J|A|C|S
ARTICLES
Published on Web 05/18/2007

Metastable Water Clusters in the Nonpolar Cavities of the Thermostable Protein Tetrabrachion

Hao Yin,[†] Gerhard Hummer,*[,‡] and Jayendran C. Rasaiah*[,†]

Contribution from the Department of Chemistry, University of Maine, Orono, Maine 04469-5706, and Laboratory for Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received January 26, 2007; E-mail: Gerhard.Hummer@NIH.gov; rasaiah@maine.edu

Abstract: Water expulsion from the protein core is a key step in protein folding. Nevertheless, unusually large water clusters confined into the nonpolar cavities have been observed in the X-ray crystal structures of tetrabrachion, a bacterial protein that is thermostable up to at least 403 K (130 °C). Here, we use molecular dynamics (MD) simulations to investigate the structure and thermodynamics of water filling the largest cavity of the right-handed coiled-coil stalk of tetrabrachion at 365 K (92 °C), the temperature of optimal bacterial growth, and at room temperature (298 K). Hydrogen-bonded water clusters of seven to nine water molecules are found to be thermodynamically stable in this cavity at both temperatures, confirming the X-ray studies. Stability, as measured by the transfer free energy of the optimal size cluster, decreases with increasing temperature. Water filling is thus driven by the energy of transfer and opposed by the transfer entropy, both depending only weakly on temperature. Our calculations suggest that cluster formation becomes unfavorable at ~384 K (110 °C), signaling the onset of drying just slightly above the temperature of optimal growth. "Drying" thus precedes protein denaturation. At room temperature, the second largest cavity in tetrabrachion accommodates a five water molecule cluster, as reported in the X-ray studies. However, the simulations show that at 365 K the cluster is unstable and breaks up. We suggest that the large hydrophobic cavities may act as binding sites for two proteases, possibly explaining the unusual thermostability of the resulting protease-stalk complexes (up to ~393 K, 120 °C).

1. Introduction

Interior water plays an essential role in modulating the structure, stability, function, and dynamics of biomolecules.[1−5] Water may occur as an integral part of the protein to which it is usually hydrogen bonded, or it may appear in a weakly polar or even hydrophobic cavity[5−7] or channel.[8] The presence of water in nonpolar hydrophobic cavities has been reported on the basis of X-ray crystallography[6] and nuclear magnetic resonance (NMR) studies,[7,9−11] and NMR has also been used to probe the dynamics of water in proteins.[9−11] Interior water in nonpolar cavities is either amorphous[7] or present as a dimer or a cluster of several water molecules stabilized by hydrogen bonding within the cluster and by van der Waals interactions and hydrogen bonding with the peptide walls of a protein cavity.[12−18] Water in these predominantly nonpolar cavities can be metastable, with filling and emptying occurring upon changes in solvent conditions[19] or the local polarity of the cavity.[20,21] Here, we examine the thermodynamic and structural properties of such unusual and interesting forms of water occupying the large cavities of the bacterial protein tetrabrachion.[14−16]

Tetrabrachion is a surface layer protein of the hyperthermophilic archaebacterium *Staphylothermus marinus* found in geothermal marine environments. It consists of a 70-nm stalk that branches into four arms of 24 nm length.[15,16] The stalk

[†] University of Maine.
[‡] National Institutes of Health.

(1) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *36*, 389−415.
(2) Helms, V. *Chem. Phys. Phys. Chem.* **2006**, *8*, 23−33.
(3) Takanao, K.; Yamagata, Y.; Yutani, K. *Protein Eng.* **2003**, *16*, 5−9.
(4) Fischer, S.; Verma, C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9613−9615.
(5) Levitt, M.; Park, B. H. *Structure* **1993**, *1*, 223−226.
(6) Yu, B.; Blaber, M.; Gronenborn, A. M.; Clore, G. M.; Caspar, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 103−108.
(7) Ernst, J. A.; Clubb, R. T.; Zhou, H.-X.; Gronenborn, A. M.; Clore, G. M. *Science* **1995**, *267*, 1813−1816.
(8) Roux, B.; Nina, M.; Pomès, R.; Smith, J. C. *Biophys. J.* **1996**, *71*, 670−681.
(9) Denisov, V. P.; Peters, J.; Hörlein, H. D.; Halle, B. *Nat. Struct. Biol.* **1996**, *3*, 505−509.
(10) Denisov, V. P.; Venu, K.; Peters, J.; Hörlein, H. D.; Halle, B. *J. Phys. Chem. B* **1997**, *101*, 9380−9389.
(11) Gottschalk, M.; Dencher, N. A.; Halle, B. *J. Mol. Biol.* **2001**, *311*, 605−621.
(12) Rashin, A. A.; Iofin, M.; Honig, B. *Biochemistry* **1986**, *25*, 3619−3625.
(13) Williams, M.; Goodfellow, J.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 1224−1235
(14) Stetefeld, J.; Jenny, M.; Schulthess, T.; Landwehr, R.; Engel, J.; Kammerer, R. A. *Nat. Struct. Biol.* **2000**, *7*, 772−776.
(15) Peters, J.; Nitsch, M.; Kühlmorgen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; Engelhardt, H.; Pfander, J.-P.; Müller, S.; Goldie, K.; Engel, A.; Stetter, K.-O.; Baumeister, W. *J. Mol. Biol.* **1995**, *245*, 385−401.
(16) Peters, J.; Baumeister, W.; Lupas, A. *J. Mol. Biol.* **1996**, *257*, 1031.
(17) Damjanovic, A.; Garcia-Moreno, B.; Lattman, E. E.; García, A. E. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 433−449.
(18) Bakowies, D.; van Gunsteren, W. F. *J. Mol. Biol.* **2002**, *315*, 713−736.
(19) Collins, M. D.; Hummer, G.; Quillin, M. L.; Matthews, B. W.; Gruner, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16668−16671. Collins, M. D.; Quillin, M. L.; Hummer, G.; Matthews, B. W.; Gruner, S. M. *J. Mol. Biol.* **2007**, *367*, 752−763.
(20) Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. E.; Sweet, B. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G. *Science* **2000**, *287*, 1615−1622.
(21) Taraphder, S.; Hummer, G. *J. Am. Chem. Soc.* **2003**, *125*, 3931−3940.

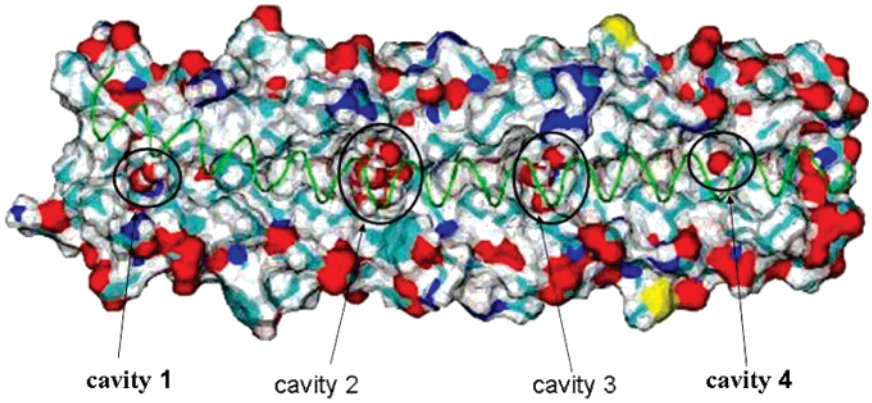

Figure 1. Structure of the tetrabrachion stalk segment after energy minimization of the crystal structure.[14] Three of the four helices are shown as molecular surfaces, and the fourth one (on top) is indicated as a green backbone tube. Water molecules in the four cavities, numbered 1 to 4 from left to right along the central channel, are shown as VDW models (red and white; black circles). Solvent water molecules are not shown for clarity.

contains a four-stranded parallel coiled coil that is involved in binding to a thermostable protease.[22] In addition to being right-handed, it is unusual also in other ways. Thermodynamically, even at 130 °C, full denaturation of the stalk requires 6 M of guanidine.[16] Structurally, the coiled coil presents several unusually large hydrophobic cavities located along its axis.[14] The cavities are connected by a narrow and continuous central channel that is lined exclusively with aliphatic side chains. In the crystal structure, determined at ~100 K, all cavities are occupied by water, with nine water molecules in the largest of the four cavities, five in the second largest, and two and one in the remaining cavities, respectively (Figure 1). Because of the lack of polar groups in the cavity walls, and the resulting weak water−protein interactions, these water molecules form tight clusters in two of the four cavities of the crystal structure.

Computer simulations and theoretical studies[23−31] have been used to probe the conditions under which protein cavities are occupied by water. Transient water penetration into the protein interior was observed even at ambient conditions.[32] At high temperature[32] and pressure,[19,33,34] water penetration is associated with protein unfolding. Previous studies of water in simple nonpolar cavities, including fullerenes, have shown that the thermodynamics of filling is controlled by a fine balance between the energy and entropy of transfer into a confined environment, in which thermodynamic stability is determined collectively by the bulk solvent activity, the cavity size and polarity, hydrogen bonding between water molecules,[35] and their interactions with the cavity wall.[36]

Here, we will explore the following questions. Are the tetrabrachion cavities, in particular the largest one, filled with water at 365 K (92 °C), the optimal growth temperature of *Staphylothermus marinus*? And can we gain insights into a possible functional role of such large nonpolar cavities in a protein whose primary function is to stabilize the surface layer of the organism? We studied the thermodynamics of water penetration into the largest cavity by using MD simulations at 365 K and also at 298 K. The simulations at 298 K were done to quantify the effect of temperature on filling. On the basis of a simple yet powerful new formalism, we are able to determine not only free energies but also the energies and entropies of the interior water from MD simulations. From these we predict the drying temperature of the largest cavity.

The outline of the paper is as follows. After a discussion of the underlying theory and computer simulation methods, we present results of a comprehensive study of the thermodynamics of filling the largest cavity in tetrabrachion. In particular, we show that a marginally stable hydrogen-bonded cluster of six to nine water molecules forms even at 92 °C, with a slightly negative (i.e., favorable) free energy of transfer from the bulk into the protein cavity. We then show that the smaller water cluster with five water molecules in the second largest cavity is only stable at room temperature but unstable at high temperatures where it breaks up to penetrate the protein. From the entropy and energy of transfer, we estimate that even the large cavity will begin to empty at temperatures above that of optimal growth. In conclusion, we present a hypothesis that the near coexistence between a filled ("wet") and empty ("dry") state of the cavity may be functionally relevant. In particular, we suggest that drying of the nonpolar cavity may explain the unusually thermostable protein complex formed between the tetrabrachion stalk and its protease.

2. Theory

To quantify the amount of water filling a cavity inside a protein, we determine the probability $P(N)$ of finding exactly N water molecules inside a fixed volume V covering the cavity. For practical reasons, we choose spherical volumes V whose center is given as an average of the positions of backbone alpha-carbon atoms of residues lining the cavity. The water occupancy probabilities $P(N)$ will depend on the thermo-

(22) Mayr, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J. *Curr. Biol.* **1996**, *6*, 739−749.
(23) Wade, R. C.; Mazo, M. H.; McCammon, J. A.; Quiocho, F. A. *J. Am. Chem. Soc.* **1990**, *112*, 7057−7059.
(24) Wade, R. C.; Mazo, M. H.; McCammon, J. A.; Quiocho, F. A. *Biopolymers* **1991**, *31*, 919−931.
(25) Helms, V.; Wade, R. C. *Biophys. J.* **1995**, *69*, 81824.
(26) Dunitz, J. D. *Science* **1994**, *264*, 670.
(27) Zhang, L.; Hermans, J. *Proteins: Struct., Funct., Genet.* **1996**, *24*, 433−438.
(28) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 7683−7689.
(29) Olano, R. L.; Rick, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 7991−8000.
(30) Woo, H.-J.; Dinner, A. R.; Roux, B. *J. Chem. Phys.* **2004**, *121*, 6392−6400.
(31) Imai, T.; Hiraoka, R.; Kavalenko, A.; Hirata, F. *J. Am. Chem. Soc.* **2005**, *127*, 15334−15335.
(32) García, A. E.; Hummer, G. *Proteins: Struct., Funct., Genet.* **2000**, *38*, 261−272.
(33) Hummer, G.; Garde, S.; García, A.; Paulaitis, M.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1552−1555.
(34) Kitchen, D. B.; Reed, L. H.; Levy, R. M. *Biochemistry* **1992**, *31*, 10083−10093.
(35) Carey, C.; Cheng, Y. K.; Rossy, P. J. *Chem. Phys.* **2000**, *258*, 415−425.
(36) Vaitheeswaran, S.; Yin, H.; Rasaiah, J. C.; Hummer, G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17002−17005.

dynamic conditions, such as temperature, pressure, salt concentration, osmolyte concentrations, etc. Following our previous studies of water in nonpolar cavities and nanotubes,[36,37] the $P(N)$ can be calculated from a series of independent simulations with $N = 0$, 1, etc. water molecules in the cavity. Those simulations can be combined to construct, in effect, a "grand-canonical" distribution for water in the cavity volume V for a fixed temperature T and a given total chemical potential μ of water.

The grand partition function Ξ is the generating function for the canonical ensemble:

$$\Xi(\mu,V,T) = \sum_{N=0}^{\infty} e^{\beta\mu N}\, Q(N,V,T) \tag{2.1}$$

where $Q(N,V,T)$ is the "canonical" partition function for N molecules in the spherical cavity with volume V at a temperature T, and $\beta = 1/k_BT$ with k_B as Boltzmann's constant. (Note that this is not strictly a canonical ensemble because the protein environment of the cavity is itself fluctuating. More formally, $Q(N,V,T)$ is the restricted partition function with the occupancy number held fixed at N.) The probability of finding N molecules in the volume V is

$$P(N) = \frac{e^{\beta\mu N}\, Q(N,V,T)}{\Xi(\mu,V,T)} \tag{2.2}$$

It follows from this that the Helmholtz free energy ΔA_N of transfer of N water molecules from the bulk into the cavity is given by

$$\frac{\Delta A_N}{N} = -\frac{k_BT}{N} \ln \frac{P(N)}{P(0)} \tag{2.3}$$

In eq 2.3, a small term $pV_{N,\mathrm{bulk}} \approx 0.002$ kJ/mol is ignored,[37] where p is the pressure and $V_{N,\mathrm{bulk}}$ is the volume per water molecule in the bulk. Note that the procedure does not require that the shape of the cavity is close to a sphere.

To obtain the occupancy number distribution $P(N)$, we perform simulations in canonical ensembles with fixed numbers of $N = 1$, 2.... particles in the cavity volume V, instead of using simulations with a fluctuating number of particles in the cavity. This procedure helps avoid sampling issues associated with the insertion and removal of particles in grand-canonical simulations. To determine $P(N)/P(0)$, the ratios $P(N+1)/P(N)$ are calculated for N ranging from zero to $N_{\max}$, where $N_{\max}$ is the maximum number of water molecules that can be accommodated in the cavity of known size without encountering large steric overlaps. $P(0)$ is then obtained by normalization,

$$\sum_{N=0}^{\infty} P(N) = 1.$$

In previous studies of water in nonpolar cavities and in carbon nanotubes,[36,37] it was shown that in an equilibrated system the ratio of occupancy probabilities is given by

$$\frac{P(N+1)}{P(N)} = \frac{zV}{N+1}\langle \exp\{-\beta(U_{N+1} - U_N)\}\rangle_N \tag{2.4}$$

where the fugacity z is related to the excess chemical potential $\mu_{\mathrm{bulk}}^{\mathrm{ex}}$ and density ρ of bulk water through $z = \rho \exp(\beta\mu_{\mathrm{ex}}^{\mathrm{bulk}})$. The canonical ensemble average of $\langle \exp\{-\beta(U_{N+1} - U_N)\}\rangle_N = \exp(-\beta\mu_{\mathrm{ex}}^{N})$ defines μ_N^{ex}, the excess chemical potential for a cavity filled with N water molecules. U_N is the total configurational energy of the system (protein, water, ions, etc.), with exactly N water molecules in the cavity, and U_{N+1} is the energy of the system with an additional water molecule randomly inserted into the cavity volume V.

Equation 2.4 allows us to "combine" the canonical simulations for different occupancies N to estimate ratios of occupancy probabilities. The average of $\langle \exp\{-\beta(U_{N+1} - U_N)\}\rangle_N$ for N water molecules in the cavity was determined from the ratio of the distributions of particle insertion energies $f_{\mathrm{ins},N}(\Delta U)$ into the cavity with N water molecules and the particle removal energies $g_{\mathrm{rem},N+1}(\Delta U)$ from the cavity containing $N + 1$ water molecules using the relation[36,37]

$$\frac{f_{\mathrm{ins},N}(\Delta U)}{g_{\mathrm{rem},N+1}(\Delta U)} = \exp(\beta\Delta U)\, \langle \exp\{-\beta(U_{N+1} - U_N)\}\rangle_N \tag{2.5}$$

Following Bennett,[38] the canonical ensemble average of $\langle \exp\{-\beta(U_{N+1} - U_N)\}\rangle_N$ is obtained by fitting $\ln[f_{\mathrm{ins},N}(\Delta U)/g_{\mathrm{rem},N+1}(\Delta U)]$ in the overlap region of the distribution functions to a straight line with respect to ΔU at a fixed slope β.

The excess chemical potential $\mu_{\mathrm{bulk}}^{\mathrm{ex}}$ of bulk water as a function of temperature can be calculated by particle insertion[39] or removal[40] methods, or again by Bennett's method of overlapping histograms,[38] which is our method of choice.

From the dependence of $P(N)$ on temperature, we can estimate the energy and entropy of transferring water from bulk into the cavity. This method was used previously in determining the energy of transfer of water from the bulk into small nonpolar cavities and periodically replicated nanotubes.[36,37] However, since it involves differences in total energies of the system, it was found to be insufficiently accurate for studies of the transfer energy of water into the cavities of large proteins. Here, we instead estimate the energy of transfer as the difference in the average energy of the N-water cluster in the cavity and of N water molecules in the bulk phase. Per water molecule, the latter is given simply in terms of the average removal energy of a single water molecule from bulk water, $U/N = \langle U_{\mathrm{rem}}\rangle_{\mathrm{bulk}}/2$. Assuming a relatively rigid cavity and thus negligible reorganization contributions of the protein upon filling, the former is approximately the average removal energy of the N-water cluster. The energy of transfer per water molecule is then given approximately by $\Delta U/N = \langle U_{\mathrm{rem}}\rangle_N/N - \langle U_{\mathrm{rem}}\rangle_{\mathrm{bulk}}/2$. At 298 and 365 K, the removal energies $\langle U_{\mathrm{rem}}\rangle_{\mathrm{bulk}}$ of one TIP3P water molecule[41] from bulk are -80.2 kJ/mol and -73.0 kJ/mol, respectively.

The entropy of transfer per water molecule in units of Boltzmann's constant k_B was calculated from the difference between the free energies and energies of transfer per water molecule,

$$\Delta S = (\Delta U - \Delta A_N)/T \tag{2.6}$$

where $\Delta S = \Delta S(N)$ and $\Delta U = \Delta U(N)$ depend on N.

3. Simulations

We performed a series of MD simulations of the tetrabrachion coiled-coil stalk segment in which the occupancy of the largest cavity (cavity 2) was varied between $N = 0$ and $N = 11$. The coordinates of the crystal structure of the polypeptide chain fragment of tetrabrachion were obtained from the protein data bank.[14] The crystal structure, with hydrogen atoms appropriately added, has 3283 protein atoms and 17 water molecules inside the cavities. The protein was solvated in a rectangular box of size $77 \times 114 \times 66$ Å^3 with 7239 TIP3P water molecules[41] and 16 Na^+ ions[42] to neutralize the system. The total number of water molecules contained in the whole system is 7256.

(37) Vaitheeswaran, S.; Rasaiah, J. C.; Hummer, G. *J. Chem. Phys.* **2004**, *121*, 7955−7965.
(38) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245−268.
(39) Widom, B. *J. Chem. Phys.* **1963**, *39*, 2808−2812.
(40) Shing, K. S.; Gubbins, K. E. *Mol. Phys.* **1982**, *46*, 1109−1128.
(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(42) 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.

MD simulations were carried out with the sander module of AMBER 6.0 (University of California at San Francisco) with the parm94 force field.[42] The particle mesh Ewald method was used for the long-range electrostatics.[43] A real-space interaction cutoff of 10 Å was applied in the MD simulations. Figure 1 shows the structure of the protein with the cavities filled with two, nine, five, and one water molecules, respectively, at the end of the energy minimization. The system was then heated gradually from 0 K to the desired temperatures of 298 and 365 K, respectively, over 20-ps periods in MD runs at constant volume with time steps of 2 fs. Equilibration at the desired temperature was achieved over periods of 500 ps in *NPT*-ensemble MD simulations, with the pressure tuned to 1 atm, using a time step of 1 fs. The temperature and pressure were maintained using Berendsen thermostats and barostats[44] with time constants of 1 ps. After equilibration, data were collected during production runs of 1 ns, again with a time step of 1 fs and temperature and pressure scaling.

We used the same procedure for different numbers $N = 0$, 1, 2, ..., 11 of water molecules in the large cavity. Initial conformations were obtained by simply transferring water molecules between the cavity and the bulk, keeping the overall number of water molecules in the simulation box constant. For example, to change from $N = 9$ to $N = 8$ water molecules inside the cavity, we transfer one water molecule from the cavity to a position in the bulk far from the protein surface and then minimize the energy of the whole system.

Configurations of the system for each occupancy N were saved every 0.5 ps during the 1 ns MD production runs. The resulting set of 2000 configurations was used to determine the excess chemical potential of the water molecules in the cavity, from which the free energy of transfer from the bulk reservoir into the cavity was calculated. Checks were made periodically during all equilibration and production runs to ensure that water molecules did not enter or leave the cavity such that the desired cavity occupancy N was always maintained.

Additional simulations of bulk TIP3P water were performed at a constant pressure of 1 atm and at temperatures between 298 and 365 K. From those simulations, we obtained bulk number densities of $\rho = 33.00$ and 30.72 nm^{-3} at 298 and 365 K, respectively, and corresponding excess chemical potentials of −25.53 and −22.43 kJ/mol.

4. Results and Discussions

In all simulations, including those at 365 K, we found the protein to be stable. Typical root-mean-square deviations (RMSDs) from the crystal structure after 1 ns of MD are 1.2 Å at 298 K and 1.6 Å at 365 K, as calculated for the C_α carbon atoms of the peptide backbones.

The free energy differences between occupancies with N and $N + 1$ water molecules were calculated by performing virtual water removals and insertions into the cavity volume for structures along the simulation trajectories. With the shape of the cavity being close to a sphere, we chose a sphere of radius 6 Å as the insertion volume. Note that the results should be independent of the detailed size and shape of the insertion volume as long as it fully covers the cavity space accessible to water and does not overlap with any other cavity (i.e., if any additional volume covered by a larger volume $V' \supset V$ is practically inaccessible to water such that test-particle insertions into those regions always produce large positive energies with a negligible Boltzmann factor). We define the center of the cavity by the alpha carbons of the residues on the four strands of tetrabrachion at the periphery of this cavity (residues 69−79, 17−27, 118−128, and 168−178; see Figure 2).

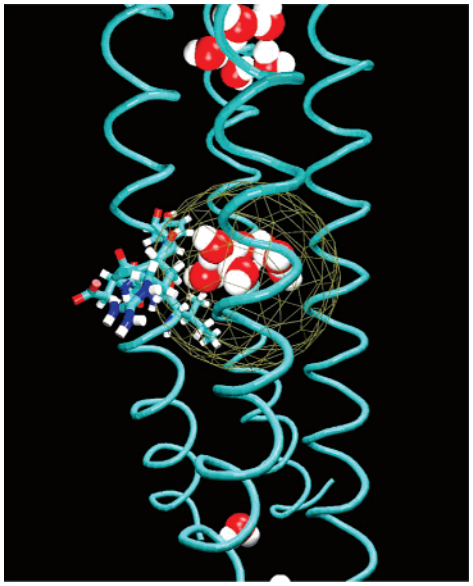

Figure 2. Configuration of tetrabrachion showing the largest cavity occupied by nine water molecules after 500 ps of equilibration at 365 K. The insertion and removal sphere of 6 Å radius is also shown. For clarity, only the residues at the surface of this cavity and for one of the four helices are shown. The cluster at the top indicates the five-water cavity. At the bottom, one can see the two-water cavity.

In all cases, we obtain broad overlap regions of the insertion and removal energies consistent with eq 2.5. Figure 3 shows the overlap region for the insertion and removal of a water molecule from cavity 2 when $N = 2$ and $N + 1 = 3$ and illustrates the calculation of the excess chemical potential of cavity water for $N = 0$, 2, and 4 water molecules, respectively, from the intercept of the plot of $\ln[f_{\mathrm{ins},N}(\Delta U)/g_{\mathrm{rem},N+1}(\Delta U)]$ vs ΔU with slope $1/k_BT$.

From the insertion and removal data we obtain the ratio of probabilities $P(N)/P(0)$ of finding exactly N and 0 water molecules in the cavity and, in turn, the free energy ΔA_N of transferring N water molecules from bulk into the protein cavity 2.

Figure 4 shows the resulting free energies ΔA_N of transfer of water from the bulk into the cavity. The free energy profiles as a function of the occupancy N have two minima: one at the empty state, $N = 0$, and the other corresponding to a filled state with N between about 6 and 10 water molecules. However, at both 298 and 365 K, the transfer free energy of the filled state is negative; i.e., the cavity is preferentially filled. The empty and filled states are separated by a free energy barrier of about $2-3\ k_BT$, indicating that the filling transition is cooperative (i.e., intermediate occupancies of $N \approx 1$ to 4 are unfavorable). Similar barriers were observed previously for water filling carbon nanotubes[37,45,46] and nonpolar cavities.[36]

(43) Darden, T.; York, D.; Petersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(44) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(45) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. *Nature* **2001**, *414*, 188−190
(46) Waghe, A.; Rasaiah, J. C.; Hummer, G. *J. Chem. Phys.* **2002**, *117*, 10789−10795.

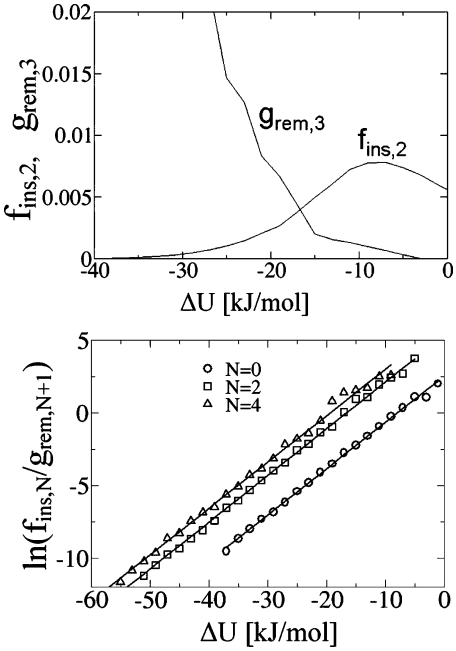

Figure 3. Illustration of the histogram analysis for cavity 2 of tetrabrachion at $T = 365$ K. The top panel shows the overlapping region of insertion and removal energies, respectively, for $N = 2$. The bottom panel shows plots of $\ln[f_{ins,N}(\Delta U)/g_{rem,N+1}(\Delta U)]$ for $N = 0$, 2, and 4. The intercept of the fitted straight line of slope $1/k_BT$ with the $\Delta U = 0$ axis gives the excess chemical potential $-\beta\mu_N^{ex}$.

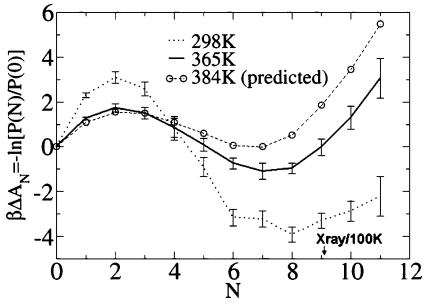

Figure 4. Free energy of transfer of water from the bulk into the interior of cavity 2 as a function of the water occupancy N at 298, 365, and 384 K. Error bars at 298 and 365 K indicate one standard deviation, as estimated from block analyses of calculated data. The free energy of transfer at the predicted transition temperature of 384 K for drying was calculated from the energy and entropy of transfer at 365 K. Occupancy $N = 9$ found in X-ray studies[14] at 100 K is indicated by an arrow.

In particular, transferring a single water molecule into this cavity in tetrabrachion is thermodynamically unfavorable, despite the large size of the cavity (which holds nine water molecules in the X-ray structure) and the presence of four carbonyl oxygen atoms available for hydrogen bonds (residues on the four peptide chains).

Adding a second water molecule is again unfavorable. Only for $N > 2$ does the free energy decrease until it reaches a minimum at an occupancy N of about seven to eight water molecules. For larger occupancies, the free energy increases because of steric overlap and the resulting loss of entropy (see below). Based on Figure 4, we estimate that the cavity is typically filled by seven water molecules at 365 K and by eight water molecules at 298 K. These numbers and the trend with temperature are consistent with the experimental occupancy of nine water molecules in the crystal structure determined at 100 K.[14] The barrier height for filling the cavity with water increases with decreasing temperature.

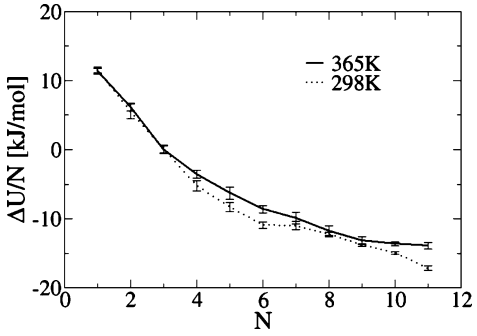

Figure 5. Energy $\Delta U_N/N$ of transfer per water molecule as a function of the occupancy N. Results are shown for transfer of water from bulk into cavity 2 for 298 K (dotted line) and 365 K (solid line).

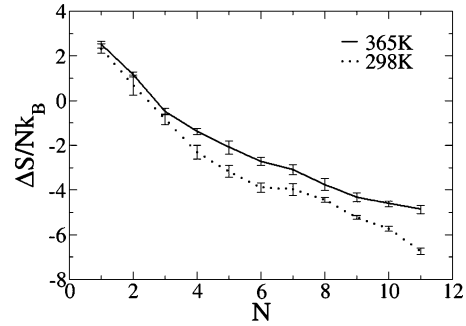

Figure 6. Entropy of transfer per molecule of water from the bulk into the large nonpolar cavity as a function of the occupancy N at 298 K (dotted line) and 365 K (full line).

To characterize the thermodynamic driving force for water molecules to penetrate into this protein cavity, we separate the transfer free energy into energy and entropy contributions using eq 2.6.

Figure 5 shows transfer energies per water molecule, $\Delta U_N/N$, as a function of the occupancy number N. The error in the energy of transfer was calculated from block averages, with blocks of 200 ps duration. We find that the transfer energies at 298 and 365 K are nearly identical, indicating only small contributions from the specific heat of transfer and suggesting that neglecting reorganization energies in eq 2.6 is a good approximation. We also find that the energy of transfer of one and two water molecules into the cavity is large and positive (i.e., unfavorable). However, increasing occupancy produces successively lower transfer energies per molecule which become negative for $N > 3$. These increasingly favorable energies reflect the formation of hydrogen bonds between water molecules that form a single connected cluster inside the cavity. This energy lowering provides the main driving force for water to fill the large nonpolar cavity of tetrabrachion.

Figure 6 shows the entropy of transfer from the bulk into the large nonpolar cavity. The entropy of transfer per water molecule, like the energy of transfer, decreases with increasing cavity occupancy N, indicating strong entropy−enthalpy compensation. The entropy, like the energy, is positive for the transfer of one and two water molecules but negative for $N > 3$.

One and two water molecules have more free space to move in the cavity than in bulk water, and the entropy of transfer is positive as expected. For 3 to 11 water molecules, the entropy

Table 1. Thermodynamics of Transfer of TIP3P Water from the Bulk into Cavity 2 of Tetrabrachion at 365 K and 1 atm of Pressure

N	$\beta\Delta A$	$\Delta A/N$ (kJ/mol)	error	$\Delta U/N$ (kJ/mol)	error	$\Delta S/Nk_B$	error
0	0	0	0	0	0	0	0
1	1.26	3.82	±0.21	11.42	±0.53	2.50	±0.24
2	1.73	2.63	±0.27	6.16	±0.53	1.17	±0.26
3	1.50	1.52	±0.26	0.03	±0.62	−0.49	±0.29
4	0.85	0.65	±0.32	−3.55	±0.59	−1.38	±0.30
5	0.10	0.06	±0.18	−6.29	±0.91	−2.09	±0.36
6	−0.75	−0.38	±0.13	−8.62	±0.52	−2.72	±0.21
7	−1.09	−0.47	±0.15	−9.90	±0.80	−3.11	±0.32
8	−0.97	−0.37	±0.09	−11.79	±0.81	−3.76	±0.30
9	−0.02	−0.01	±0.13	−13.17	±0.59	−4.33	±0.23
10	1.31	0.40	±0.15	−13.61	±0.29	−4.61	±0.15
11	3.06	0.84	±0.24	−13.93	±0.48	−4.86	±0.24

change is negative and unfavorable due to cluster formation and interactions with the cavity walls.

The calculated entropies at the two temperatures are again nearly identical. From the energy, entropy, and free energy values in Tables 1 and 2, we find that the small changes in entropy and energy are thermodynamically consistent. For instance, for $N = 7$ the difference in transfer entropies at the two temperatures of $+0.87\ k_B$ per molecule would, at constant heat capacity, correspond approximately to a change in energy of 2.4 ± 1.7 kJ/mol, compared to the directly calculated change of 1.1 ± 1.4 kJ/mol. From the energy and entropy of transfer at 298 and 365 K, we can estimate a specific heat of transfer per molecule using $\Delta C_V = \mathrm{d}\Delta U/\mathrm{d}T = T\mathrm{d}\Delta S/\mathrm{d}T$ of 16 ± 21 and 36 ± 25 J K^{-1} mol^{-1}, respectively.

To explore the trends in the energy and entropy of transfer in more detail, we need to look at the structure of cavity water as a function of the occupancy N. The positions and orientations of a single water molecule in the cavity provide information on the preferred binding sites and a hint of protein−water interactions at higher cavity occupancy. By following energy changes in a 1 ns trajectory, we observe that a single water molecule is bound to or moving between hydrogen-bonding sites at the carbonyl groups of the four residues on the four peptides in the cavity wall (Figure 7). The removal-energy fluctuation of the water molecule is displayed in Figure 8, illustrating the rapid motion of water between low-energy binding sites. The cavity free volume is large, resulting in a positive and favorable entropy of transfer of a single molecule. The balance between the energy and entropy is reflected in a positive (unfavorable) free energy of transfer of one water molecule into the cavity.

Table 2. Thermodynamics of Transfer of TIP3P Water from the Bulk into Cavity 2 of Tetrabrachion at 298 K and 1 atm of Pressure

N	$\beta\Delta A$	$\Delta A/N$ (kJ/mol)	error	$\Delta U/N$ (kJ/mol)	error	$\Delta S/Nk_B$	error
0	0	0	0	0	0	0	0
1	2.29	5.67	±0.25	11.44	±0.33	2.33	±0.23
2	3.09	3.83	±0.32	5.55	±1.092	0.70	±0.57
3	2.57	2.12	±0.26	0.06	±0.54	−0.83	±0.32
4	0.82	0.51	±0.33	−5.21	±0.76	−2.31	±0.44
5	−0.90	−0.45	±0.21	−8.29	±0.64	−3.17	±0.35
6	−3.15	−1.30	±0.15	−10.93	±0.49	−3.88	±0.26
7	−3.22	−1.14	±0.12	−10.99	±0.62	−3.97	±0.30
8	−3.91	−1.21	±0.10	−12.21	±0.26	−4.44	±0.15
9	−3.30	−0.91	±0.09	−13.85	±0.19	−5.22	±0.11
10	−2.87	−0.71	±0.11	−14.92	±0.19	−5.73	±0.12
11	−2.21	−0.50	±0.20	−17.17	±0.33	−6.73	±0.21

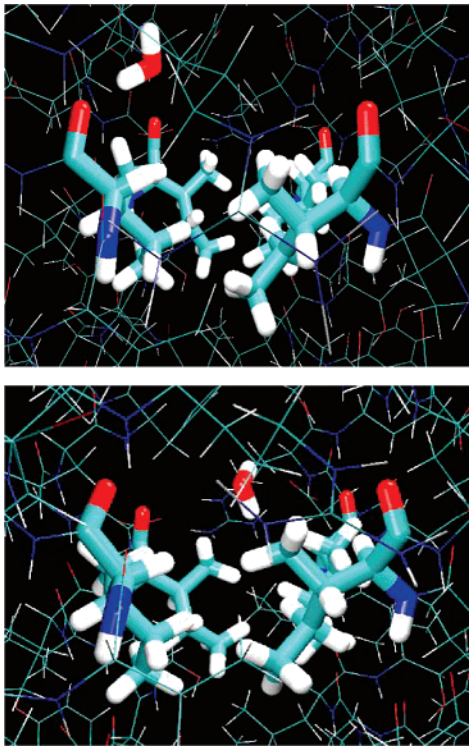

Figure 7. Lowest (Top) and highest (bottom) removal-energy equilibrium structures of a single water molecule in cavity 2 at 365 K. The four hydrophilic carbonyl oxygen atoms are represented in red.

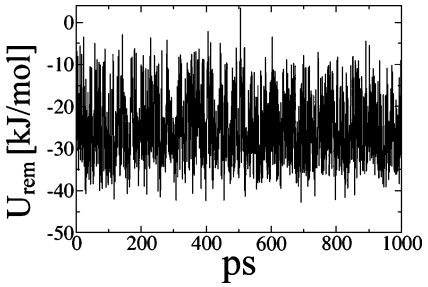

Figure 8. Removal energy of a single water molecule in cavity 2 of tetrabrachion as a function of time at 365 K. The time averaged removal energy is −25.08 kJ/mol.

The addition of the second water molecule to the cavity could lead to the formation of a water dimer or the binding of each water molecule to different carbonyl oxygens of the isoleucine residues lining the walls of the cavity. Figure 9 shows snapshots of configurations corresponding to the lowest and highest removal energies of two water molecules in the cavity. This energy includes water−dimer and water−protein interactions.

The second water molecule prefers to form a hydrogen bond with the first water molecule rather than with the protein wall of the cavity for both the lowest and highest energy configurations observed in the simulations. This implies that hydrogen bonding and van der Waals interactions between the water and the protein are much weaker than the hydrogen bonding between two water molecules.

Equilibration after successive additions of more water molecules clearly shows that the water molecules are preferentially hydrogen bonded to each other to form a single cluster of increasing size. There is also a tendency for the water molecules in the cluster to be hydrogen bonded to one or more of the hydrophilic sites of the cavity wall. Equilibrium configurations

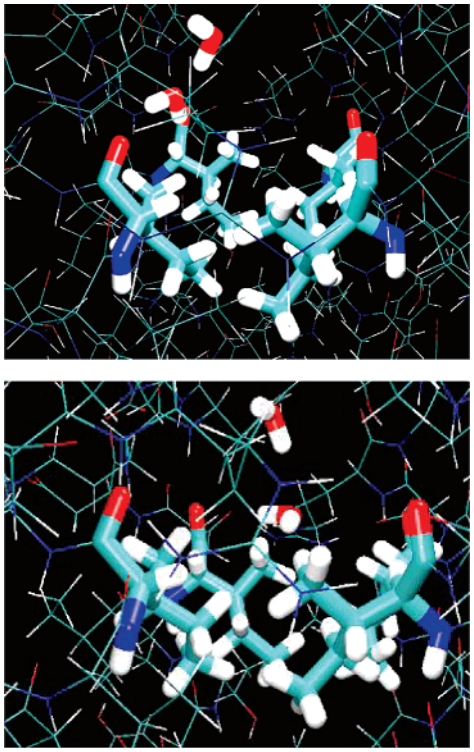

Figure 9. Lowest (top) and highest (bottom) removal energy equilibrium structures of two water molecules in cavity 2 at 365 K.

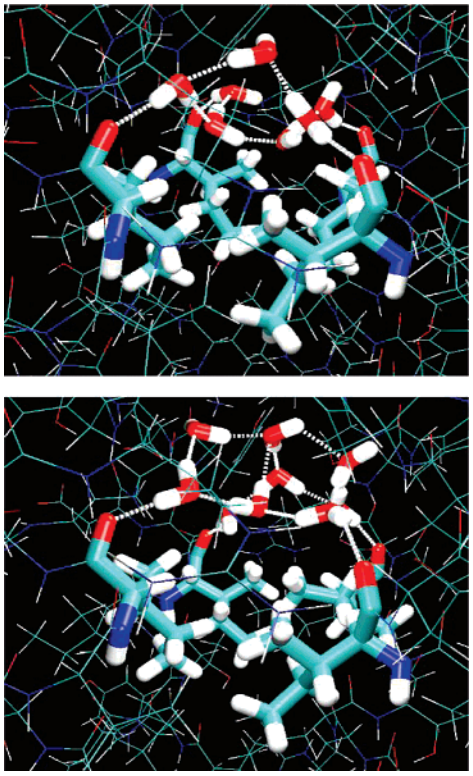

Figure 10. Lowest removal-energy equilibrium structures of seven (top) and nine water molecules (bottom) in cavity 2 at 365 K. Hydrogen bonds are shown as dotted lines.

of the structures of seven and nine water molecules inside the cavity are shown in Figure 10, with close-ups of the seven and eight water clusters in Figure 11.

The large clusters in the largest protein cavity of tetrabrachion are remarkably similar to the clusters found earlier in our studies of idealized nonpolar cavities.[36]

From the energies and entropies of transfer at 365 K (see Table 1), we predict that the transfer free energy of clusters with seven to nine water molecule cross over from negative to positive values at ~384 K, signaling the onset of the emptying of cavity 2 (see Figure 4). This is still about 20 °C below the temperature (403 K) at which tetrabrachion unfolds in the presence of strong denaturants (1% dodecyl sulfate or 6 M guanidine or 70% sulfuric acid),[15,16] which suggests that drying precedes denaturation and cavity water contributes to the stability, and possibly function, of this protein.

In the following, we briefly discuss the hydration of the other, smaller cavities in tetrabrachion. Besides the largest cavity (cavity 2) three other nonpolar cavities (see Figure 1) have been located in the coiled-coil segment by crystallography.[14] Cavities 1 and 4 are close to the bulk reservoir. At 365 K, we find that water molecules escaped from the cavities into the bulk, without water re-entering during the relatively short time of our 1 ns simulation. The second largest cavity (cavity 3) has a five water molecule pentagonal cluster that also broke up at 365 K. Three of the water molecules escape from the cavity, leaving two water molecules behind toward the end of the 1 ns simulation at 365 K.

One of the water molecules released during the disintegration of the cluster is temporarily immobilized by hydrogen bonding to the carbonyl oxygen of a residue in the cavity wall. It then adjusts its orientation to form a new hydrogen bond with a different carbonyl oxygen atom of a residue in the cavity wall

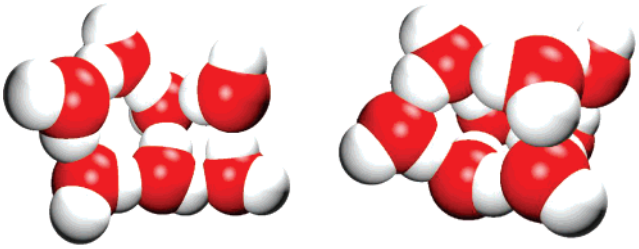

Figure 11. Lowest removal energy equilibrium structures of the hydrogen-bonded seven-water cluster at 365 K (left) and of the eight-water cluster in cavity 2 at 298 K (right).

and eventually penetrates the protein. In this way, three water molecules escape from cavity 3. Even after the water molecules escape, the protein remains stable, consistent with the backbone RMSD of 1.6 Å reported above.

The behavior of water in the smaller cavities (1, 4, and 3; see Figure 1) at 298 K is rather different. At the lower temperature, the water molecules in cavities 1 and 4 exchange with bulk water (i.e., water molecules go in and out of the cavity during the simulations). However, the five-water pentagonal hydrogen-bonded cluster in cavity 3 remains stable, in contrast to its behavior at 365 K. Figure 12 shows the lowest energy configuration of this cluster in our MD simulation. The energy of transfer from the bulk reservoir to form this cluster in cavity 3 is favorable, about −7.7 kJ/mol. However, removal of one or more water molecules at 298 K destabilizes the depleted cluster, and the remaining water molecules penetrate the protein wall individually, as they do at high temperature, before escaping into the bulk reservoir. The movement of water in and out of the smaller cavity 3 prevents us from using the methods applied to the larger cavity 2 to determine the free energy of transfer as

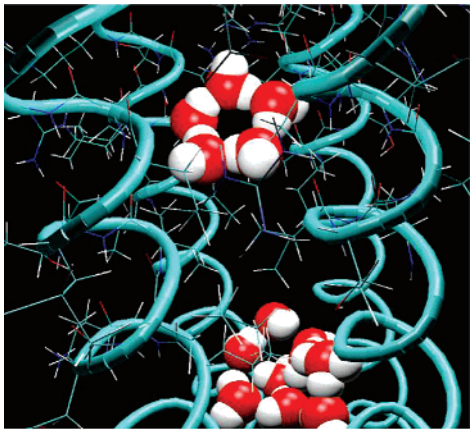

Figure 12. The lowest removal-energy configuration of the five-water molecule pentagonal cluster in the second largest cavity (cavity 3) of tetrabrachion during the 298 K MD simulation. The four protein strands of tetrabrachion are represented as tubes, and the water molecules, by van der Waals spheres. The protein residues around the pentagonal water molecule cluster are shown using a line model. The cluster at the bottom indicates the nine-water cluster in the largest cavity (cavity 2).

a function of the occupancy, unless a restraining potential is applied. We leave this for a future investigation.

The cavity occupancy is sensitive to the activity of the solvent. The calculations at 365 and 298 K were carried out at a pressure of 1 atm. Increased hydrostatic pressure in deep sea thermal marine environments where tetrabrachion is found[14−16] will produce changes in the chemical potential and density of the solvent tending to favor filling;[32,33] however, increased salt and osmolyte concentrations would favor drying. The change in the transfer free energy is related to changes in the solvent density and chemical potential by[36,37]

$$\Delta A_N(T,P) - \Delta A_N(T,1\text{atm}) = -NkT \ln\left(\frac{\rho(T,P)}{\rho(T,1\text{atm})}\right) - N(\mu^{\text{ex}}_{\text{bulk}}(T,P) - \mu^{\text{ex}}_{\text{bulk}}(T,1\text{atm})) \quad (4.1)$$

Although the density increases with pressure, it appears as a logarithmic term in the free energy change, and its effect is typically smaller than that due to an increase in the excess chemical potential of water. Pressure favors filling[19,33] and a slightly larger optimal cluster size. It may also raise the transition temperature at which the cavities begin to empty.

What are possible functional implications of our results? Consistent with extensive studies of artificially created cavities in lysozyme by Matthews and collaborators,[47,48] we found that tetrabrachion is stable, independent of the cavity being filled or empty. The observed stability is also consistent with experimental results for a lysozyme mutant under pressures up to at least 2 kbar.[19]

However, our thermodynamic results indicate that, at the temperature of optimal growth of *Staphylothermus marinus*, the two largest cavities of its surface-stabilizing protein tetrabrachion are near the midpoint of their filling/emptying transitions. Such drying of molecular pores[45,46] has been implicated, for instance, in the function of biological channels.[49,50] Here, the drying may play a role in protein complex formation. Interestingly, the two protease molecules are attached to the tetrabrachion stalk essentially on top of the largest cavity.[16] It is therefore conceivable that the cavities are used in binding the proteases, possibly through nonpolar molecular anchors penetrating into the cavity and replacing the water clusters. By having a cavity close to the midpoint of an emptying/filling transition, its large nonpolar surface can be made available for such binding processes at essentially zero cost in free energy. Similar effects have been suggested in binding of small ligands to proteins[51] and the formation of protein complexes.[52] Indeed, the complex between the tetrabrachion stalk and its two proteases appears to be stably bound at temperatures at least as high as 120 °C (398 K). We therefore speculate that the unusually large nonpolar cavity in tetrabrachion may serve as a binding element for possible nonpolar anchors of its proteases.

5. Conclusions

The right-handed coiled-coil tetramer structure of tetrabrachion is a model for understanding the stability of thermophilic proteins and for the design of de novo proteins.[14] The stability of this protein at high temperatures is related to its unusual structure, resulting in a series of large, almost entirely nonpolar, yet water-filled cavities.

We have carried out a comprehensive study of the thermodynamics of transfer of water from the bulk phase into the largest cavity in tetrabrachion at 298 and 365 K. We have shown that water molecules form a thermodynamically stable hydrogen-bonded cluster inside the largest cavity (cavity 2; see Figure 1) both at the high temperature corresponding to the hydrothermal marine environment in which tetrabrachion exists (365 K) and at room temperature (298 K). The typical cluster at 365 K consists of seven water molecules, hydrogen bonded to each other and stabilized further through hydrogen bonds and weaker van der Waals interactions with the cavity walls. The filled and empty states are separated by a barrier of 2−3 k_BT, with the absolute barrier height increasing with temperature decreasing.

The entropy of transfer per molecule of the thermodynamically most stable cluster of seven water molecules from the bulk into the cavity is approximately −25.8 J/(K·mol) at 365 K, and the energy of transfer per molecule is −9.9 kJ/mol. Filling is driven by the favorable transfer energy at this temperature (see Table 1) and also at room temperature (298 K) (Table 2). The energy and entropy of transfer are found to depend only weakly on temperature, with the estimated entropy being slightly more negative at 298 K than at 365 K (Figure 6). A wider range of thermodynamically stable cluster sizes (N = 6 to 9) of comparable free energy exist at 298 K than at 365 K.

From our estimates of the entropy and energy of transfer, we predict that the transfer free energy of clusters with seven to nine water molecules would become unfavorable above temperatures of ∼384 K, signaling the onset of drying above (see Figure 4). This result suggests that emptying occurs before denaturation (with the protein being stable at temperatures >400 K).[16] In our simulations tetrabrachion is stable at 365 K

(47) Erickson, A. E.; Baase, W. A.; Matthews, B. W. *J. Mol. Biol.* **1993**, *229*, 747−769.
(48) Morton, A.; Baase, W. A.; Matthews, B. W. *Biochemistry* **1995**, *34*, 8564−8575.
(49) Beckstein, O.; Biggin, P.; Sansom, M. *J. Phys. Chem. B* **2001**, *105*, 12902−12905.
(50) Hummer, G. *Mol. Phys.* **2007**, *105*, 201−207.
(51) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *J. Med. Chem.* **2006**, *49*, 6177−6196.
(52) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808−813.

even if the large cavity 2 is empty. Based on this and the experimental observation that the stalk and two proteases are stably bound even at 398 K, we suggest that a possible functional role of the largest nonpolar cavity (cavity 2) in tetrabrachion is to serve as the binding element for nonpolar anchors of its proteases.

We note that the present formalism to study the hydration thermodynamics of confined systems should prove useful in other cases where cavities can be occupied by multiple water molecules, as shown before for simple spherical cavities and fullerenes,[36] nanotubes,[37,45,46] and here proteins. Displacement of water is an important factor in the thermodynamics of ligand binding. Our method should help in calculating accurate ligand-binding free energies, even in cases with an unknown number of water molecules in the binding site, or when the number and positions of water fluctuate.

Acknowledgment. J.C.R. and H.Y. acknowledge support from the National Science Foundation (Grant No: CHE 0549187). G.H. was supported by the Intramural Program of the NIDDK, NIH. We thank Dr. Guogang Feng and Aparna Waghe for their comments, the University of Maine Supercomputing Center for generous allocations of computer time on the Beowulf cluster, and Dr. John Koskie and Justin Bronder of the Supercomputer Center for their assistance.

JA070456H