CONFIGURATIONAL ENTROPY OF NATIVE PROTEINS

M. KARPLUS, T. ICHIYE, AND B. M. PETTITT
Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

ABSTRACT Simulations of the residual configurational entropy of a protein in the native state suggest that it is nearly an order of magnitude larger than the entropy of denaturation. The implications of this result are discussed.

It is common practice (Schulz and Schirmer, 1978; Privalov, 1979) to estimate the configurational entropy contribution to protein denaturation, ΔS_{conf}, by setting that of the native protein equal to zero and calculating the entropy of the random coil state from the number of possible conformations, generally taken to be of equal energy for simplicity. This leads to a value for ΔS_{conf} in the range of 4 to 6 cal/(mol-of-residue °K) and to a contribution to the free energy of denaturation of 1.2 to 1.8 kcal/(mol-of-residue) at room temperature. Experimental values of the denaturation entropy at ambient temperatures are in this range, although there is some uncertainty concerning the meaning of the results because of their significant temperature dependence due to the large value of the heat capacity of denaturation (Privalov, 1979). We demonstrate here that the residual configurational entropy of a folded protein is nearly an order of magnitude larger than ΔS_{conf}. The implications of this result for the standard model of the denaturation entropy and more generally for the thermodynamic properties of proteins are examined.

Recent developments (Karplus and McCammon, 1983) in the treatment of the internal dynamics of native proteins make possible an estimate of the residual entropy due to the inherent motional freedom of the folded polypeptide chain. Molecular, harmonic, and quasiharmonic dynamics approaches can be used for this purpose (Karplus and Kushick, 1981; Brooks and Karplus, 1983; Levy et al., 1984). Although the harmonic approximation yields an overestimate of the classical entropy (Edholm and Berendsen, 1984; Ichiye, 1985), a harmonic calculation is useful because the essential quantum corrections can be introduced most easily. As an example, we consider a well-studied model system, the bovine pancreatic trypsin inhibitor (BPTI), for which both molecular dynamics and harmonic dynamics results are available (Karplus and McCammon, 1983; Brooks and Karplus, 1983). A calculation of the harmonic spectrum of this molecule, which consists of 58 residues, has been made in Cartesian coordinates, including all heavy atoms and polar hydrogens (580 atoms) (Brooks and Karplus, 1983), with a potential function that has been described previously (Brooks, et al., 1983). The internal vibrational entropy, S_p^v, evaluated from the quantum-mechanical partition function is found to be 2,006 cal/(mol-°K) or 34.6 cal/(mol-of-residue-°K). This corresponds to a free energy contribution from the configurational entropy of -602 kcal/mol or -10.4 kcal/(mol-of-residue) at 300°K. In a normal mode calculation, which included only dihedral angle degrees of freedom, a value of -320 kcal/mol was obtained (Levitt et al., 1985). This reduced value is due primarily to neglect of contributions from the coupling between dihedral and bond angle degrees of freedom (Karplus and Kushick, 1981; Brady and Karplus, 1985).

That the calculated value of S_p^v, normally set equal to zero by convention, is nearly an order of magnitude larger than the estimates of ΔS_{conf}, raises a paradox that has to be explored; i.e., why does the simple model for ΔS_{conf} that ignores S_p^v agree with the experimental results? The paradox can be resolved by considering the difference between the folded and denatured state in more detail. For a denatured protein, which is assumed to approximate a random-coil polymer, there are two contributions to the configuration entropy (Karplus and Kushick, 1981). The first is that due to the local fluctuations in the neighborhood of a well-defined structure and the second corresponds to the existence of more than one such structure. In standard treatments the former is neglected and only the latter is considered. To include both contributions, we consider a model in which the protein molecule has N conformations, each of which can be treated as a disjoint multidimensional harmonic well. The total configurational entropy, S_{conf}, can be written:

$$S_{conf} = \sum_{I=1}^{N} \omega_I S_I^v - k_B \sum_{I=1}^{N} \omega_I \ln \omega_I, \qquad (1)$$

Dr. Ichiye's present address is Department of Chemistry, University of California, Berkeley, CA 94720.
Dr. Pettitt's present address is Department of Chemistry, University of Houston, Houston, TX 77004.

where ω_I is the Boltzmann weighting factor of the zero of energy of well I, and S_I^v is the vibrational entropy of well I. For the one-dimensional case,

$$S_I^v/k_B = \frac{\theta_v^I/T}{e^{\theta_v^I/T} - 1} - \ln(1 - e^{-\theta_v^I/T}) \quad (2)$$

with $\theta_v^I = h\nu_I/k_B$ and ν_I the vibrational frequency of harmonic well I. To apply Eq. 1 to a protein, the one-dimensional expression given in Eq. 2 has to be replaced by the full vibrational partition function result for a multidimensional oscillator (e.g., in the case of BPTI, the system has 1,740 degrees of freedom).

In Eq. 1, the second term, the so-called "entropy of mixing" is the one usually equated to the total configurational entropy of the denatured state; under the assumption of equal weights ($\omega_I = \omega$), it reduces to the standard expression, $k_B \ln N$, for the entropy. The first term in Eq. 1 represents the Boltzmann weighted sum of the configurational entropies of the individual conformers. The denatured state has contributions from both terms in Eq. 1 with the sum extending over the N allowed conformations. The entropy for the folded protein is also described by Eq. 1. In the heuristic approximation used here, only the first type of term is included for the protein; i.e., it is assumed that there is a single conformation with a residual entropy S_p^v associated with the vibrational degrees of freedom.

The configurational contribution to the denaturation entropy can now be written:

$$\Delta S_{conf} = S_{rc} - S_p^v = \left(\sum_{I=1}^{N} \omega_I S_I^v - k_B \sum_{I=1}^{N} \omega_I \ln \omega_I\right) - S_p^v, \quad (3)$$

where S_{rc} is the entropy of the denatured (random coil) state. Since the calculation for BPTI shows that $S_p^v \gg (-k_B \cdot \Sigma_{I=1}^{N} \omega_I \ln \omega_I)$, the usual approximation for ΔS_{conf} is valid if

$$\sum_{I=1}^{N} \omega_I S_I^v \simeq S_p^v, \quad (4)$$

i.e., if the weighted average of the configurational entropies of the individual denatured conformers is the same as that of the folded protein. To examine the possible validity of Eq. 4, we have calculated the vibrational entropy of single conformers of a number of blocked amino acids with the potential function used for the protein (Brooks et al., 1983). The values of S_I^v were found to range from 19 cal/(mol °K) for Ala to 48 cal/(mol-°K) for Trp with an average of 34 cal/(mol-°K). The exact values depend upon the nature of the blocking group, but the order of magnitude should be valid. A value for comparison is that for decaglycine in a single extended configuration, which yields 28 cal/(mol-of-residue-°K) (Karplus and Kushick, 1981). The effect of solvent on the configurational entropy has been neglected in these calculations; model calculations (Swaminathan et al., 1982; Brady and Karplus, 1985) indicate that the amplitudes of the motions which determine the entropy are insensitive to solvent.

The above results suggest that the fluctuations, and therefore the configurational entropy in a folded protein, are similar to those of a random coil in a single-potential well. This leads to the following argument. The vibrational entropy of a protein is essentially an extensive property; i.e., a protein is large enough so that the $S_p^v \simeq n\langle S_{res}^v \rangle$, where $\langle S_{res}^v \rangle$ is the average over the different amino acids of the configurational entropy for a single conformation (i.e., $\langle S_{res}^v \rangle \simeq 34$ cal/[mol-of-residue-°K]). This conclusion is supported by a comparison of calculations for BPTI and lysozyme (Ichiye, 1985; Levitt et al., 1985). If $\langle S_{res}^v \rangle$ is the same in the native protein and for the individual configurations of the denatured state, Eq. 4 is satisfied and

$$\Delta S_{conf} \simeq -k_B \sum_{I=1}^{N} \omega_I \ln \omega_I \quad (5)$$

in accord with the standard model.

An additional factor to consider is the magnitude of the anharmonic contributions to the entropy of proteins. Analyses of molecular dynamics simulations have demonstrated that the major anharmonic contribution can be ascribed to multiple conformations (Elber and Karplus, 1987; Ichiye and Karplus, 1987). In BPTI and lysozyme, an estimate of the change in entropy due to this effect (e.g., 89 atoms in lysozyme have multiple wells) yields a correction of <2% to the classical entropy. Thus, multiple conformations appear not to be very important for the residual entropy at room temperature. However, near absolute zero (1–2°K), the presence of several minima (e.g., "tunnelling" states) may contribute significantly to the heat capacity and the entropy (Singh et al., 1984).

The conclusion that the $\langle S_{res}^v \rangle$ is essentially the same for a protein in its native conformation and for a single conformer of the denatured polypeptide chain apparently rationalizes the agreement between the standard model calculations described in the first paragraph and the measured denaturation entropies. However, the large magnitude of S_p^v raises the possibility that it may have to be considered explicitly in some cases. This is most likely when apparently small perturbations are made on the protein and a quantitative estimate of the entropy change is required. One such case is concerned with ligand binding (Sturtevant, 1977; Jacrot et al., 1982; Brooks and Karplus, 1983), for which it has been suggested that the distribution of vibrational frequencies is altered in going from the unliganded to the liganded protein. As a model for this effect, the results obtained for BPTI with the frequencies calculated from the normal mode analysis can be compared with those from a set of adjusted frequencies (Brooks and Karplus, 1983); the latter represents a "perturbed" system with higher frequencies that might result when BPTI is bound to trypsin. We consider the vibrational free energy, which changes from −41.5 to −37.7 kcal/mol at 100°K in the presence of the perturbation; at 300°K the

corresponding values are −336.4 and −325.1 kcal/mol, respectively. At all temperatures for an increase in the vibrational frequencies the enthalpy increases while the entropy decreases, leading to a significant destabilizing effect on the system. The change in enthalpy contrasts with that obtained in a previous analysis (Sturtevant, 1977), in which the zero-point contribution was neglected. An example to which the present considerations may be applicable is the enzyme hexokinase, for which inelastic neutron scattering has indicated that there is a change in the low-frequency spectrum on ligand binding (Jacrot et al., 1982).

Another case of interest concerns the effects of single site mutations on the stability of proteins. The most specific data are available for T_4 lysozyme mutants (Hawkes et al., 1984). Several single amino acid substitutions have been studied that leave the crystal structure essentially unchanged, within the accuracy of the x-ray data. The observed changes in the denaturation entropy at 320°K relative to the wild type value [236 cal/(mol-°K)] varied from −59 to +5 cal/(mol°K); changes in the denaturation enthalpy at 320°K ranged from −21 to 0.6 kcal/mol, resulting in some enthalpy-entropy compensation. That such large changes in the denaturation entropy are introduced by altering one residue does not appear unreasonable in view of the residual per residue entropy of 34 cal/(mol°K). Specific calculations are needed, however, to obtain a quantitative explanation of the observed results. In this regard, it is important to note that differences in thermodynamic quantities are being considered; it is possible that the observed results arise from changes in the properties of the native state, the denatured state, or both (Hawkes et al., 1984). An analysis of mutant data for staphylococcus nuclease suggests that the properties of the denatured state play an important role (Shortle and Mieker, 1986).

This work is supported in part by a grant from the National Institutes of Health.

Received for publication 31 March 1987 and in final form 17 May 1987.

REFERENCES

Brady, J., and M. Karplus. 1985. Configuration entropy of the alanine dipeptide in vacuum and in solution: a molecular dynamics study. *J. Am. Chem. Soc.* 107:6103–6105.

Brooks, B. R., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy minimization and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Edholm, O., and H. J. C. Berendsen. 1984. Entropy estimation from simulations of non-diffusive systems. *Mol. Physics.* 51:1011–1028.

Elber, R., and M. Karplus. 1987. Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. *Science (Wash. DC).* 235:318–321.

Hawkes, R., M. G. Grutter, and J. Schellman. 1984. Thermodynamic stability and point mutations of bacteriophage T4 lysozyme. *J. Mol. Biol.* 175:195–212.

Ichige, T. 1985. The Internal Dynamics of Proteins. Ph.D. Thesis. Harvard University, Cambridge, MA.

Ichiye, T., and M. Karplus. 1987. Anisotropy and anharmonicity of atomic fluctuations in proteins: analysis of a molecular dynamics simulation. *Proteins.* In press.

Jacrot, B., S. Cusack, A. J. Dianoux, and D. M. Engelman. 1982. Inelastic neutron scattering analysis of hexokinase dynamics and its modification on binding of glucose. *Nature (Lond.).* 300:84–86.

Karplus, M., and J. N. Kushick. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules.* 14:325–332.

Karplus, M., and J. A. McCammon. 1983. Dynamics of proteins: elements and function. *Annu. Rev. Biochem.* 53:263–300.

Levitt, M., C. Sander, and P. S. Stern. 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease, and lysozyme. *J. Mol. Biol.* 181:423–447.

Levy, R. M., M. Karplus, J. N. Kushick, and D. Perahia. 1984. Evaluation of the configurational entropy for proteins: application to molecular dynamics simulations of an α-helix. *Macromolecules.* 17:1370–1374.

Privalov, P. L. 1979. Stability of proteins. *Adv. Protein Chem.* 33:167.

Shortle, D., and A. K. Meeker. 1986. Mutant forms of staphylococcai nuclease with altered patterns of guanidine hydrochloride and urea denaturation. *Proteins.* 1:81–85.

Schulz, G. E., and H. Schirmer. 1978. Principles of Protein Structure. Springer-Verlag, New York, Inc., New York. Chapter 8.

Singh, G. P., H. J. Schink, L. Lohnysen, F. Parak, and F. Huncklinger. 1984. Excitation in metmyoglobin crystals at low temperatures. *Z. Physik.* B55:23–26.

Sturtevant, J. M. 1977. Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. USA.* 74:2236–2240.

Swaminathan, S., T. Ichiye, W. F. van Gunsteren, and M. Karplus. 1982. Time dependence of atomic fluctuations in proteins: analysis of local and collective motions in bovine pancreatic trypsin inhibitor. *Biochemistry.* 21:5230–5241.