J|A|C|S
ARTICLES
Published on Web 06/05/2004

Hydration Free Energies and Entropies for Water in Protein Interiors

L. Renee Olano† and Steven W. Rick*,†,‡

Contribution from the Department of Chemistry, University of New Orleans, New Orleans, Louisiana, 70148 and Chemistry Department, Southern University of New Orleans, New Orleans, Louisiana 70126

Received January 16, 2004; E-mail: srick@uno.edu

Abstract: Free energy calculations for the transfer of a water molecule from the pure liquid to an interior cavity site in a protein are presented. Two different protein cavities, in bovine pancreatic trypsin inhibitor (BPTI) and in the I76A mutant of barnase, represent very different environments for the water molecule: one which is polar, forming four water−protein hydrogen bonds, and one which is more hydrophobic, forming only one water−protein hydrogen bond. The calculations give very different free energies for the different cavities, with only the polar BPTI cavity predicted to be hydrated. The corresponding entropies for the transfer to the interior cavities are calculated as well and show that the transfer to the polar cavity is significantly entropically unfavorable while the transfer to the nonpolar cavity is entropically favorable. For both proteins an analysis of the fluctuations in the positions of the protein atoms shows that the addition of a water molecule makes the protein more flexible. This increased flexibility appears to be due to an increased length and weakened strength of protein−protein hydrogen bonds near the cavity.

1. Introduction

The interiors of globular proteins are well packed, but cavities large enough to contain at least one water molecule are found in nearly all globular proteins and can account for about 1% of the total protein volume.[1,2] Buried water molecules have been observed by nuclear magnetic resonance (NMR), X-ray crystallography, and neutron diffraction methods[3,4] and occupy about 18% of the cavity sites, giving on average one buried water per 27 residues.[2] Biologically important roles for buried waters have been assigned for ligand binding (for example, HIV protease[5] and hemoglobin[6]), protein−protein (including antibody−antigen) association,[7] protein stability,[8] and protein flexibility.[9−11] The ability to measure or predict the position of buried water molecules can be important to understanding the catalytic mechanism of enzymes. For example, the reaction mechanism of carboxy peptidase A may depend on the number of water molecules in the active site, which is not known from the crystal structure.[12] The amount of experimental data on buried water has increased rapidly over the past few years. These data, together with computational studies,[9,11,13−17] suggest interesting questions concerning the conditions which favor the occupation of water molecules.

Some buried waters that are observed by X-ray crystallography are also observed by NMR spectroscopy and neutron diffraction and under a variety of crystallization conditions.[3] Bovine pancreatic trypsin inhibitor (BPTI), for example, has four internal water molecules (one isolated and three in a cluster) which are observed by X-ray, neutron, and NMR methods, and the isolated water site may be occupied in the gas phase as well.[18−22] This water molecule, labeled W122 in the 5PTI crystal

† University of New Orleans.
‡ Southern University of New Orleans.

(1) Rashin, A. A.; Iofin, M.; Honig, B. *Biochemistry* **1986**, *25*, 3619−3625.
(2) Williams, M. A.; Goodfellow, J. M.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 1224−1235.
(3) Levitt, M.; Park, B. H. *Structure* **1993**, *1*, 223−226.
(4) Schoenborn, B. P.; Garcia, A.; Knott, R. *Prog. Biophys. Mol. Biol.* **1995**, *64*, 105−119.
(5) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneiderm, J.; Kent, S. B. H. *Science* **1989**, *245*, 616−621.
(6) Royer, W. E., Jr.; Pardanani, A.; Gibson, Q. H.; Peterson, E. S.; Friedman, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14526−14531.
(7) Bhat, T. N.; Bentley, G. A.; Boulot, G.; Greene, M. I.; Tello, D.; DallAcqua, W.; Souchon, H.; Schwarz, F. R.; Mariuzza, R. A.; Poljak, F. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1089−1093.
(8) Takano, K.; Funahashi, J.; Yamagata, Y.; Fujii, S.; Yutani, K. *J. Mol. Biol.* **1997**, *274*, 132−142.
(9) Fischer, S.; Verma, C. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9613−9615.
(10) Smith, S.; Smith, J. C.; Verma, C. S. *J. Phys. Chem. B* **2001**, *105*, 8050−8055.
(11) Mao, Y.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2000**, *122*, 2950−2951.
(12) Abashkin, Y. G.; Burt, S. K.; Collins, J. R.; Cachau, R. E.; Russo, N.; Erickson, J. W. Density functional modelling of ligand-metal interactions and enzymatic activities in metalloenzymes. In *Metal-ligand interactions—structure and reactivity*; Russo, N., Salahub, D., Eds.; Kluwer Academic: Norwell, MA, 1995; pp 1−22.
(13) Wade, R. C.; Mazor, M. H.; McCammon, J. A.; Quiocho, F. A. *J. Am. Chem. Soc.* **1990**, *112*, 7057−7059.
(14) Wade, R. C.; Mazor, M. H.; McCammon, J. A.; Quiocho, F. A. *Biopolymers* **1991**, *31*, 919−931.
(15) Helms, V.; Wade, R. C. *Biophys. J.* **1995**, *69*, 810−824.
(16) Zhang, L.; Hermans, J. *Proteins: Struct. Funct. Gen.* **1996**, *24*, 433−438.
(17) Roux, B.; Nina, M.; Pomès, R.; Smith, J. C. *Biophys. J.* **1996**, *71*, 670−681.
(18) Deisenhofer, J.; Steigemann, W. *Acta Crystallogr., Sect. B* **1975**, *31*, 238−250.
(19) Wlodawer, A.; Walter, J.; Huber, R.; Sjölin, L. *J. Mol. Biol.* **1984**, *180*, 301−329.
(20) Wlodawer, A.; Nachman, J.; Gilliland, G. L.; Gallagher, W.; Woodward, C. *J. Mol. Biol.* **1987**, *198*, 469−480.

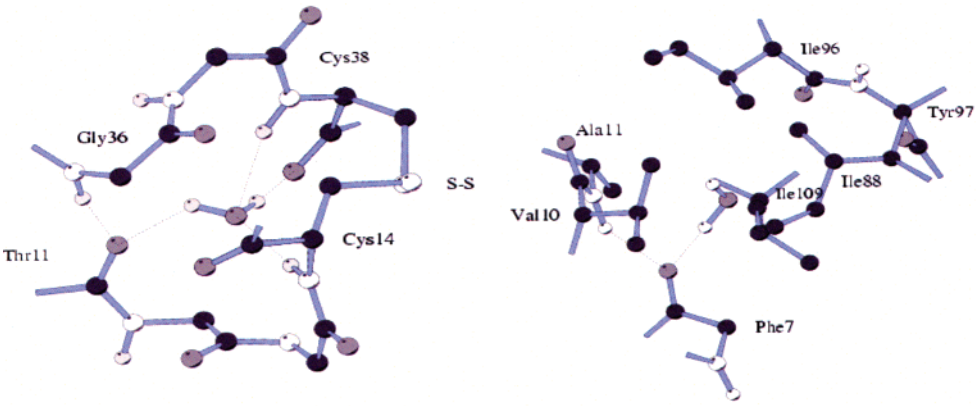

Figure 1. (a, left) Water molecule W122 in the interior of BPTI, showing the hydrogen bonds to the protein backbone. The coordinates are from the 5PTI crystal structure.[19] (b, right) Water molecule W274 in the interior of the I76A mutant of barnase, from the 1BRI crystal structure.[23] The figure was made using MOLSCRIPT.[24]

structure of BPTI,[19] which forms four hydrogen bonds to the protein, can be thought of as an integral part of the protein (see Figure 1a). Another tetrahedrally coordinated water is observed in most structures of HIV-1 protease complexes with peptidomimetic inhibitors.[25]

The occupancy of water in nonpolar cavities, most often considered to be empty, is more controversial. NMR, which may be able to detect water molecules with lower occupancies and more disorder, has detected buried water in nonpolar cavities which are not visible in X-ray structures.[26,27] In hen egg-white lysozyme, NMR experiments detected water in two different cavities, in which each can form at most one hydrogen bond with the protein.[27] There are X-ray structures indicating water in nonpolar cavities as well. A 1994 analysis of 75 structures of resolution higher than 2.5 Å revealed that 18% of the buried water molecules had two or fewer hydrogen bonds to the protein or other buried waters.[2] However, Zhang and Hermans have argued that these water molecules mostly have low occupancies and high *B*-factors and are artifacts of the crystal refinement process.[16] After the Zhang and Hermans paper, X-ray structures of barnase[23] and human lysozyme[8] mutants (in which cavities are created by mutations to smaller, nonpolar residues such as alanine and glycine) reveal buried waters, with one or two hydrogen bonds. From their *B*-factors and occupancies, in the terminology of Zhang and Hermans these waters would be characterized as "strong waters" and not as easily dismissed as artifacts. The barnase mutant (Ile 76 → Ala) structure reveals a buried water which makes only one hydrogen bond to the protein in a largely hydrophobic environment (Figure 1b). Free energy calculations of the transfer of water molecules from the liquid to interior cavities find that the process is spontaneous only for polar cavities and there is a good correlation with a negative free energy change and the presence of an observed water molecule in that cavity.[13−16] However, the calculated free energy to transfer to one of the crystallographically unoccupied cavities, which is largely nonpolar and can form only two possible hydrogen bonds, is 0.2 ± 1.5 kcal/mol, giving the surprising result that there is no free energy cost in transferring a water molecule from the liquid to an environment often characterized as hydrophobic.[14] Taken together, the experimental and computational studies indicate that some nonpolar cavities may in fact be hydrated. It is possible that water in these cavities is stabilized by polarizability[26] or entropy.[26,28] Alternatively, the stabilization may be through C−H···O hydrogen bonds as has recently been proposed for the peptide backbone and other systems.[29,30] Contacts between oxygen atoms on water and C−H groups have also been observed by neutron diffraction.[31]

For larger cavities, even if the protein can form no hydrogen bonds, water can be energetically stabilized through water−water hydrogen bonds. There are a number of proteins with cavities large enough to hold two or more water molecules which show evidence of containing water even though though these cavities are completely lined with nonpolar groups. These proteins include human human interleukin-1β (hIL-1β),[26,32] hen egg-white lysozyme,[27] staphylococcal nuclease,[33] and trypsin.[34] In hIL-1β, there is a large central hydrophobic cavity which can accommodate two to four water molecules surrounded by hydrophobic residues and which offers no potential hydrogen bonds to the protein. NMR data indicate that there is water in this cavity, which was not indicated by a previous crystal structure.[26] The presence of water in this cavity has been the subject of debate,[16,35,36] but a recent crystal structure indicated that there are water molecules in this cavity with partial occupancies.[32] Unpublished free energy calculations cited in ref 16 indicate that the hIL-1β does not contain water, so there is a significant disagreement between these calculations and the

(21) Otting, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1989**, *111*, 1871−1875.
(22) Woenckhaus, J.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 9586−9587.
(23) Buckle, A. M.; Cramer, P.; Fersht, A. R. *Biochemistry* **1996**, *35*, 4298−4305.
(24) Kraulis, P. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(25) Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543−585.
(26) Ernst, J. A.; Clubb, R. T.; Zhou, H. X.; Gronenborn, A. M.; Clore, G. M. *Science* **1995**, *267*, 1813−1817.
(27) Otting, G.; Liepinsh, E.; Halle, B.; Frey, U. *Nat. Struct. Biol.* **1997**, *4*, 396−404.
(28) Denisov, V. P.; Venu, K.; Peters, J.; Hörlein, H. D.; Halle, B. *J. Phys. Chem. B* **1997**, *101*, 9380−9389.
(29) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750−4755.
(30) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963−4968.
(31) Bon, C.; Lehman, M. S.; Wilkinson, C. *Acta Crystallogr., Sect. D* **1999**, *55*, 978−987.
(32) Yu, B.; Blaber, M.; Gronenberg, A. M.; Clore, G. M.; Caspar, D. L. D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 103−108.
(33) Dwyer, J. J.; Gittis, A. G.; Karp, D. A.; Lattman, E. E.; Spencer, D. S. *Biophys. J.* **2000**, *79*, 1610−1620.
(34) Kossiakoff, A. A.; Sintchak, M. D.; Shpungin, J.; Presta, L. G. *Proteins: Struct. Funct. Gen.* **1992**, *12*, 223−236.
(35) Matthews, B. W.; Morton, A. G.; Dahlquist, F. W. *Science* **1995**, *270*, 1847−1848.
(36) Ernst, J. A.; Clubb, R. T.; Zhou, H. X.; Gronenborn, A. M.; Clore, G. M. *Science* **1995**, *270*, 1848−1849.

NMR and X-ray data. Other large cavities, including one in ribonuclease S, are reported to be empty of water.[37] For the larger cavities as well as the single molecule cavities, no agreement has been reached about the conditions for hydration.

The entropy of transfer between liquid and buried water, with a rigid hydrogen bond arrangement (for example, W122 in BPTI), can be estimated based on the water to ice entropy change as −7 cal/(mol K).[38] For less rigid waters, as indicated by the librational amplitudes, the entropy may be smaller and the entropy change may compensate for a small enthalpy change. For buried water molecules in BPTI, the transfer entropy is estimated to be as low as −1 cal/(mol K), and this is for water molecules with three hydrogen bonds to the protein.[28] Direct measurements for BPTI in the gas phase using mass spectroscopy indicate a large entropy change for binding of -62 ± 5 cal/(mol K).[22] When compared to the entropy of vaporization at the same temperature (around 0 °C), this gives a entropy of transfer from liquid water to, putatatively, the W122 site in gas phase BPTI equal to −33 cal/(mol K), much larger than the Dunitz estimate from ice or crystal hydrates. A calculation using inhomogeneous fluid solvation theory of the entropy of the HIV-1 protease buried water also finds a large entropy change (−9.8 cal/(mol K)), which is also larger than the Dunitz estimate.[39] It is important to fully characterize the entropy of buried water since it is assumed to play a role in ligand−protein and protein−protein association.[28,40−42] The release of bound water, displaced by a ligand or a protein, or the formation of an ordered water in the binding interface may lead to a large entropy change which will affect the thermodynamics of the process. If factors such as entropy and polarization are important to the stability of proteins, then commonly used procedures for predicting buried water positions, such as GRID[43] and AUTOSOL,[44] which are based on an energetic analysis of putative binding sites, may have to be modified in order to successfully predict hydration in a range of environments.

Another significance of the entropy change is its connection to protein flexibility.[9−11,45,46] Part of the entropy change is the vibrational entropy of the protein due to changes in conformational freedom. Exactly opposite conclusions about the influence of strongly ordered water molecules on protein flexibility have been reached (and for the same protein[47]).[9,11] For the binding of other, larger ligands and for the dimerization of proteins, greater conformational freedom was found for the associated state.[45,46] This increase in conformational freedom is for the protein−ligand or protein−protein complex relative to the pure protein in water, reflecting how the flexibility changes when water is displaced by the ligand or protein. These studies indicate that no consensus has been reached as to whether a molecule which binds rigidly and fits well in the binding site increases or decreases the flexibility of the protein. This influence on protein flexibility may be strongly dependent on the type and strength of the protein−water interaction, and water in the two cavities presented here may influence the flexibility in different ways.

Because the ligand is also the solvent, experimental measurements of the binding thermodynamics for water are not possible using many of the methods developed for other ligands. Computational methods are therefore particularly useful. There are very interesting experimental data pertaining to the thermodynamics of water binding. The free energy for the binding of water has been measured in the gas phase for BPTI.[22] In water, ΔG can be estimated from comparing the stabilities of engineered proteins with hydrated cavities (created by mutations) to the wild-type stabilities.[8,23] To find the ΔG for the binding process, an estimate of the free energy of the mutant protein with its cavity empty must be made, for example by comparison with similar mutations in other proteins. For both the I76A mutant of barnase and the I106A mutant of lysozyme, this analysis gives a ΔG of about −2 kcal/mol, despite the different numbers of hydrogen bonds formed.[8,23] In general, it has been found from mutagenesis experiments that buried waters affect the stability of proteins over a wide range of values.[8]

2. Methods

The free energy of hydration of the interior cavity, ΔG_{hyd}, is given by the sum of two terms: (1) the removal of a water molecule from the pure fluid (ΔG_{wat}) and (2) the addition of a water molecule to the protein cavity ($\Delta G_{protein}$):[13]

$$(N+1)H_2O_{liq} + \text{protein}\{\text{empty cavity}\} \xrightarrow{\Delta G_{hyd}} NH_2O_{liq} + \text{protein}\{H_2O\} \quad (1)$$

$$(N+1)H_2O_{liq} \xrightarrow{\Delta G_{wat}} NH_2O + H_2O_{ni} \quad (2)$$

$$H_2O_{ni} + \text{protein}\{\text{empty cavity}\} \xrightarrow{\Delta G_{protein}} \text{protein}\{H_2O\} \quad (3)$$

where H_2O_{liq} is liquid water and H_2O_{ni} is a noninteracting water molecule. To correctly and reversibly calculate the free energy for eq 3, the simulation would have to allow for the water molecule to search the entire simulation box, which would require prohibitively long simulation times. Also, if $\Delta G_{protein}$ is positive, the water molecule will not occupy the cavity of interest. These problems can be avoided by first localizing the molecule in the binding site and then allowing the molecule to interact with the protein.[16,17,48−50] The process represented by eq 3 becomes a two-step process

$$H_2O_{ni} + \text{protein}\{\text{empty cavity}\} \xrightarrow{\Delta G_{loc}} \text{protein}\{H_2O^{loc}_{ni}\} \quad (4)$$

$$\text{protein}\{H_2O^{loc}_{ni}\} \xrightarrow{\Delta G_{inter}} \text{protein}\{H_2O\} \quad (5)$$

where $H_2O^{loc}_{in}$ indicates a noninteracting molecule which is localized in the protein cavity. If the molecule is localized using a harmonic potential

$$U_{harm}(r) = k_{harm}(r_O - r_X)^2 \quad (6)$$

where k_{harm} is the force constant, r_O is the position of the oxygen atom of the water molecule, and r_X is the center of the binding site; then the

(37) Varadarajan, R.; Richards, F. M. *Biochemistry* **1992**, *31*, 12315−12327.
(38) Dunitz, J. D. *Science* **1994**, *264*, 670.
(39) Li, Z.; Lazardis, T. *J. Am. Chem. Soc.* **2003**, *125*, 6636−6637.
(40) Ladbury, J. E. *Chem. Biol.* **1996**, *3*, 973−980.
(41) Holdgate, G. A.; Tunnicliffe, A.; Ward, W. H. J.; Westin, S. A.; Rosenbrock, G.; Barth, P. T.; Taylor, I. W. F.; Pauptit, R. A.; Timms, D. *Biochemistry* **1997**, *36*, 9663−9673.
(42) Snider, M. J.; Wolfenden, R. *Biochemistry* **2001**, *40*, 11364−11371.
(43) Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849−857.
(44) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1991**, *113*, 5860−5862.
(45) Tidor, B.; Karplus, M. *J. Mol. Biol.* **1994**, *238*, 405−414.
(46) Phelps, D. K.; Rossky, P. J.; Post, C. B. *J. Mol. Biol.* **1998**, *276*, 331−337.
(47) These two studies examined the same protein, but slightly different cavities.
(48) Mann, G.; Hermans, J. *J. Mol. Biol.* **2000**, *302*, 979−989.
(49) Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, *119*, 2707−2714.
(50) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.

free energy to localize the molecule is given by[17,49,50]

$$\Delta G_{\mathrm{loc}} = -kT \ln[\rho(\pi kT/k_{\mathrm{harm}})^{3/2}] \quad (7)$$

where k is Boltzmann's constant, T is temperature, and ρ is the bulk density of water. This method correctly satisfies the conditions regarding a standard state which is the pure liquid of density ρ.[16,17,48−50] Note that some previous studies of cavity hydration[13−15] have not satisfied conditions of the standard state and reversibility.[49,50] Simulations of BPTI were run without restraint to find the mean square fluctuation in the oxygen position, $\langle \delta r^2 \rangle$. The optimal force constant was found from $k_{\mathrm{harm}} = 6kT/\langle \delta r^2 \rangle$ to give $k_{\mathrm{harm}} = 3$ kcal/mol/Å^2.[17]

Because the protein may rotate and translate during the course of the simulation, the center of the restraint potential, r_{o}, must rotate along with the protein. This was implemented by rotating the crystal structure geometry onto the simulation structure every time step by minimizing the root-mean-square deviation between the C_α atoms. The value of r_{X} is taken to be the location of the oxygen atom of the crystal water, after the structure is rotated onto the current coordinates. The free energy for the hydration process is then given by

$$\Delta G_{\mathrm{hyd}} = \Delta G_{\mathrm{wat}} + \Delta G_{\mathrm{protein}} = \Delta G_{\mathrm{wat}} + \Delta G_{\mathrm{loc}} + \Delta G_{\mathrm{inter}} \quad (8)$$

The initial state of eq 1 is the empty cavity. To ensure that other water molecules do not enter the cavity as λ goes to zero, an additional short-ranged interaction between the position of the reference cavity water oxygen and the oxygen atoms on the solvent water is added with this form: $U_{\mathrm{r}} = \sum_k \epsilon_{\mathrm{r}}(r_{\mathrm{Oh}}/\sigma_{\mathrm{r}})^{-12}$.

To calculate the free energy of the process given by eq 5, free energy perturbation can be used, with a potential energy of the water molecule in the binding pocket is given by

$$U_\lambda = \lambda\left[\sum_j 4\epsilon_{\mathrm{Oj}}\left[\left(\frac{r_{\mathrm{Oj}}}{\sigma_{\mathrm{Oj}}}\right)^{-12} - \left(\frac{r_{\mathrm{Oj}}}{\sigma_{\mathrm{Oj}}}\right)^{-6}\right] + \sum_{i=1}^{3}\sum_j q_i q_j / r_{ij}\right] + (1-\lambda)\sum_k \epsilon_{\mathrm{r}}\left(\frac{r_{\mathrm{Ok}}}{\sigma_r}\right)^{-12} + (1-\lambda)k_{\mathrm{harm}}(r_{\mathrm{O}} - r_{\mathrm{X}})^2 \quad (9)$$

where O denotes the position of the oxygen atom of the cavity water, the sum over i is over the atoms of the cavity water, the sum over j is over all the other atoms (all protein, ions, and all other water molecules), and the sum over k is over the oxygen positions of all other water molecules. The parameter λ therefore scales in the Lennard-Jones and electrostatic interactions of the cavity water molecule with the surrounding molecules while simultaneously scaling out both the repulsive term, keeping other water molecules from the site ($\epsilon_{\mathrm{r}} = 0.152$ kcal and $\sigma_{\mathrm{r}} = 2.0$ Å), and the harmonic restraint term. The protein is fully hydrated at $\lambda = 1$.

The calculation of ΔG from free energy perturbation requires the calculation of averages of $\langle \exp[-(U_{\lambda_i} - U_{\lambda_{i+1}})/kT] \rangle_{U_{\mathrm{l}}}$ (see, for example, ref 51). The bias from the repulsive potential U_{r} keeping the cavity empty as λ goes to zero can be corrected using

$$\langle \exp[-(U_{\lambda_i} - U_{\lambda_{i+1}})/kT] \rangle^0_{U_{\mathrm{l}}} = \frac{\langle \exp[-(U_{\lambda_i} - U_{\lambda_{i+1}})/kT] \exp[(1-\lambda_i)U_{\mathrm{r}}/kT] \rangle_{U_{\mathrm{l}}}}{\langle \exp[(1-\lambda_i)U_{\mathrm{r}}/kT] \rangle_{U_{\mathrm{l}}}} \quad (10)$$

(51) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.

where $\langle ... \rangle^0$ denotes the average without U_{r}. If the value of the perturbation term $U_{\lambda_i} - U_{\lambda_{i+1}}$ is uncorrelated with U_{r}, then

$$\langle \exp[-(U_{\lambda_i} - U_{\lambda_{i+1}})/kT] \rangle^0_{U_{\mathrm{l}}} = \frac{\langle \exp[-(U_{\lambda_i} - U_{\lambda_{i+1}})/kT] \rangle_{U_{\mathrm{l}}} \langle \exp[(1-\lambda_i)U_{\mathrm{r}}/kT] \rangle_{U_{\mathrm{l}}}}{\langle \exp[(1-\lambda_i)U_{\mathrm{r}}/kT] \rangle_{U_{\mathrm{l}}}} = \langle \exp[-(U_\lambda - U_{\lambda_{i+1}})/kT] \rangle_{U_{\mathrm{l}}} \quad (11)$$

This is the assumption we use. For the barnase cavity, water molecules never attempt to enter so U_{r} is never much greater than zero. For BPTI, this term is required, but is effectively zero except at small λ values. At larger values of λ, the interactions with the cavity water are sufficient to keep other molecules out of the cavity.

The entropy can be found from a finite difference approximation of the temperature derivative which requires calculating the free energy at two different temperatures ($T \pm \Delta T$),[52] given by

$$\Delta S = -(\Delta G(T + \Delta T) - \Delta G(T - \Delta T))/(2\Delta T) \quad (12)$$

The entropies are about an order of magnitude more uncertain than the ΔG and so require longer simulations.[52] Using a finite difference approximation to the entropy is equivalent to assuming that the free energy is linear over this temperature range. Therefore, rather than calculating ΔS through eq 12, the free energy at the three temperatures ($T - \Delta T$, T, $T + \Delta$T) can be fit to a line and the slope of the line can be used to get ΔS. A temperature difference of 15 K is used, which in previous studies of aqueous solvation has been shown to be effective.[52−54] A similar method has been used to estimate the entropy of binding between nucleic acids.[55]

All protein molecular dynamics were performed using the Amber 6 suite of programs with the Cornell et al. 1994 force field[56] and TIP3P water.[57] The protein contributions to the free energies, $\Delta G_{\mathrm{proteins}}$, were determined with free energy perturbation in the module Gibbs from the Amber 6.0 suite of programs.[58] A minimum of 500 ps of simulation was performed at each of 12 λ values ranging from $\lambda = 0.95$ to $\lambda = 0.025$ with a prior 20 ps of equilibration performed at each λ value. In the wild-type BPTI (PDB entry 5PTI) all λ values were equilibrated for 220 ps. The barnase mutant (PDB entry 1BRI) 298 K simulations included an additional 500 ps of molecular dynamics on all values of λ from a different equilibration point, as well as an addition 2000 ps at the end points in an attempt to minimize error in the free energy and other structural data. Only forward (insertion) values of ΔG were used in the protein simulations with the exception of 0.025 → 0.0, where the backward (deletion) value was used due to increased noise at the end points. Insertion and deletion free energies were not averaged as the magnitudes of errors in the two measurements are not identical.[59] The insertion free energy values were less noisy than their deletion counterparts, but values for ΔG_{hyd} using both forward and reverse as well as half steps were very similar.

The free energy for the removal of a water molecule from the pure liquid, ΔG_{wat}, was calculated using our group's own program, using the separated−shifted scaling method to avoid singularities.[60,61] The

(52) Smith, D. E.; Haymet, A. D. J. *J. Chem. Phys.* **1993**, *98*, 6445−6454.
(53) Rick, S. W. *J. Phys. Chem. B* **2000**, *104*, 6884−6888.
(54) Lüdemann, S.; Schreiber, H.; Abseher, R.; Steinhauser, O. *J. Chem. Phys.* **1996**, *104*, 286.
(55) Norberg, J.; Nilsson, L. *J. Phys. Chem.* **1995**, *99*, 13056−13058.
(56) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(57) Jorgensen, W. L.; Madura, J. D. *Mol. Phys.* **1985**, *56*, 1381.
(58) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. I.; Vincent, J. J.; Crowlet, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *Amber 6*; University of California: San Francisco, 1999.
(59) Lu, N.; Kofke, D. A. *J. Chem. Phys.* **2001**, *114*, 7303−7311.
(60) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025−9031.

Table 1. Protein ΔG Values at 283, 298, and 313 K[a]

	$\Delta G_{protein}$ (kcal/mol)			ΔG_{hyd} (kcal/mol)	
T (K)	BPTI	barnase	ΔG_{wat} (kcal/mol)	BPTI	barnase
283	−11.0(2)	−1.2(2)	6.36(6)	−4.6(2)	5.1(2)
298	−10.9(1)	−1.5(1)	6.18(4)	−4.7(1)	4.7(1)
313	−10.3(2)	−1.3(2)	5.99(5)	−4.3(2)	4.7(2)

[a] Numbers in parentheses represent 95% confidence limits.

free energy calculations used 12 λ values, ranging from 0.05 to 0.95, and ran for 5 ns at each λ value. These simulations used 256 molecules, a 1 fs time step, and SHAKE for bond constraints.[51] The simulations were done in the isothermal−isobaric (constant T, P, N) ensemble, by coupling to a pressure bath and a Nosé−Hoover temperature bath.[62−64]

The I76A mutant of barnase and BPTI structures were obtained and the original counterions were removed. Three proteins are in the unit cell of the 1BRI structure; only the third (labeled C) was used. The histidine residues were assumed to be neutral and changed to reflect protonation of the N_δ atom (HID). In 5PTI all the original crystal waters were kept, while for 1BRI only those waters surrounding the protein chain C were retained. The proteins were again loaded into xLeap, and cross-linking of the appropriate cysteine residues was performed for 5PTI using the connect command. Chlorine ions were added to create a neutral simulation box; two ions were added to 1BRI and six were added to 5PTI. Additional solvent waters were added to create an 8 Å box around the two proteins: 3132 and 4769 water molecules were added to 5PTI and 1BRI, respectively. Simulations were performed with particle-mesh Ewald. A 7 Å cutoff was used for nonbonded pairs during simulations for Lennard-Jones and real space Ewald interactions.

Steepest descent minimization was performed for 10 ps, and the proteins were warmed to 298 K over the course of 26 ps for 5PTI and 140 ps for 1BRI using the Sander program in the canonical ensemble. After equilibrating at 298 K for 20 ps, an additional 20 ps of isothermal−isobaric equilibration was performed at 283, 298, and 313 K. Temperature coupling during the P, N, and T simulations was performed using Berendsen coupling[65] with the temperature relaxation time τ_T optimized for each structure in an effort to maintain good temperature control for the cavity water. The resulting couplings were $\tau_T = 0.5$ ps for 1BRI and $\tau_T = 0.2$ ps for 5PTI. The cavity water was considered part of the solute for all calculations. The constant pressure calculations were performed with isotropic position scaling, and SHAKE was used to constrain all bonds.

3. Results

Thermodynamics of Cavity Hydration. The free energies for hydration of the two proteins are shown in Table 1. The free energy for adding a water to the empty cavity (see eq 5) is negative for both proteins. The hydration process is a competition between $\Delta G_{protein}$ and the free energy of removing a water molecule to the pure liquid, ΔG_{wat}. For BPTI our calculations find that $\Delta G_{protein}$ is greater in magnitude than ΔG_{wat} and for this cavity ΔG_{hyd} is negative. For barnase, $\Delta G_{protein}$ is not large enough in magnitude to outweigh ΔG_{wat} and the resultant ΔG_{hyd} is positive. Our calculated ΔG_{hyd} for barnase is much larger than the −2.0 kcal/mol estimate based on the stability of the barnase mutant, taken to have a filled cavity, relative to other Ile → Ala mutants with empty cavities.[23] Our calculations then indicate that the polar cavity in BPTI is hydrated and the nonpolar cavity in barnase is not hydrated. This conclusion about the occupancy of polar and nonpolar cavities agrees with the calculations of Zhang and Hermans for other proteins but not with X-ray data for barnase.

Table 2. Protein Entropy and Enthalpy Values[a]

	protein			hydration	
	BPTI	barnase	water	BPTI	barnase
ΔG (kcal/mol)	−10.7(1)	−1.4(1)	6.18(4)	−4.5(1)	4.8(1)
ΔS (cal/(mol K))	−24(9)	4(9)	12(3)	−12(9)	16(9)
$-T\Delta S$ (kcal/mol)	7(3)	−1(3)	−3.7(8)	3(3)	−5(3)
ΔH (kcal/mol)	−19(3)	0(3)	9.85(8)	−9(3)	10(3)

[a] Numbers in parentheses represent 95% confidence limits.

From the temperature dependences of the free energies, the entropies can be found from a linear fit as described above (Table 2). The error estimates for the entropies are about an order of magnitude larger than the error bars for the free energies, as is typical.[52] The entropic contribution to ΔG_{wat} is unfavorable and is in good agreement with the experimental value of −12.24 cal/(mol K).[66] The entropic contribution to $\Delta G_{protein}$ is also unfavorable and by a larger amount than for the pure liquid. Therefore, the hydration process for the BPTI cavity is entropically unfavorable by -12 ± 9 cal/(mol K). From the gas phase mass spectroscopic data on BPTI, ΔS for the transfer of a water molecule from the vapor to the protein at 0 °C is -62 ± 5 cal/(mol K),[22] giving an entropy change of -33 ± 5 cal/(mol K) for the transfer from the liquid phase to gas phase BPTI.[67,68] This value is much larger than our calculated value, perhaps indicating a difference in vibrational modes between gas phase and aqueous phase proteins. The entropy of hydration for the barnase cavity is positive ($\Delta S_{hyd} = 16 \pm 9$ cal/(mol K)), which means that even though the process has a positive free energy, it is entropically favorable.

The entropy change due to turning on the protein−water interactions (see eq 5) is given by $\Delta S_{inter} = \Delta S_{protein} - \Delta S_{loc}$. From the temperature derivative of eq 7, ΔS_{loc} equals −5.4 cal/(mol K), giving $\Delta S_{inter} = -19 \pm 9$ cal/(mol K). The study of Fischer et al. estimated the ΔS from the vibrational normal modes with solvent modeled using a distance-dependent dielectric constant. This gives a ΔS of −13.4 cal/(mol K) for the vibrational entropy difference between hydrated and empty BPTI. If the rotational entropy of gas phase water is subtracted ($\Delta S_{rot} = 10.6$ cal/(mol K)), then a value of ΔS equal to −24 cal/(mol K) is found.[10] This entropy change is for a process comparable to eq 5: a translationally restrained but rotationally free water plus an unhydrated BPTI going to hydrated BPTI. The Fischer et al. estimate of −24 cal/(mol K) contains a part due to the ordering of the water molecule in the protein environment and also a contribution from a change in the low-frequency vibrational modes, indicating a more flexible protein.[10] That value is close to our value of -19 ± 9 cal/(mol K).

The enthalpy changes are found using $\Delta H = \Delta G + T\Delta S$. For BPTI, $\Delta H_{protein}$ is -19 ± 3 kcal/mol, which is about the enthalpy change expected upon forming four hydrogen bonds. The measured value for gas phase BPTI is -21.3 ± 1.0 kcal/mol [22] and a calculated value for gas phase BPTI is −19.8 kcal/mol,[9] indicating that $\Delta H_{protein}$, unlike $\Delta S_{protein}$, is similar for the

(61) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1994**, *116*, 3949.
(62) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384.
(63) Nosé, S. *Mol. Phys.* **1984**, *52*, 255−268.
(64) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(65) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(66) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016−2027.
(67) Entropy of vaporization calculated from the method of Kuntz and Kauzmann is −29.4 cal/(mol K) at 0 °C.
(68) Kuntz, I. D., Jr.; Kauzmann, W. *Adv. Protein Chem.* **1974**, *28*, 239−345.

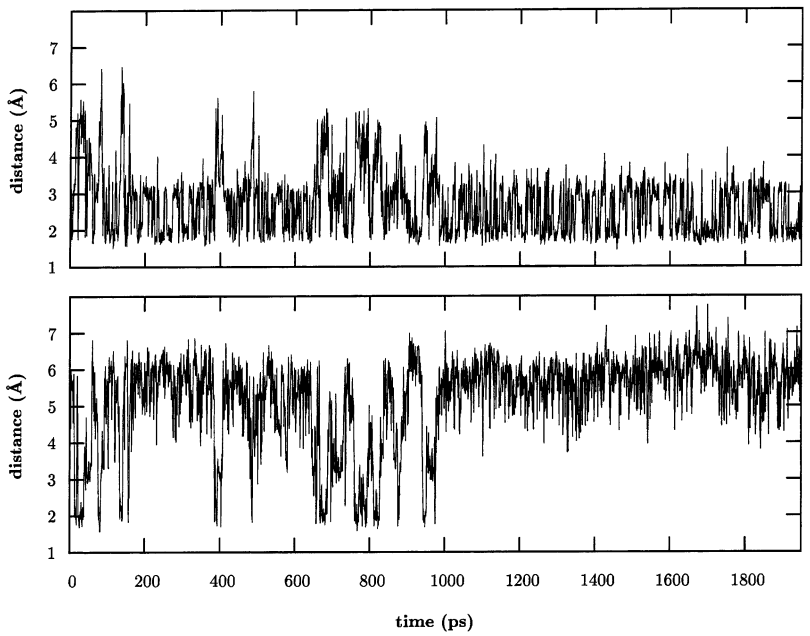

Figure 2. Hydrogen bond distance between the H of the cavity water (WTP) and the oxygens of (a) Phe 7 and (b) Tyr 97.

gas and liquid phases. The binding site of the water in these experiments is not known, so comparisons with our calculated values should be considered as speculative. The calculated enthalpy change for removing a water molecule from the pure liquid is 9.85 ± 0.08 kcal/mol, which is close to the experimental value of 9.974 kcal/mol.[66] Since the removal of a water molecule eliminates about two hydrogen bonds, this value of ΔH_{water} is consistent with an enthalpic cost of 5 kcal/mol per hydrogen bond. The overall enthalpy change for hydrating the cavity in BPTI, ΔH_{hyd}, is exothermic. For the barnase cavity, $\Delta H_{\text{protein}}$ is about zero, less than would be expected upon the formation of one hydrogen bond. The small $\Delta H_{\text{protein}}$ may indicate the protein−water hydrogen bond that is formed is weak or may indicate that the binding of the water weakens protein−protein interactions. The small value of $\Delta H_{\text{protein}}$ gives a ΔH_{hyd} which is endothermic.

Dynamics of the Interior Water. As can be seen in Figure 1a, the BPTI interior water makes four hydrogen bonds to the protein. Of these four atoms (the carbonyl oxygens on residues Thr 11 and Cys 38 and the amide hydrogens on residues Cys 14 and Cys 38), only one, the carbonyl oxygen on Thr 11, makes a hydrogen bond to another atom in the protein. Our simulations reveal that the four hydrogen bond partners of the interior water do not form any new hydrogen bonds as the water molecule is removed. Therefore, the addition of the water to the BPTI cavity creates four new hydrogen bonds. In the pure liquid a water has (about) four hydrogen bonds, but when the water is removed, the solvent can rearrange and remake two hydrogen bonds, so only two are lost. The creation of four new hydrogen bonds in the BPTI cavity, as opposed to two lost in the liquid, and the fact that these hydrogen bonds are with atoms that are, for the most part, not involved in other hydrogen bonds, explain the large enthalpy change, ΔH_{hyd}. Analysis of the 2500 ps trajectory at $\lambda = 1$ indicates that the water protons hydrogen bonded to the O on Thr 11 and the O on Cys 38 do not flip during the time scale of our simulations. This is agreement with the literature which indicates the time scale for hydrogen exchange in BPTI will be on the order of 45 ns.[69] Both of the water protons are within 2.8 Å of their respective hydrogen bond donors for more than 99.5% of the simulation length. The water O−Cys 38 amide H distance is within 2.8 Å 98.2% of the time. The water O−Cys 14 amide H distance is within 2.8 Å only 82.5% of the time, indicating that this hydrogen bond is the least stable of the four hydrogen bonds.

In the 1BRI structure of barnase,[23] the interior water molecule makes only one hydrogen bond to the protein, to the Phe 7 O atom (see Figure 1b). This atom also makes a hydrogen bond to another protein atom (the amide H on Ala 11). The water in this cavity is very dynamic. The water rotates to allow both protons to share the single available site of hydrogen bonding. The binding at the O on Phe 7 is split between the two water proteins with an average residency time of 4.6 ps (see Figure 2). The water molecule is also observed to move within the protein and form a hydrogen bond to the carbonyl oxygen atom on Tyr 97. This atom is far enough from Phe 7 that the water molecule cannot simultaneously form hydrogen bonds to both oxygen atoms and the water moves back and forth between the two sites. The water molecule occupies the site by the Tyr 97 13.1% of the simulation time at $\lambda = 0.95$. In the secondary binding site, as in the first binding site, only one hydrogen bond is formed. The protons flip in this site about once every 14 ps (Figure 2). As in BPTI there are no hydrogen bonds present in the empty cavity that were not found in the hydrated protein.

Protein Flexibility and the Interior Water. The mean square fluctuation in the atomic positions, $\langle \Delta r_i^2 \rangle$, provides a measure of the mobility of the protein atoms. This is calculated from

$$\langle \Delta r_i^2 \rangle = \frac{1}{N} \sum_{j=1}^{N} (\mathbf{r}_i(t_j) - \bar{\mathbf{r}}_i)^2 \qquad (13)$$

(69) Fischer, S.; Verma, C. S.; Hubbard, R. E. *J. Phys. Chem. B* **1998**, *102*, 1797−1805.

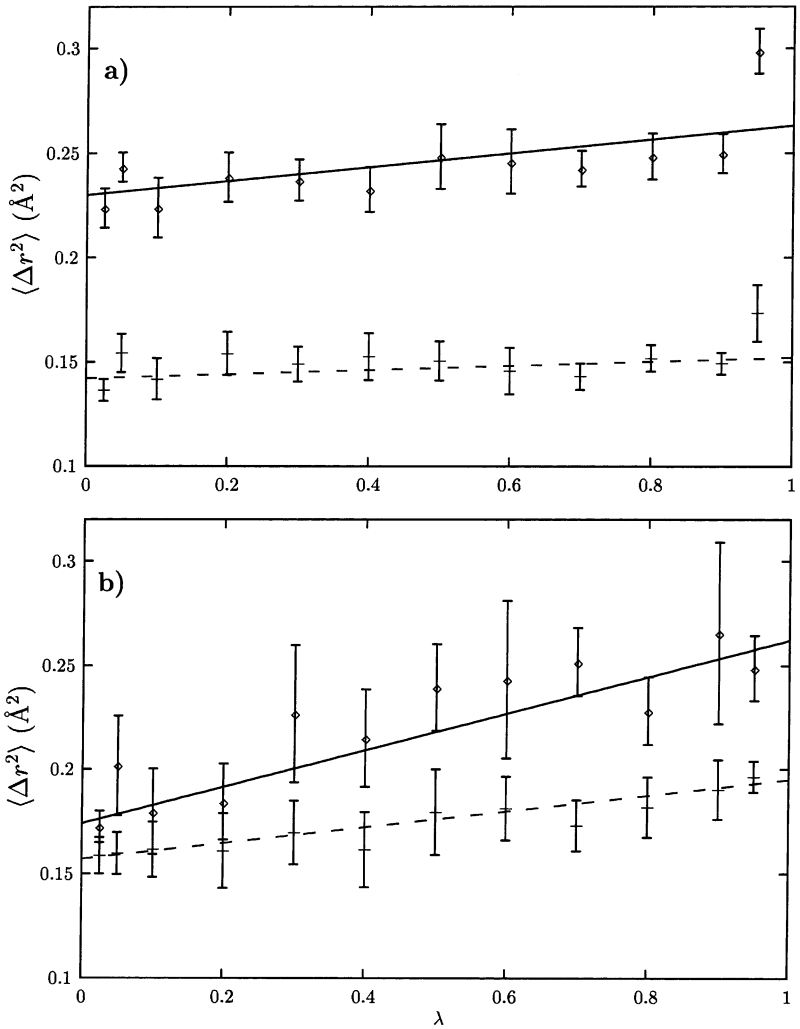

Figure 3. Mean square deviation of full protein backbone (◇) and protein backbone atoms within 6 Å of the cavity water (+) from the reference structure of (a) 1BRI and (b) 5PTI. Errors reported are 95% confidence limits.

where $\mathbf{r}_i(t_j)$ is the position of atom i at time t_j and $\bar{\mathbf{r}}_i$ is the average position of atom i over the N stored time values. To eliminate motion due to translations and rotations of the protein, each configuration is rotated and translated onto a reference structure (the original crystal structure) prior to calculating $\langle \Delta r_i^2 \rangle$ from eq 13. In general, the backbone heavy atoms show the smallest $\langle \Delta r_i^2 \rangle$ and these are the values which can be calculated with the most precision. For this reason, we choose to focus on these atoms to examine the differences in $\langle \Delta r_i^2 \rangle$ with and without the interior water molecule. For this comparison, we report both the value averaged over all heavy backbone atoms and the value averaged over all heavy backbone atoms with 6 Å of the crystal water position. The $\langle \Delta r^2 \rangle$ as a function of the parameter λ ($\lambda = 0$ is the empty and $\lambda = 1$ is the hydrated state) shows that its value increases as the water−protein interactions are scaled in (Figure 3). These results suggest that the protein gets more flexible when interior cavities are occupied. The change in $\langle \Delta r^2 \rangle$ is larger for BPTI than for barnase, perhaps because the water interacts more strongly with BPTI. For both proteins, $\langle \Delta r^2 \rangle$ is smaller for the atoms close to the water, indicating that this is a more rigid than average region of the proteins, and also the mean square deviation of the heavy backbone atoms close to the water changes *less* than the average of all atoms. Thus, even though they are closer to the water, the flexibility of those atoms is affected less than average by the addition of the water, perhaps, again, because these regions of the two proteins are more rigid. For BPTI, $\langle \Delta r^2 \rangle$ for all backbone atoms increases by 0.088 Å^2 and for the close atoms it increases by 0.038 Å^2 with hydration. For barnase, the changes were smaller in magnitude with increases of 0.033 and 0.10 Å^2 for the entire protein and close residues, respectively. One atom which decreases its value of mean square fluctuations when the water is present is Ile 88 C$^\delta$ on barnase. This atom borders the cavity, and the rotation of Ile 88 from its position in the wild-type structure is what creates the cavity.[23] While other atoms show an increase in $\langle \Delta r_i^2 \rangle$, for this atom $\langle \Delta r_i^2 \rangle$ decreases from 0.43 ± 0.03 Å^2 to 0.23 ± 0.02 Å^2. The increased mobility of this atom in the absence of water may be due to the space created by the loss of the water. The atoms which border the cavity for BPTI, which are more polar, and the polar atom bordering the barnase cavity (Phe 7 O) all appear to get more flexible as the water is added.

The increase in protein flexibility seen on hydration may be understood at least in part by looking at the hydrogen bonds formed near the cavity. Of the four protein atoms which form hydrogen bonds to the water in the BPTI cavity, only one of these atoms (Thr 11 O) forms a second hydrogen bond with another protein atom (see Figure 1a). As the water molecule is

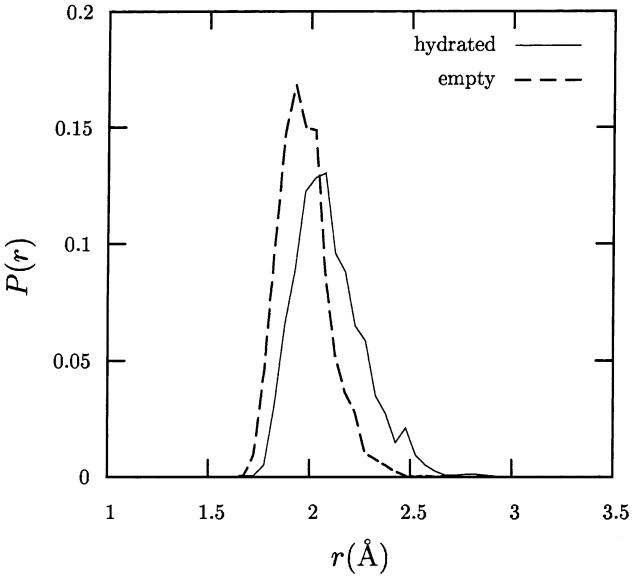

Figure 4. Distribution of hydrogen bond lengths between the H on Gly 36 and the O on Thr 11 in both the empty and hydrated simulations of 5PTI.

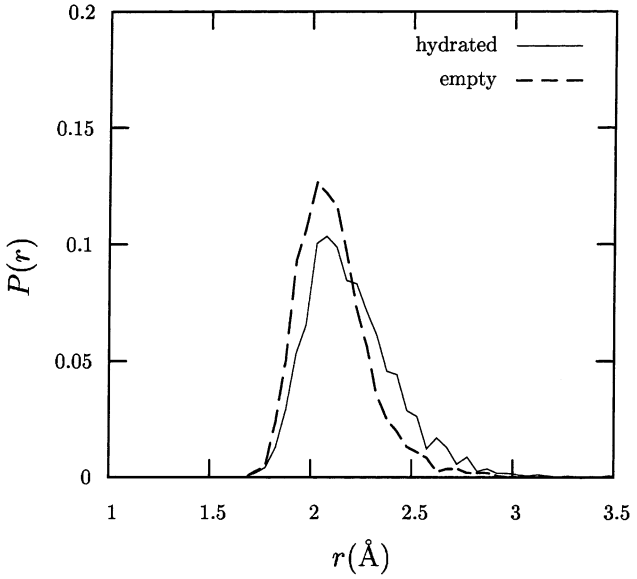

Figure 5. Distribution of hydrogen bond lengths between the H on Ala 11 and the O on Phe 7 in both the empty and hydrated simulations of 1BRI.

removed from the cavity, the bond between the O on Thr 11 and the H on Gly 36 becomes slightly tighter with the average length, decreasing to 1.92 Å from 2.05 Å (Figure 4). For the barnase cavity, the Phe 7 O has one hydrogen bond partner other than the water (Figure 1b). This hydrogen bond, with Ala 11 H, also is tighter for the empty cavity with a length of 2.06 Å than for the hydrated state with a length of 2.18 Å (Figure 5). In both cases, it appears that the hydrogen bonds between protein atoms at the surface of the cavity get longer due to the presence of the buried water. The protein−protein hydrogen bond is stretched in the process of forming the water−protein hydrogen bond. This may explain why the protein gets more flexible as the cavity gets hydrated. It may also be part of the reason that $\Delta H_{\text{protein}}$ is about zero for barnase, despite forming a new hydrogen bond. The new hydrogen bond that is formed must weaken the protein−protein interactions, including the Phe 7 O−Ala 11 H hydrogen bond.

Changes in Cavity Volumes. The addition of a water molecule may also change the size of the cavity. Cavity volumes can be found by rolling a probe sphere around the van der Waals surface as implemented in the Molecular Surface Package[70] and VOIDOO[71] programs. However, we found that our results were very sensitive to the size of the probe sphere radius and consistent results using the same probe radius for a sequence of structures for both proteins could not be achieved. A simple method for estimating the volume is to define the cavity as an irregular polyhedron with vertices defined by atoms on the edges. The volume of the tetrahedron is found from

$$V = \begin{vmatrix} x_4 & y_4 & z_4 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_1 & y_1 & z_1 & 1 \end{vmatrix} (6)^{-1} \qquad (14)$$

where x_i, y_i, and z_i are the coordinates of the four atoms.[72] For barnase, the polyhedron is defined by six points, making two tetrahedra which share the same base. The base was composed of $C_{\gamma 1}$ Ile 88, $C_{\gamma 2}$ Ile 96, and $C_{\delta 1}$ Ile 109. The first tetrahedron used O Phe 7 as its vertex, and the second used O Tyr 97. Using this definition, the original three protein chains from the crystal structure were analyzed. The volume of the cavity as defined in protein C is 23.5 Å^3 and falls between the volumes for A and B, 22.6 and 24.1 Å^3, respectively. The similar cavity volumes agrees with the similarity of the three structures as measured by the root-mean-square-deviation (rmsd) overlap of the residues within 6 Å of the interior water. The atoms in this region on protein C differed from those on A by 0.17 Å and those on B by 0.13 Å. Determination of cavity volumes for a snapshot taken every picosecond for the 1700 ($\lambda = 0$) and 2000 ps ($\lambda = 1$) simulations was performed using the simple polyhedron method. The differences between the two cavities are small with a total average volume of 28.70 (3.38) Å^3 for the hydrated cavity versus 26.40 (3.36) Å^3 for the empty cavity. The numbers in parentheses indicate standard deviations. The difference between the volumes of the hydrated and empty cavities is smaller than the fluctuations of the cavity. The volumes are about the same as those calculated from the crystal structures. For the tetrahedron defined by the three base atoms and O Phe 7, which is the site the water molecule occupies the majority of the time, the volume difference is 12.95 (2.25) Å^3 for the hydrated protein versus 12.02 (2.09) Å^3 for the nonhydrated protein. This is more than half the volume of the polyhedron made up of the two tetrahedra, indicating that this tetrahedron is the larger of the two. The empty cavity is slightly tighter in both volume and its fluctuations, as given by the standard deviation.

A similar treatment of the BPTI cavity can be made using vertices that are the sites involved in hydrogen bonding: O Thr 11, O Cys 38, N Cys 14, and N Cys 38. The volumes were again calculated for snapshots taken over a trajectory and averaged. The hydrated cavity had a volume of 8.78 (0.69) Å^3, while the empty cavity had a slightly larger average volume of

(70) Connolly, M. L. *Science* **1983**, *221*, 709−713.
(71) Kleywegt, G. J.; Jones, T. A. *Acta Crystallogr., Sect. D* **1994**, *50*, 178−185.
(72) Kopcsak, P. J. *Matrix Analysis and Analytical Geometry: a new approach*; Chemistry Publishing Company, Inc.: New York, 1968.

9.27 (0.80) Å^3. The hydrated cavity does appear to be slightly smaller, perhaps due to the attractive forces between the water and the protein. The cavity size in the crystal structure is 8.92 Å^3. For comparison, the volume of the tetrahedron made up of nearest neighbors in an ice crystal is 10.8 Å^3 (using an oxygen−oxygen nearest neighbor distance of 2.760 Å).[73] Comparisons between the cavities are complicated by the fact that different atom types are used to define them (either carbon, nitrogen, or oxygen atoms) and the atoms will have their own excluded volumes, leaving a different amount of free volume for the water molecule. However, it appears that a water molecule in the BPTI has less space than a molecule in ice.

4. Discussion

The calculated free energy changes for the hydration of the cavities in the two different proteins vary considerably depending on the type of cavity. The entropy change also varies considerably. For the BPTI cavity, the hydration process is entropically unfavorable, and for barnase it is entropically favorable. The polar cavity in BPTI, in which the water can form four hydrogen bonds, would be predicted to be hydrated, based on the calculated free energy change of −4.7 kcal/mol (Table 1). The less polar cavity of the barnase mutant is not predicted to be hydrated. This result, that polar cavities are hydrated and nonpolar are not, is consistent with calculations of cavities on other proteins,[16] but our ΔG_{hyd} for the barnase mutant is in apparent disagreement with the X-ray results of Buckle et al.[23]

Before discussing the differences between the simulations and experiments, it is worth stating that there is some ambiguity in the X-ray data as well. The unit cell of the protein contains three protein molecules, and only one (molecule C) contains electron density well-defined enough to indicate a water molecule in that position. The electron density in the other two structures is too weak to assign a water molecule to this site. The three protein structures in the unit cell are very similar (with root-mean-square deviations of 0.169 for protein A and 0.128 for protein B from C), so it is not obvious why only one of the structures contains the buried water. The simulation and X-ray (for protein structure C) results may be different for several reasons. The experiments are done under different conditions, including not only in the crystal environment but also at a temperature of 4 °C and at a pH of 7.5.[23] The temperature difference does not seem significant enough to change results, and since the calculations find that the process is entropically favorable, ΔG_{hyd} should be even less at 25 °C than at 4 °C. Another explanation may be that the interior water in the X-ray structure is thermodynamically unstable, in agreement with the calculated results, but is kinetically trapped. This does not seem too likely, since the cavity is not too far from the surface and our simulations reveal that the water molecule can leave the cavity (when unconstrained) on a short time scale. The differences may also be due to problems with the potential models. For the Cornell et al. force field, the parameters which describe the interactions between nonpolar groups and water are chosen to reproduce the solvation free energies for a set of molecules, including methane, ethane, and butane.[56] These solvation energies depend on not only the strength of the water−solute interaction but also the water−water potential. For example, several models with much different methane−water interactions have similar solvation free energies.[74] The water−methane interaction for Cornell et al. TIP3P models has a potential energy minimum equal to −0.28 kcal/mol, much less than the ab initio value of −0.71 kcal/mol.[75] Other nonpolarizable models have similarly weak methane−water interactions, while a polarizable model has a deeper minimum (−0.57 kcal/mol), while still giving the same solvation free energy.[74] This may indicate a possible area of improvement for potential models. Stronger interactions between water and nonpolar groups would tend to decrease the value of ΔG_{hyd}, bringing it in closer agreement with the apparent X-ray result. It appears that a closer look at both the X-ray data and the potential models is necessary before the hydration of this cavity and perhaps other hydrophobic cavities is fully understood.

The dynamics of the water in the cavity depends on the type of cavity. Mobility in the nonpolar barnase cavity is high with the water molecule undergoing two types of motion not seen in the polar BPTI cavity. The molecule rotates to interchange which of its hydrogen atoms is near the Phe 7 O atom, on a 5 ps time scale, and also translates, about once every 14 ps, to a secondary binding site 7.3 Å away to form a single hydrogen bond with the Try 97 O atom (Figure 2). For the water molecule in BPTI, no such rotations or translations to other regions of the protein are seen.

For both proteins, the addition of the water molecule to the interior cavity increases the flexibility of the protein, as seen in an increase in the mean square fluctuations in atomic positions, $\langle \Delta r^2 \rangle$ (Figure 3). The interior water molecules may increase the protein flexibility by increasing and weakening the length of nearby protein−protein hydrogen bonds (Figures 5 and 4). Other studies have indicated that buried water molecules shield charge−charge interactions of the protein leading to a higher dielectric constant; this would lead to an increased protein flexibility.[33] Two studies of BPTI in the gas phase have examined how protein flexibility changes, based on calculations of the change in vibrational entropy upon binding the water molecule. These calculations reached opposite conclusions, as stated in the Introduction, with one concluding that the flexibility increases[9] and the other that it decreases.[11] In the simulations of Mao et al., the water molecule did not stay in the cavity, so this may explain the difference.[11] Our results for aqueous BPTI agree with the results of Fischer and Verma[9] and also a later study by Fischer, Smith, and Verma.[10] The binding of molecules larger than water may increase flexibility as well. A study by Tidor and Karplus of the dimerization of the protein insulin, in the gas phase, demonstrated that the change in vibrational entropy suggests that the monomer protein gets more flexible upon binding to form the dimer. Another study showed that the binding of an inhibitor to the rhinovirus capsid protein again increases the flexibility of the protein, as seen by an increase in $\langle \Delta r^2 \rangle$.[46] This capsid protein study is different in that the binding process involved the displacement of water, whereas the other studies involved the displacement of empty space, so this study is not looking at quite the same thing.

(73) Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Oxford University Press: New York and Oxford, 1969.
(74) Rick, S. W.; Berne, B. J. *J. Phys. Chem. B* **1997**, *101*, 10488.
(75) Szczęśniak, M. M.; Chałasiński, G.; Cybulski, S. M.; Cieplak, P. *J. Chem. Phys.* **1993**, *98*, 3078.

5. Conclusions

Our calculations found that the thermodynamics for the hydration of a protein cavity depend significantly on the properties of the cavity. The entropy of hydration, ΔS_{hyd}, varies considerably for the two proteins, which, given the large differences in the mobility of the buried water molecules, is perhaps not surprising. The ΔS_{hyd} for the barnase cavity is positive, indicating that the process is entropically favorable. For the BPTI cavity ΔS_{hyd} is negative. Different entropies for different hydration sites on proteins are also indicated by the mass spectroscopic data on gas phase BPTI[22] and by the librational amplitudes of water in BPTI,[28] although all the reported ΔS_{hyd} values are negative. Those hydration sites are likely to be more hydrophilic than the barnase cavity. The entropy estimates based on the librational amplitudes find that the entropy changes are closer to zero for the water molecules that form fewer than four hydrogen bonds.[28] Our results show that ΔS_{hyd} can be negative for water molecules which form only one hydrogen bond. For the BPTI cavity, ΔS_{hyd} is -12 ± 9 cal/(mol K). The gas phase mass spectroscopic data on BPTI gives an entropy change of -33 ± 5 cal/(mol K) for the transfer from the liquid phase to gas phase BPTI, although the binding site of the water is uncertain.[22] Our calculated value is significantly larger than this but is smaller than the Dunitz lower bound estimate of -7 cal/(mol K).[38] That value is certainly within the error bars of our calculations, but our result, together with the gas phase measurements for BPTI and the theoretical estimate for HIV-1 protease,[39] indicates that entropy changes may in fact be less than the "lower bound" value of Dunitz. The lower entropy implies that the addition of the buried water has a significant influence on the protein. The magnitude of the influence may be larger in the gas phase than in the liquid phase. Our results, particularly the ΔS_{hyd} values, agree with the conclusions of Takano et al. that "all water molecules do not contribute equally to stability, owing to differences in the environment of water molecules in proteins, such as the number of hydrogen bonds." [8]

Acknowledgment. Support from the National Science Foundation under Contract No. CHE-0213488 is gratefully acknowledged.

JA049701C