thermodynamic contributions of bound water molecules into account.

Acknowledgment. This work was supported by the National Science Foundation (MCB-031667).

References and Notes

(1) Niefind, K.; Putter, M.; Guerrra, B.; Issinger, O. G.; Schomburg, D. *Nat. Struct. Biol.* **1999**, *6*, 1100−1103.

(2) Zheng, J. H.; Trafny, E. A.; Knighton, D. R.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. *Acta Crystallogr., Sect. D* **1993**, *49*, 362−365.

(3) Otwinowski, Z.; Schevitz, R. W.; Zhang, R.-G. Sigler, P. B. *Nature* **1988**, *335*, 321−329.

(4) Joachimiak, A.; Haran, T. E.; Sigler, P. B. *EMBO J.* **1994**, *13*, 367−372.

(5) Bhat, T. N.; Bentley, G. A.; Fishmann, T. O.; Boulot, G.; Poljak, R. J. *Nature* **1990**, *347*, 483−485.

(6) Bhat, T. N.; Poljak, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1089−1093.

(7) Quiocho, F. A.; Wilson, D. K.; Vyas, N. K. *Nature* **1989**, *340*, 404−407.

(8) Lam, P. Y. S.; Jadhav, P. K. Erickson-Viitanen, S. *Science* **1994**, *263*, 380−3844.

(9) Tame, J. R. H.; Murshudov, G. N.; Dodson, E. J.; Neil, T. K.; Wilkinson, A. J. *Science* **1994**, *264*, 1578−1581.

(10) Sleigh, S. H.; Tame, J. R. H.; Dodson, E. J.; Wilkinson, A. J. *Biochemistry* **1997**, *36*, 9747−9758.

(11) Holdgate, G. A.; Tunnicliffe, A.; Ward, W. H. J. Timms, D. *Biochemistry* **1997**, *36*, 9663−9673.

(12) Mikol, V.; Papageorgiou, C.; Borer, X. *J. Med. Chem.* **1995**, *38*, 3361−3367.

(13) Binda, C.; Hubalek, F.; Li, M.; Herzig, Y.; Sterling, J.; Edmondson, D.; Mattevi, A. *J. Med. Chem.* **2004**, *47*, 1767−1774.

(14) Ni, H. H.; Sotriffer, C. A.; McCammon, J. A. *J. Med. Chem.* **2001**, *44*, 3043−3047.

(15) Wang, Y.-X.; Freedberg, D. I.; Wingfield, P. T. Torchia, D. A. *J. Am. Chem. Soc.* **1996**, *118*, 12287−12290.

(16) Hubalek, F.; Binda, C.; Li, M.; Herzig, Y.; Sterling, J.; Youdim, M. B. H.; Mattevi, A.; Edmonds, K. A. *J. Med. Chem.* **2004**, *47*, 1760−1766.

(17) Ladbury, J. E. *Chem. Biol.* **1996**, *3*, 973−980.

(18) Clarke, C.; Woods, R. J.; Gluska, J.; Cooper, A.; Nutley, M. A.; Boons, G. J. *J. Am. Chem. Soc.* **2001**, *123*, 12238−12247.

(19) Loris, R.; Maes, D.; Poortmans, F.; Wyns, L.; Bouckaert, J. *J. Biol. Chem.* **1996**, *271*, 30614−30618.

(20) Yokota, A.; Tsumoto, K.; Shiroishi, M.; Kondo, H.; Kumagai, I. *J. Biol. Chem.* **2003**, *278*, 5410−5418.

(21) Sunnerhagen, M.; Denisov, V. P.; Venu, K.; Bonvin, A. M. J. J.; Carey, J.; Halle, B.; Otting, G. *J. Mol. Biol.* **1998**, *282*, 847−858.

(22) Sharrow, S. D.; Edmonds, K. A.; Goodman, M. A.; Novotny, M. V.; Stone, M. J. *Protein Sci.* **2005**, *14*, 249−256.

(23) Wang, R. X.; Lu, Y. P.; Wang, S. M. *J. Med. Chem.* **2003**, *46*, 2287−2303.

(24) Rarey, M.; Kramer, B.; Lengauer, T. *Proteins: Struct., Funct., Genet.* **1999**, *4*, 17−28.

(25) Fornabaio, M.; Spyrakis, F.; Mozzarelli, A.; Cozzini, P.; Abraham, D. J.; Kellogg, G. E. *J. Med. Chem.* **2004**, *47*, 4507−4516.

(26) Li, Z.; Lazaridis, T. *J. Am. Chem. Soc.* **2003**, *125*, 6636−6637.

(27) Li, Z.; Lazaridis, T. *J. Phys. Chem.* **2005**, *109*, 662−670.

(28) O'Keefe, S. J.; Tamra, J.; Kincaid, R. L.; Tocci, M. J.; O'Neil, E. A. *Nature* **1992**, *357*, 692−694.

(29) Mikol, V.; Kallen, J.; Pflugl, G.; Walkinshaw, M. D. *J. Mol. Biol.* **1993**, *234*, 1119−1130.

(30) Lazaridis, T. *J. Phys. Chem. B* **1998**, *102*, 3542−3550.

(31) Lazaridis, T. *J. Phys. Chem. B* **2000**, *104*, 4964−4979.

(32) Lazaridis, T.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 4294−4316.

(33) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639−659.

(34) Miller, S.; Janin, J.; Lesk, A. M.; Chothia, C. *J. Mol. Biol.* **1987**, *196*, 641−656.

(35) Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1993**, *232*, 660−679.

(36) MacKerell, A. D., Jr.; Bashford, D.; Bellott, R. L.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(37) Brooks, C. L.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.

(38) Wang, J.; Szewczuk, Z.; Yue, S.-Y.; Tsuda, Y.; Konishi, Y.; Purisima, E. O. *J. Mol. Biol.* **1995**, *253*, 473−492.

(39) Luo, R.; Gilson, M. K. *J. Am. Chem. Soc.* **2000**, *122*, 2934−7.

(40) Pickett, S. D.; Sternberg, M. J. E. *J. Mol. Biol.* **1993**, *231*, 825−839.

(41) Nicholls, A.; Sharp, K. A.; Honig, B. *Protein Sci.* **1991**, *11*, 281−296.

(42) Creamer, T. P. *Protein Sci.* **2000**, *40*, 443−450.

(43) D'Aquino, J. A.; Freire, E.; Amzel, L. M. *Protein Sci.* **2000**, (Suppl. 4), 93−107.

(44) Garcia-Sosa, A. T.; Firth-Clark, S.; Mancera, R. L. *J. Chem. Inf. Model.* **2005**, *45*, 624−633.