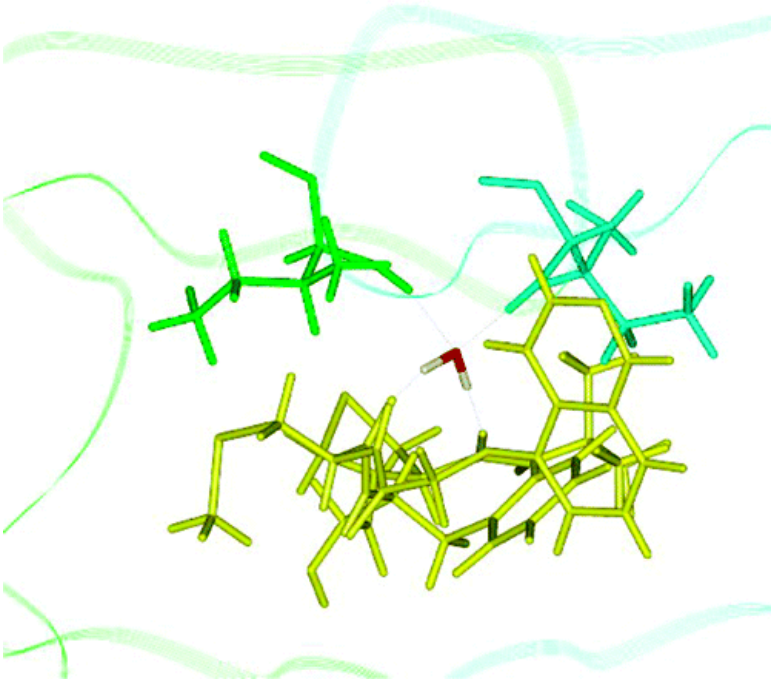
Communication

# Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease

Zheng Li, and Themis Lazaridis

*J. Am. Chem. Soc.*, **2003**, 125 (22), 6636-6637• DOI: 10.1021/ja0299203 • Publication Date (Web): 10 May 2003





J|A|C|S
COMMUNICATIONS

Published on Web 05/10/2003

# Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease

Zheng Li and Themis Lazaridis*

*Department of Chemistry, City College of CUNY, Convent Avenue and 138th Street, New York, New York 10031*

Received December 25, 2002; E-mail: themis@sci.ccny.cuny.edu

Binding between biomolecules is usually accompanied by the formation of direct interactions with displacement of water from the binding sites. In some cases, however, the interactions are mediated by ordered water molecules,[1] whose effect on binding affinity and the other thermodynamic functions is unclear. In this work, we compute the contribution of one such water molecule to the thermodynamic properties using statistical mechanical formulas for the energy and entropy. The requisite correlation functions are obtained by molecular dynamics (MD) simulations. We find that the entropic penalty of ordering is large but is outweighed by the favorable water−protein interactions. We also find a large negative contribution from this water molecule to the heat capacity.

Most theoretical work so far has focused on water in internal protein cavities and on the free energy only.[2,3] Dunitz[4] provided an empirical upper bound of 7 cal/mol K (about 2 kcal/mol at 300 K) for the entropy cost of transferring a water molecule from the bulk to a binding site. Similar estimates for the enthalpy of water ordering gave −3.8 kcal/mol.[1] Therefore, the free energy of ordering should be favorable. However, these estimates do not consider the specific binding interface and interactions that the bound water experiences with the biomolecule or the inhibitor, and thus they cannot provide much guidance as to whether water displacement is favorable or unfavorable in specific cases.

One approach that could prove useful in this regard is the inhomogeneous fluid solvation theory.[5,6] In this approach, the solvation energy and entropy are decomposed into the solute−solvent terms ($E_{sw}$, $S_{sw}$) and solvent reorganization terms ($\Delta E_{ww}$, $\Delta S_{ww}$). $S_{sw}$ is expressed as an integral over the solute−solvent correlation function $g_{sw}(\mathbf{r},\omega)$:

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

where $k$ is Boltzmann's constant, $\rho$ is the density of bulk water, and $\Omega$ is the integral over $\omega$ (the orientation of the solvent with respect to the solute).

Using the identity

$$g_{sw}(\mathbf{r},\omega) = g^{tr}_{sw}(\mathbf{r}) g^{or}_{sw}(\omega|\mathbf{r}) \qquad (2)$$

we can decomposed the above integral into a translational and an orientational contribution:

$$S_{sw} = -k\rho \int g^{tr}_{sw}(\mathbf{r}) \ln g^{tr}_{sw}(\mathbf{r})\, d\mathbf{r} - k\rho/\Omega \int d\mathbf{r}\, g^{tr}_{sw}(\mathbf{r}) \int g^{or}_{sw}(\omega|\mathbf{r}) \ln g^{or}_{sw}(\omega|\mathbf{r})\, d\omega \qquad (3)$$

If we further assume that $g^{or}_{sw}(\omega|\mathbf{r})$ is independent of $\mathbf{r}$ within a certain region of space $v$ ($g^{or}_{sw}(\omega|\mathbf{r}) \approx g^{or}_{sw}(\omega)$) and restrict the range of integration within $v$, eq 3 becomes

$$S^{v}_{sw} = -k\rho \int_v g^{tr}_{sw}(\mathbf{r}) \ln g^{tr}_{sw}(\mathbf{r})\, d\mathbf{r} - kN_{wat}/\Omega \int g^{or}_{sw}(\omega) \ln g^{or}_{sw}(\omega)\, d\omega \qquad (4)$$

where $N_{wat}$ is the number of water molecules in that region of space ($N_{wat} = \rho \int_v g^{tr}_{sw}(\mathbf{r})\, d\mathbf{r}$). The solute−solvent energy can also be written as an integral:

$$E_{sw} = \rho/\Omega \int g_{sw}(\mathbf{r},\omega) u_{sw}(\mathbf{r},\omega)\, d\mathbf{r}\, d\omega \qquad (5)$$

where $u_{sw}$ is the potential energy, but it is more easily evaluated directly from a simulation.

The solvent terms, $\Delta E_{ww}$ and $\Delta S_{ww}$, can also be expressed as integrals over the space around the solute. Thus, the contribution of specific regions of space to the solvation properties can be determined. This approach can be applied over regions occupied by bound water molecules in biomolecular complexes to provide a rigorous estimate of the contribution of such molecules to the thermodynamic functions. The contribution of a solvent molecule to the solvation energy, entropy, and free energy is equivalent to the energy, entropy, and free energy difference between bulk and "interfacial" solvent molecules (upon insertion of a solute into the solvent, a number of bulk solvent molecules become interfacial solvent molecules).

One case where the displacement of a bound water molecule was thought to be favorable to binding is HIV-1 protease. The crystal structures of the HIV-1 protease with a number of inhibitors[7−10] show an ordered water molecule donating two hydrogen bonds to the inhibitor and accepting two hydrogen bonds from the protein "flaps". NMR studies showed that this water molecule has a long residence time.[11] In the work reported by Lam et al.,[12] cyclic urea inhibitors were designed to displace and mimic the interactions of this water molecule and were found to bind more strongly to HIV-1 protease. The crystal structure of the HIV-1 protease-DMP450 complex[13] shows the oxygen atom of the cyclic urea carbonyl group accepting two hydrogen bonds from the protein flaps.

Here we calculate the contributions of this water molecule to the energy, entropy, and heat capacity of solvation using the inhomogeneous fluid theory.[5,6] MD simulations were performed on two HIV-1 protease-inhibitor complexes: the complex with KNI-272 (pdb code 1HPX) where the ordered water molecule is present, and the complex with DMP450 (pdb code 1DMP) where it has been displaced. The simulations started from the crystal structures and lasted 8 ns at 300 K, keeping the protein and the inhibitors fixed. In addition to the crystal water molecules, a 15 Å sphere of TIP3P water molecules was added around the active site. The charmm22 force field was used for the protein. Partial charges for the inhibitors were obtained with QUANTA (Accelrys, Inc.) using charmm template charges.

We first calculated the translational correlation function $g^{tr}_{sw}(r,\theta',\phi')$, where $r$, $\theta'$, and $\phi'$ are spherical coordinates of the water oxygen with respect to its average position. The radial distribution function ($g^{tr}_{sw}(\mathbf{r})$ averaged over $\theta'$ and $\phi'$) is shown in Figure 1. This function gives the local density relative to bulk water, which can be seen to be very high. From the Euler angles of this water molecule in each frame, we also calculated an average orientational correlation function $g^{or}_{sw}(\theta,\phi,\psi)$ over the region occupied by this water molecule (Figure 2 shows the distribution of



10.1021/ja0299203 CCC: $25.00 

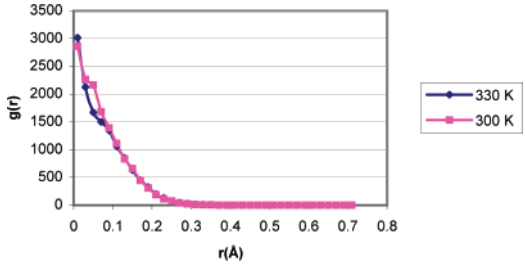



***Figure 1.*** Radial distribution function $g^{r}_{sw}(r)$ at 300 and 330 K. $r$ is the distance of the key water molecule from its average position.

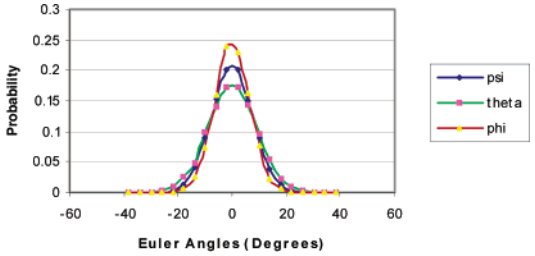



***Figure 2.*** Probability distribution of the Euler angles around their average values.

each angle). The calculation of the integrals in eq 4 was done using either the full three-dimensional functions $g^{tr}_{sw}(r,\theta',\phi')$ and $g^{or}_{sw}(\theta,\phi,\psi)$ or the factorization approximations $g^{tr}_{sw}(r,\theta',\phi') = g^{r}_{sw}(r)g_{sw}(\theta')g_{sw}(\phi')$ and $g^{or}_{sw}(\theta,\phi,\psi) = g_{sw}(\theta)g_{sw}(\phi)g_{sw}(\psi)$, with similar results. The values below are those obtained by the factorization approximations. The translational and the orientational contributions to the solute−solvent entropy were found to be −12.5 cal/mol K each, giving $S_{sw} = -25.0$ cal/mol K. The solute−solvent energy $E_{sw}$ was calculated directly from the simulation to be −28.2 kcal/mol.

The calculation of the solvent reorganization terms in this case is facilitated by the fact that this water molecule does not interact significantly with other water molecules when bound to the protease (the calculated interaction energy $E_{ww}$ is −0.3 kcal/mol). Therefore, the $\Delta E_{ww}$ and $\Delta S_{ww}$ are simply the energy and entropy of removing a water molecule from bulk water, that is, $\Delta E_{ww} = +10.1$ kcal/mol and $\Delta S_{ww} = +15.2$ cal/mol K.[14] Therefore, $\Delta E = -18.1$ kcal/mol and $\Delta S = -9.8$ cal/mol K. This value is larger than the upper bound estimated by Dunitz,[4] showing the extremely high degree of ordering of this water molecule. The contribution of the water molecule to the solvation free energy at 300 K is ($P\Delta V$ term is negligible)

$$\Delta G_{solv} = E_{sw} + \Delta E_{ww} - T(S_{sw} + \Delta S_{ww}) = -15.2 \text{ kcal/mol} \quad (6)$$

To examine the dependence of the calculated solvation properties on the protein/inhibitor configuration, we repeated the calculation on the structure obtained after 1 ns of MD simulation, with the inhibitor and the residues of the protein within a radius 13 Å free to move. We obtained $S_{sw} = -23.3$ cal/mol, $E_{sw} = -28.4$ kcal/mol, and $\Delta G_{solv} = -15.9$ kcal/mol, quite similar to the values above.

To calculate the contribution of this water molecule to the heat capacity, we repeated the MD simulation at 330 K. We obtained $S_{sw} = -24.5$ cal/mol K and $E_{sw} = -28.0$ kcal/mol. From the properties of bulk water, we obtained $\Delta E_{ww} = +9.6$ kcal/mol and $\Delta S_{ww} = +13.5$ cal/mol K.[14] Therefore, $\Delta E = -18.4$ kcal/mol and $\Delta S = -11.0$ cal/mol K, and the contribution of the water molecule to the heat capacity of solvation is[6]

$$\Delta C_p = \left(\frac{d\Delta H}{dT}\right)_p \cong \frac{\delta E_{sw}}{\delta T} + \frac{\delta \Delta E_{ww}}{\delta T} = 7 - 17 = -10 \text{ cal/mol K} \quad (7)$$

An alternative calculation gives

$$\Delta C_p = T\left(\frac{d\Delta S}{dT}\right)_p = -12 \text{ cal/mol K} \quad (8)$$

The negative value of $\Delta C_p$ is due to the fact that the decrease of solvent reorganization energy and entropy with temperature is faster than the decrease of the protein−water interaction energy and entropy. In other words, the interactions of the water molecule in the binding site are less susceptible to temperature than bulk water. This finding is in agreement with the proposal that ordered water molecules contribute to the negative heat capacity observed in protein−DNA complexes.[15]

For the protease/DMP450 complex, we calculated the interaction energy of the carbonyl group with the protein to be −16.9 kcal/mol. This value is more negative than $\Delta G_{solv}$ for the water molecule. However, the cost of desolvating a carbonyl group is about +5 kcal/mol.[16] Therefore, the displacement of the key crystal water molecule by insertion of a carbonyl group per se should be unfavorable. On the other hand, two carbonyl groups interacting with the key water molecule are eliminated in going from KNI-272 to DMP450, reducing the desolvation cost. Consideration of the above contributions leads to the conclusion that the specific way water displacement was achieved in this case was favorable for binding. Of course, because these two inhibitors are entirely different, one would need a comprehensive account of all other differences for a reliable attribution of the gain in binding affinity.

The present work shows that inhomogeneous fluid solvation theory can yield the thermodynamic contributions of ordered water molecules. One advantage of this method over standard free energy simulations is that it is also applicable to water molecules that are not fully buried; such calculations are in progress. Other advantages include the facility by which it yields all thermodynamic functions (not just the free energy), easier assessment of sampling adequacy, and a lower computational cost. This approach could be useful in rational drug design by estimating which bound water molecules would be most favorable to displace.

**Acknowledgment.** This work was supported by the National Science Foundation (DBI-9979621 and HRD-0206162).

## References

(1) Ladbury, J. E. *Chem. Biol.* **1996**, *3*, 973.
(2) Wade, R. C.; Mazor, M. H.; McCammon, J. A.; Quiocho, F. A. *Biopolymers* **1991**, *31*, 919.
(3) Roux, B.; Nina, M.; Pomes, R.; Smith, J. C. *Biophys. J.* **1996**, *71*, 670.
(4) Dunitz, J. D. *Science* **1994**, *264*, 670.
(5) Lazaridis, T. *J. Phys. Chem.* **1998**, *102*, 3531.
(6) Lazaridis, T. *J. Phys. Chem. B* **2000**, *104*, 4964.
(7) Baldwin, E. T.; Bhat, T. N.; et al.; Erickson, J. W. *Structure* **1995**, *3*, 581.
(8) Louis, J. M.; Dyda, F.; Nashed, N. T.; Kimmel, A. R.; Davies, D. R. *Biochemistry* **1998**, *37*, 2105.
(9) Hong, L.; Zhang, X. J.; Foundling, S.; Hartsuck, J. A.; J., T. *FEBS Lett.* **1997**, *420*, 11.
(10) Kervinen, J.; Thanki, N.; Zdanov, A.; Tina, J.; Barrish, J.; Lin, P. F.; Colonno, R.; Riccardi, K.; Samanta, H.; Wlodawer, A. *Protein Pept. Lett.* **1996**, *3*, 399.
(11) Wang, Y.-X.; Freedberg, D. I.; Wingfield, P. T.; et al.; Torchia, D. A. *J. Am. Chem. Soc.* **1996**, *118*, 12287.
(12) Lam, P. Y. S.; Jadhav, P. K.; et al.; Erickson-Viitanen, S. *Science* **1994**, *263*, 380.
(13) 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. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. *Chem. Biol.* **1996**, *3*, 301.
(14) Lazaridis, T.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 4294.
(15) Morton, C. J.; Ladbury, J. E. *Protein Sci.* **1996**, *5*, 2115.
(16) Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1993**, *232*, 660.

JA0299203