Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 9613–9615, August 1999
Biophysics

Binding of buried structural water increases the flexibility of proteins

(bovine pancreatic trypsin inhibitor/gas-phase hydration/free energy of binding/vibrational entropy)

STEFAN FISCHER*†‡ AND CHANDRA S. VERMA§

*Computational and Structural Chemistry, Hoffmann-La Roche/Pharma Research, CH-4070 Basel, Switzerland; and §Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, United Kingdom

Edited by Hans Frauenfelder, Los Alamos National Laboratory, Los Alamos, NM, and approved June 21, 1999 (received for review May 10, 1999)

ABSTRACT **Water deeply buried in proteins is considered to be an integral part of the folded structure. Such structural water molecules make strong H bonds with polar groups of the surrounding protein and therefore are believed to tighten the protein matrix. Surprisingly, our computational analysis of the binding of a buried water molecule to bovine pancreatic trypsin inhibitor shows that the protein actually becomes more flexible, as revealed by an increase in the vibrational entropy. We find that this effect must be common in proteins, because the large entropic cost of immobilizing a single water molecule [$-T\Delta S = 20.6$ kcal/mol (1 kcal = 4.18 kJ) for the lost translational and rotational degrees of freedom] can only be partly compensated by water–protein interactions, even when they are nearly perfect, as in the case of bovine pancreatic trypsin inhibitor ($\Delta E = -19.8$ kcal/mol), leaving no room for a further decrease in entropy from protein tightening. This study illustrates the importance of considering changes in protein flexibility (which in this case favor binding by 3.5 kcal/mol) for the prediction of ligand binding affinities.**

Contributions of water to the structural, dynamic, and functional properties of proteins are well known.(1, 2) However, detailed experimental information about the interactions with single water molecules has only become available recently (3–6) and complemented by computational studies (7–9). Of great interest is a recent study by Woenckhaus *et al.* (10), who have used mass spectrometry to measure the enthalpy and entropy changes associated with the binding of the first water molecule to fully dehydrated bovine pancreatic trypsin inhibitor (BPTI). This water molecule is deeply buried and makes four optimal H bonds to the protein (11), so that the enthalpy of the first hydration step ($\Delta H = -21.3$ kcal/mol; 1 kcal = 4.18 kJ) is significantly more negative than for the second and subsequent hydration steps. They found that this enthalpy is compensated for by a large decrease in entropy ($\Delta S = -62$ cal·mol^{-1}·K^{-1}). Based on qualitative estimates, they concluded that about half of this decrease must result from a conformational locking of the protein induced by the tight binding of the water molecule. Here, we calculate the thermodynamic parameters of this binding step by using molecular mechanics to get a quantitative understanding of the binding entropy. Because the experimental system was in the gas phase, it is particularly amenable to a treatment by normal-mode analysis. We arrive at the unexpected conclusion that, in spite of the four newly created H bonds, the binding of the first water actually loosens the protein, as indicated by an increase in the vibrational entropy. The remaining decrease in the experimental entropy can be fully accounted for by the loss of the rotational and translational degrees of freedom of the water molecule.

Well known formulae relate the translational entropy (S_{trans}) to the mass, the rotational entropy (S_{rot}) to the moments of inertia and the vibrational entropy (S_{vib}) to the vibrational spectrum (shown in Fig. 1*a*) of a molecule(12). Their use in the context of the dimerization of insulin has been described in detail (13). We applied them in the same manner to the binding of one water molecule (denoted W) to fully dehydrated bovine pancreatic trypsin inhibitor (denoted P): P + W → PW. Computing the total entropy $S = S_{trans} + S_{rot} + S_{vib}$ for P, W, and the complex PW yields the binding entropy $\Delta S = S(PW) - S(P) - S(W)$. Contributions to the binding enthalpy from the interactions between the protein and the water molecule were obtained from the difference between the minimized potential energy of the complex and the isolated protein: $\Delta E = E(PW) - E(P) - E(W)$. The resulting values are summarized in Table 1.

After energy minimization, the rms difference between P and PW is only 0.1 Å over the backbone atoms, showing that binding of W does not perturb the protein structure. Moreover, binding does not contract the protein, whose collision cross-section (calculated as $5\pi R_g^2/3$, where R_g is the radius of gyration)(14) remains unchanged at 581 Å^2, in agreement with the value 548 ± 30 Å^2 measured for dehydrated bovine pancreatic trypsin inhibitor (15). These structural results are a first indication that there is no global tightening of the protein.

Dissecting the components of the interaction energy in the PW complex shows that the four quasi-optimal H bonds between W and the surrounding protein backbone contribute −20 kcal/mol to the electrostatic energy. This is partly compensated for by a weakening in the electrostatic interactions of the protein with itself (+3.6 kcal/mol), resulting in a net contribution to ΔE_{elec} of −16.4 kcal/mol (Table 1). Because ΔE_{elec} is the major component of ΔE, the formation of the four H bonds dominates the enthalpy and drives the binding process. Although the strong H bonds can be expected to tighten the protein "cage" around the water, the observed increase in the electrostatic self-energy of the protein is indicative of a global weakening in the protein interactions. That the latter actually dominates the former in their effect on global protein flexibility is reflected in the vibrational entropy.

The entropic cost of immobilizing a single water molecule ($\Delta S_{trans} + \Delta S_{rot}$) is −68.6 cal·mol^{-1}·K^{-1} (Table 1). In view of this large negative value, it is not necessary to invoke a change in conformational entropy, for example because of some locking of the protein, to account for the negative binding entropy observed experimentally. The positive change in the vibrational entropy ($\Delta S_{vib} = +11.7$ cal·mol^{-1}·K^{-1}) indicates

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviation: BPTI, bovine pancreatic trypsin inhibitor.
†Present address: Im Neuenheimerfeld 368, IWR, Biocomputing/ Universität Heidelberg, D-69120 Heidelberg, Germany.
‡To whom reprint requests should be addressed. E-mail: stefan.fischer@iwr.uni-heidelberg.de.

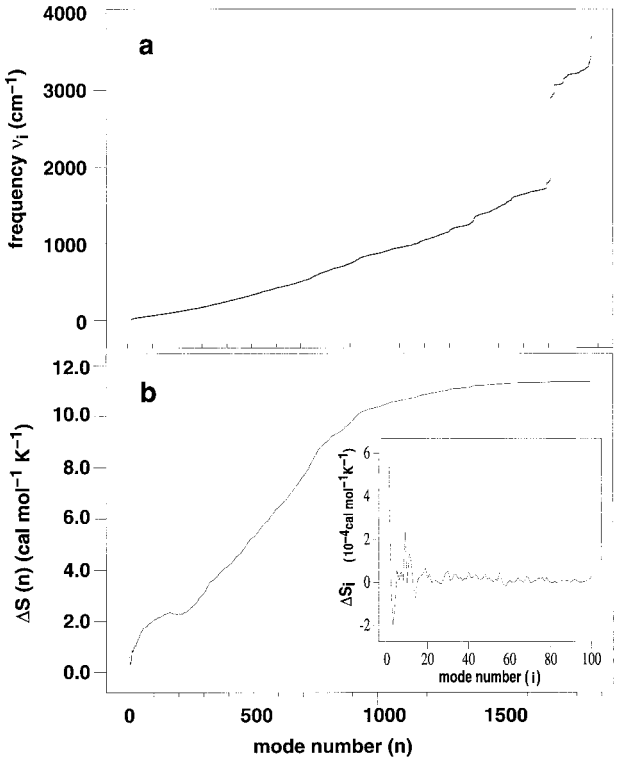

FIG. 1. (*a*) Vibrational spectra of unbound (P) and complexed (PW) bovine pancreatic trypsin inhibitor (cannot be distinguished on this scale). The vibrational frequencies ν_i correspond to the normal modes of the energy-minimized structure (20) obtained by diagonalizing the mass-weighted matrix of the second derivatives of the energy (21). (*b*) Cumulative change of the vibrational entropy $\Delta S(n) = \sum_{i=1}^{n} \Delta S_i$ on water binding, where $\Delta S_i = (S_i^{PW} - S_i^{P})$ is the contribution of the *i*th vibration mode. The *Inset* shows ΔS_i for the first 100 modes. $S_i = [h\nu_i/T/(e^{h\nu i/kT} - 1) - k\ln(1 - e^{-h\nu i/kT})]$ is the entropic content of one vibration mode *i* (12).

that the overall flexibility of the protein actually increases on binding. What is the origin of this entropy increase? The three translational and three rotational degrees of freedom of unbound water are replaced in the PW complex by six new libration modes of the water in its protein "cage." The corresponding libration frequencies are in the 200–800 cm^{-1} range (16), so that the combined vibrational entropy content (obtained with the equation in the caption of Fig. 1*b*) of these six modes cannot contribute more than 0.012 cal·mol^{-1}·K^{-1} to ΔS_{vib} (though they account for most of the 3 ± 0.2 kcal/mol change in the zero-point energy). Fig. 1*b* shows that the large ΔS_{vib} results from the accumulation of many small changes in the vibrational spectrum over the first 900 modes of the protein. The 20 lowest modes representing global motions of large amplitude are affected most (see Fig. 1*b Inset*), but their net contribution is minor. It is necessary to include high-frequency modes ($\nu_i < 800$ cm^{-1}), such as angle bending, to account for the full effect on ΔS_{vib}. Although a few modes get tightened, the majority of modes have an increased entropy content, showing that the binding results in a general increase in the flexibility of the protein. This result indicates that most parts of the protein, both locally and globally, are in a tense state in the absence of the structural water molecule. For instance, when considering the nonbonded interactions between all possible pairs of protein residues, one finds that 69% of these pairs have their interactions weakened on binding, particularly between residues that are around the bound water.

The total entropic change $\Delta S = -56.9 \pm 2$ cal·mol^{-1}·K^{-1} is close to the experimental value (Table 1), and its contribution to the free energy of binding ($-T\Delta S = 17.1$ kcal/mol at room temperature) compensates for a large fraction of the binding enthalpy. This good agreement with the experimental data suggests that our calculations have accounted for the relevant effects in this particular case of ligand binding (fluctuations of <10% in the values calculated for a sample of 100 conformations demonstrate that it is not fortuitous) and raises the question of how much protein tightening can be expected on binding of structural water in general. Considering that the large entropic cost of immobilizing one water molecule [$-T(\Delta S_{trans} + \Delta S_{rot}) = +20.6$ kcal/mol at 300 K) must be paid for in any case and that a perfect binding site providing four H bonds with neutral donors and acceptors will not contribute much more that −20 kcal/mol to the binding enthalpy (as seen here), there is little room for further entropic costs from conformational locking or vibrational tightening. If the net free-energy change is to favor binding to a neutral protein cavity, we conclude that structural water is more likely to increase than to decrease the flexibility of proteins in general. Similar observations have been made for the binding of water molecules to proteins in solution, where it was concluded that despite extensive H bonding to the protein, the buried water

Table 1. Calculated entropy and potential energy of the first hydration step in BPTI

	P	+	W	→	PW	Δ(PW − P − W)	Experimental*
S_{trans}	47.3		35.6		47.3	−35.6 ± 0	
S_{rot}	70.6		33.0		70.6	−33.0 ± 0	
S_{vib}	1,930.5		0.01		1,941.8	+11.7 ± 2*	
ΔS						−56.9 ± 2*	−62 ± 5
E_{elec}	−2,476.4		0†		−2,492.8	−16.4	
E_{vdW}	−264.2		0†		−268.3	−4.1	
E_{coval}	+222.9		0†		+223.4	+0.5	
						−19.8 ±1.9*	−21.3 ± 1

Entropy (S in kcal^{-1}·mol^{-1}·K^{-1} calculated as described in Tidor *et al.* (13) See also Fig. 1*b* legend for the formula of S_{vib}. The potential energy of the protein (E in kcal/mol) was calculated using the CHARMM force field (22) with parameter set 19 (23) and minimized as described (7) with a constant dielectric of 1. Its components are the electrostatic (E_{elec}), van der Waals (E_{vdW}), and covalent (E_{coval}) terms: $\Delta E = \Delta E_{elec} + \Delta E_{vdW} + \Delta E_{coval}$. Measured values of ΔS and ΔH were obtained from Woenckhaus *et al.* (10).

*To generate a representative ensemble of protein conformations, 100 conformers were produced by quenching every 20 ps of a 2-ns gas-phase molecular-dynamics simulation of the protein. For each conformer, pairs of structures (the complex PW and the dehydrated protein P) were obtained by minimization after removing three and all four structural water molecules, respectively. Rather than computing a Boltzmann average over the ensemble, 100 values of ΔS and ΔE were calculated from these pairs of structures, whose average (and SD) is shown here. The values of the energy breakdown are only illustrative and were obtained from a single conformer, i.e., the minimized crystal structure (11).

†E is zero for the isolated TIP3P model of water used here (24).

molecules do not have a significantly lower entropy than bulk water (6, 17).

Increases in protein flexibility have been reported for the binding of larger ligands, for example in the case of the human rhinovirus capsid protein, where molecular dynamics studies have shown that the binding of antiviral compounds is entropy-driven (18). In the case of the dimerization of insulin, changes in the vibrational entropy have been shown to favor binding by 7.2 kcal/mol (13). Despite being only a small ligand, the ability of the water molecule to alter the vibrational state of bovine pancreatic trypsin inhibitor and thereby to significantly contribute in an unexpected way to the binding free-energy ($-T\Delta S_{vib} = -3.5$ kcal/mol) demonstrates that accounting for changes in protein flexibility should be an integral part of methods for predicting the binding affinity of ligands (19).

C.S.V. thanks the Biotechnology and Biological Sciences Research Council for support.

1. Levitt, M. & Park, B. H. (1993) *Structure (London)* **1,** 223–226.
2. Wuthrich, K. (1993) *Cold Spring Harbor Symp. Quant. Biol.* **58,** 149–157.
3. Ernst, J. A., Clubb, R. T., Zhou, H. X., Gronenborn, A. M. & Clore, G. M. (1995) *Science* **267,** 1813–1817.
4. Takano, K., Funahashi, J., Yamagata, Y., Fujii, S. & Yutani, K. (1997) *J. Mol. Biol.* **274,** 132–142.
5. Denisov, V. P., Peters, J., Horlein, H. D. & Halle, B. (1996) *Nat. Struct. Biol.* **3,** 505–509.
6. Denisov, V. P., Venu, K., Peters, J., Horlein, H. D. & Halle, B. (1997) *J. Phys. Chem. B* **101,** 9380–9389.
7. Fischer, S., Verma, C. S. & Hubbard, R. E. (1998) *J. Phys. Chem. B* **102,** 1797–1805.
8. Prevost, M. (1998) *Folding Des.* **3,** 345–351.
9. Abseher, R., Schreiber, H. & Steinhaiser, O. (1996) *Proteins Struct. Funct. Genet.* **25,** 366–378.
10. Woenckhaus, J., Hudgins, R. R. & Jarrold, M. F. (1997) *J. Am. Chem. Soc.* **119,** 9586–9587.
11. Wlodawer, A., Walter, J., Huber, R. & Sjolin, L. (1984) *J. Mol. Biol.* **180,** 301–329.
12. McQuarrie, D. A. (1976) *Statistical Mechanics* (Harper Row, New York).
13. Tidor, B. & Karplus, M. (1994) *J. Mol. Biol.* **238,** 405–414.
14. Collings, B. A. & Douglas, D. J. (1996) *J. Am. Chem. Soc.* **118,** 4488–4489.
15. Nesatiy, V., Chen, Y-L, Collings, B. A. & Douglas, D. J. (1998) *Rapid Commun. Mass Spectrom.* **12,** 40–44.
16. Eisenberg, D. & Kauzmann, W. (1969) *The Structure and Properties of Water* (Oxford Univ. Press, London).
17. Dunitz, J. D. (1994) *Science* **264,** 670.
18. Phelps, D. K., Rossky, P. J. & Post, C. B. (1998) *J. Mol. Biol.* **276,** 331–337.
19. Ajay, W. & Murcko, M. A. (1995) *J. Med. Chem.* **38,** 4953–4967.
20. Wilson, E. B., Jr., Decius, J. C. & Cross, P. C. (1980) *Molecular Vibrations* (Dover, New York).
21. Brooks, B. R., Janezic, D. & Karplus, M. (1995) *J. Comput. Chem.* **16,** 1522–1542.
22. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem.* **4,** 187–217.
23. Neria, E., Fischer, S. & Karplus, M. (1996) *J. Chem. Phys.* **105,** 1902–1921.
24. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983) *J. Chem. Phys.* **79,** 926–935.