On the entropy of protein folding

GEORGE I. MAKHATADZE AND PETER L. PRIVALOV
Department of Biology and Biccalorimetry Center, The Johns Hopkins University, Baltimore, Maryland 21218
(Received September 6, 1995; Accepted December 28, 1995)

Abstract
The failure to appreciate that the hydration of polar groups is a major contribution to the entropy of protein unfolding has led to considerable underestimates for the loss of configurational freedom when a protein chain folds.

Keywords: configurational entropy; hydration; protein folding; stability

The entropy of protein folding/unfolding is a parameter of principal importance for understanding protein structure, but is subject to much dispute because many controversies remain regarding its estimation. It is clear that the entropy of protein unfolding in aqueous media includes two components: one is associated with the increase of configurational freedom in the polypeptide chain and the other with the hydration of groups that become exposed on unfolding, $\Delta S^{\text{unf}}(T) = \Delta S^{\text{conf}}(T) + \Delta S^{\text{hydr}}(T)$. This defines the configurational entropy of protein unfolding as the entropy in the absence of solvent, i.e., in vacuum. In theoretical studies of protein unfolding, the configurational term, $\Delta S^{\text{conf}}(T)$, is usually assumed to have a value of 15–20 J/K per mol of amino acid residue at 25°C (e.g., Nemer et al., 1966; Bryngelson & Wolynes, 1987; Lee et al., 1994). With regard to hydration entropies, $\Delta S^{\text{hydr}}(T)$, they were deemed significant only for nonpolar groups, $\Delta S^{\text{hydr}}(T)$, the hydration of which gives rise to the hydrophobic force (Kauzmann, 1959; Tanford, 1968; Dill, 1985, 1990). These have been studied in detail using the transfer of model compounds into water and found to be negative at room temperature, to be proportional to the water-accessible surface area (ASA), and to decrease in magnitude with temperature rise, vanishing at about 125°C (e.g., Gill & Wadsö, 1976; Baldwin, 1986; Privalov & Gill, 1988; Privalov & Makhatadze, 1993; Makhatadze & Privalov, 1994, 1995). Therefore, it is usually assumed that, at 120–130°C, hydration effects do not contribute to the entropy of protein unfolding in aqueous media and, at this temperature, the measured values correspond to the configurational entropy of protein unfolding (Baldwin, 1986; Murphy & Freire, 1992; Murphy et al., 1993; 1995; Lee et al., 1994; Xie & Freire, 1994).

Theoretical estimates of configurational entropy at 25°C, $\Delta S^{\text{conf}}(25°C)$, were indeed found to be close to calorimetrically determined entropies of unfolding of globular proteins in aqueous solution at 125°C, $\Delta S^{\text{unf}}(125°C)$ (Privalov, 1979; Privalov & Gill, 1988). To take two examples from many: the entropies of unfolding of barnase and ubiquitin (two single-domain globular proteins without disulfide crosslinks, the unfolding of which is complete and reversible, Fig. 1) are 18 and 14 J/K per mol of amino acid residue, respectively, at 125°C (Griko et al., 1994; Martinez et al., 1994; Wintrode et al., 1994). The correspondence of the entropy values estimated for the two different temperatures, $\Delta S^{\text{conf}}(25°C)$ and $\Delta S^{\text{unf}}(125°C)$, appears reasonable because it is known that the heat capacity increment of protein unfolding in vacuum is very small (Sturtevant, 1977; Privalov & Makhatadze, 1992) and, therefore, the configurational entropy of protein unfolding should not depend significantly on temperature. This correspondence was considered an argument for the correctness of the theory and its assumptions, in particular that the entropy of hydration of polar groups is negligibly small and can be ignored in protein unfolding (Murphy & Freire, 1992; Murphy et al., 1993, 1995; Lee et al., 1994; Spolar & Record, 1994; Xie & Freire, 1994). To draw such a conclusion is, however, not justified, as the analysis below shows.

Consider the entropies of protein unfolding in aqueous solution at 25°C, $\Delta S^{\text{conf}}(25°C)$ (Fig. 1): from calorimetric measurements, it is 8.0 for barnase (Griko et al., 1994; Martinez et al., 1994) and -1.0 for ubiquitin (Wintrode et al., 1994). The difference between the unfolding entropy values at 25°C and 125°C can be written:

$$\Delta S^{\text{conf}}(125°C) - \Delta S^{\text{conf}}(25°C) = \Delta S^{\text{hydr}}(125°C) - \Delta S^{\text{hydr}}(25°C) + \Delta S^{\text{intr}}(125°C) - \Delta S^{\text{intr}}(25°C) + \Delta S^{\text{adm}}(125°C) - \Delta S^{\text{adm}}(25°C).$$

(1)

Because configurational entropy changes are almost temperature independent, $\Delta S^{\text{conf}}(T) = \text{const}$, the difference between the entropies at these two temperatures, can be attributed only to the changes in the hydration effects of buried polar and nonpolar groups. The entropy of hydration of nonpolar groups at
125 °C is close to zero, $\Delta S_{h}^{\text{hydr}}(125 \, ^\circ\text{C}) = 0$, and if one also assumes that the hydration entropy of polar groups is negligible, $\Delta S_{h}^{\text{hydr}}(T) = 0$, then:

$$\Delta S^{\text{exp}}(125 \, ^\circ\text{C}) - \Delta S^{\text{exp}}(25 \, ^\circ\text{C}) = -\Delta S_{h}^{\text{hydr}}(25 \, ^\circ\text{C}).$$

On this basis, the entropy of hydration of nonpolar groups, $\Delta S_{h}^{\text{hydr}}(25 \, ^\circ\text{C})$, would be $-10 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for barnase and $-15 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for ubiquitin at 25 °C. However, the entropies of hydration of nonpolar groups are in precise proportion to the ASA of these groups, i.e., $\Delta S_{h}^{\text{hydr}}(T) = (\Delta S_{\text{ASA}}^{\text{hydr}} - \Delta S_{\text{ASA}}^{\text{hydr}}(25 \, ^\circ\text{C}))$, where $\Delta S_{\text{ASA}}^{\text{hydr}}(T) = -0.578 \, \text{J.K}^{-1} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$ at 25 °C (Privalov & Makhatadze, 1993) and one can therefore determine this entropy of hydration on protein unfolding from a knowledge of the ASA of these groups in the native and unfolded proteins. The ASA of groups in native proteins with known three-dimensional structure is estimated by rolling a probe of the size of a water molecule over the exposed surface (Lee & Richards, 1971; Shriake & Rupley, 1973; Chothia, 1976; Miller et al., 1987). The same method can be used for estimation of the ASA of groups in the unfolded protein by approximating it either to a β-conformation (Ooi et al., 1987; Livingstone et al., 1991) or to a fully extended conformation (Makhatadze & Privalov, 1994, 1995). The approximation of the unfolded state to a fully extended conformation assumes some mutual screening effects by neighboring side chains. Using this approach, we find (Fig. 1) that the entropy of hydration of nonpolar groups on unfolding at 25 °C, $\Delta S_{h}^{\text{hydr}}(25 \, ^\circ\text{C})$, should be $-26 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for barnase and $-29 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for ubiquitin (Makhatadze & Privalov, 1994), i.e., significantly more negative than the values calculated above. The difference, $\Delta S^{\text{hydr}}(125 \, ^\circ\text{C}) - \Delta S^{\text{hydr}}(25 \, ^\circ\text{C})$, of 16 J.K$^{-1} \cdot \text{mol}^{-1}$ for barnase or 14 J.K$^{-1} \cdot \text{mol}^{-1}$ for ubiquitin cannot be explained by an error in the surface-normalized hydration entropies of nonpolar groups because they are known with considerable accuracy from the experiments on model compounds. Neither can it be explained by an error in the determination of the ASA of groups exposed on unfolding of these proteins because this would require only half the increase of ASA of nonpolar groups on unfolding than that found by approximating the unfolded state of the proteins to an extended conformation. An error in this magnitude is highly improbable because the heat capacity increment on protein unfolding, which is mainly caused by hydration of the nonpolar groups, would then also be half that calculated from the determined ASA. But it has been shown for a number of proteins, and, in particular for barnase and ubiquitin, that the calorimetrically measured heat capacity increment on protein unfolding is very close to that calculated from the ASA of nonpolar and polar groups, the deviation usually not exceeding 10% (Griko et al., 1994; Wintrode et al., 1994; Makhatadze & Privalov, 1995). Thus, one or both of the two initial assumptions must be incorrect: (1) that the experimental entropy of protein unfolding at 125 °C is configurational entropy; (2) that the entropy of hydration of polar groups is negligibly small.

Studies of transfer of many model compounds from the gas phase into water have shown that the entropies of hydration of polar groups, $\Delta S_{p}^{\text{hydr}}(T)$, are negative at 25 °C, and become even more negative with increase of temperature, due to the negative heat capacity change upon hydration. This entropy of polar hydration, which can also be scaled with $\Delta S_{\text{ASA}}^{\text{hydr}}$ (Privalov & Makhatadze, 1993; Makhatadze & Privalov, 1994), ranges from $-0.5$ to $-1.0 \, \text{J.K}^{-1} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$, depending on the type of polar surface. Using the buried polar surface areas of barnase and ubiquitin, the entropy of polar hydration at 25 °C, $\Delta S_{p}^{\text{hydr}}(25 \, ^\circ\text{C})$, was estimated (Fig. 1) to be $-22 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for barnase and $-21 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for ubiquitin, respectively, for these proteins (Makhatadze & Privalov, 1994).

If one subtracts the negative entropy of hydration of polar and nonpolar groups from the experimentally determined values (Fig. 1),

$$\Delta S^{\text{exp}}(25 \, ^\circ\text{C}) = \Delta S^{\text{exp}}(25 \, ^\circ\text{C}) - \Delta S_{h}^{\text{hydr}}(25 \, ^\circ\text{C}) - \Delta S_{p}^{\text{hydr}}(25 \, ^\circ\text{C}),$$

the configurational entropy of protein unfolding at 25 °C would be $56 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for barnase and $50 \, \text{J.K}^{-1} \cdot \text{mol}^{-1}$ for ubiquitin, values more than twice those estimated theoretically. Calorimetric studies of many other globular proteins have led to the conclusion that configurational entropies of protein unfolding are between 48 and 62 J.K$^{-1} \cdot \text{mol}^{-1}$, depending on the compactness of the native states, with an average of 52 J.K$^{-1} \cdot \text{mol}^{-1}$ (Makhatadze & Privalov, 1994, 1995). Assuming that the averaged amino acid residue has five torsion angles, this gives about 10 J.K$^{-1} \cdot \text{mol}^{-1}$ torsion$^{-1}$.

The discrepancy between the expected and experimentally estimated entropy values raises the question as to precisely what are the theoretically calculated values that are assumed to represent the configurational entropies of protein unfolding? In particular, are they the entropies of protein unfolding in a vacuum, which are usually considered as representing the configurational entropies?

Analysis of published theoretical estimates of the entropy of protein unfolding show that most are based on enumerating the rotateable bonds (e.g., Schellman, 1955; Nemethy et al., 1966; Finkelstein & Janin, 1989; Creamer & Rose, 1992, 1994; Lee et al., 1994; Stites & Pranata, 1995). To scale the helix properties of different amino acid residues, several groups estimated relative entropy changes for side chains upon the helix-coil transition with the reference to that of Gly and Ala, the entropies of which are set to zero (Creamer & Rose, 1992, 1994; Blaber...
Entropy of protein folding

et al., 1993, 1994; Pickett & Sternberg, 1993; Koehl & Delarue, 1994; Lee et al., 1994). Some of these estimates were based on the definition of entropy as the Boltzmann sampling of states: \( \Delta S = -R \sum \rho_i \ln \rho_i \), where \( \rho_i \) is the probability of the side chain being in state \( i \), obtained through Monte Carlo simulations (Creeper & Rose, 1992, 1994) or the partition function, \( Z \), as \( \exp(-E_i/RT)/Z \) (Lee et al., 1994). The energy term, \( E_i \), was calculated using various force fields and included electrostatic contributions computed with a value of 78 for the dielectric constant. The use of this dielectric constant means that the entropies of side chains are calculated in the presence of water. An alternative approach is to use rotamer libraries derived from known protein structures (Blaber et al., 1993, 1994; Pickett & Sternberg, 1993; Koehl & Delarue, 1994). In this case, the distribution of rotamers is also dictated by the presence of solvent. The entropy changes calculated in these two ways are able quite successfully to describe experimentally observed relative helix propensities of different amino acid residues in aqueous solution (Lyu et al., 1996; O’Neil & DeGrado, 1990; Padmanabhan et al., 1990; Horovitz et al., 1992; Blaber et al., 1994; Chakrabarty et al., 1994). The results of these calculations combined with the earlier estimates for the entropy of the protein backbone by Schellman (1955) or Nemethy et al. (1966) gave the entropy value of about 18 J/K mol per residue for a typical globular protein (Murphy & Freire, 1992; Murphy et al., 1993, 1995). However, this entropy cannot be considered as an entropy of protein unfolding in vacuum, in part because its calculation uses the relative entropy changes for the side chains in water. Recently, Honig and colleagues have separately calculated configurational entropies of protein unfolding in vacuum and the expected hydration effects (Yang & Honig, 1995). The hydration contributions are taken into account using the FDPB/γ solvation model, which includes an electrostatic contribution from the finite difference solutions to the Poisson–Boltzmann equation (Silcock et al., 1994). According to these authors, the configurational entropy of protein unfolding in vacuum is 28 J K⁻¹ mol⁻¹ residue⁻¹ for backbone and, as they found previously (Nicholls et al., 1991), 18 J K⁻¹ mol⁻¹ residue⁻¹ for side chains. In total, this gives 46 J K⁻¹ mol⁻¹ residue⁻¹, which is in a good correspondence with our empirical estimate of 52 J K⁻¹ mol⁻¹ residues.

The hydration and electrostatic effects were also taken into account in another recent calculation of configurational entropy (Wang & Pirisima, 1995; Wang et al., 1995). It was found that the entropic cost of restricting the backbone of an amino acid in a helical matrix is 21 J K⁻¹ mol⁻¹ residue⁻¹. For the configurational entropy of the averaged side chain, Wang et al. (1995) calculate 20 J K⁻¹ mol⁻¹ residue⁻¹, giving the total configurational entropy change upon unfolding as about 41 J K⁻¹ mol⁻¹ residue⁻¹, i.e., 8 J K⁻¹ mol⁻¹ torsion⁻¹ (assuming five torsions in the averaged amino acid residue). This value is comparable to previous estimates based on experimental studies of the cyclization of organic compounds (10.4 J K⁻¹ mol⁻¹ torsion⁻¹ [Page & Jencks, 1971]) or the sublimation of organic compounds (9.1 J K⁻¹ mol⁻¹ torsion⁻¹ [G. I. Makhatadze & P.L. Privalov, unpub. obs.]), on protein–protein interactions (9 J K⁻¹ mol⁻¹ torsion⁻¹ [Novotny et al., 1989]). These can be compared with our empirical estimate of 10 J K⁻¹ mol⁻¹ torsion⁻¹ (see above).

It is clear from the above that the calorimetrically determined entropy of protein unfolding at 125 °C is not just the net configurational entropy of unfolding, but includes the entropy of hydration of exposed polar groups. Assuming the configurational entropy to be essentially independent of temperature, we can estimate the contribution of the hydration of polar groups at 125 °C by comparing the configurational entropy found at 25 °C with the measured entropy of protein unfolding at 125 °C:

\[
\Delta S_{\text{hydr}}^{\text{exp}}(125 \degree C) = \Delta S_{\text{hydr}}^{\text{exp}}(25 \degree C) - \Delta S_{\text{hydr}}^{\text{exp}}(125 \degree C) - \Delta S_{\text{hydr}}^{\text{exp}}(125 \degree C).
\]

Because \( \Delta S_{\text{hydr}}^{\text{exp}}(25 \degree C) = 0 \) and \( \Delta S_{\text{hydr}}^{\text{exp}}(125 \degree C) = \Delta S_{\text{hydr}}^{\text{exp}}(25 \degree C) \), Equation (4) becomes

\[
\Delta S_{\text{hydr}}^{\text{exp}}(125 \degree C) = \Delta S_{\text{hydr}}^{\text{exp}}(25 \degree C) - \Delta S_{\text{hydr}}^{\text{exp}}(25 \degree C).
\]

This difference amounts to -58 J K⁻¹ mol⁻¹ residue⁻¹ for barnase and -35 J K⁻¹ mol⁻¹ residue⁻¹ for ubiquitin and corresponds to the entropy of hydration of polar groups at that temperature. These values are twice that found at 25 °C (Fig. 1) and it is realistic to assume that this difference results mainly from the temperature dependence of the entropy of hydration of polar groups. Analysis of the transfer of polar groups into water indeed shows that their entropy of hydration is negative and increases in magnitude with temperature rise because the heat capacity increment of hydration of polargroups is negative, in contrast to the heat capacity increment of hydration of nonpolar groups (Makhatadze & Privalov, 1990; Privalov & Makhatazde, 1990, 1992). Using data for the transfer of model compounds into water and polar surface area changes upon unfolding, the entropy of hydration of polar groups on unfolding at 125 °C (Fig. 1) is expected to be -31 J K⁻¹ mol⁻¹ residue⁻¹ for barnase and -29 J K⁻¹ mol⁻¹ residue⁻¹ for ubiquitin (Makhatadze & Privalov, 1994, 1995). The remarkable correspondence of these values to those presented above becomes even more striking if one takes into account that the configurational entropy is not, in fact, totally independent of temperature, but slightly decreases with temperature increase (5 J K⁻¹ mol⁻¹ residue⁻¹) on heating to 125 °C (see Fig. 1). This slight decrease results from the fact that the heat capacity of the anhydrous protein is lower in the unfolded state than in the native state (Privalov & Makhatazde, 1992; Makhatadze & Privalov, 1995).

The analysis presented above demonstrates that the magnitude and temperature dependence of the entropy of protein unfolding in aqueous solution cannot be explained by assuming that the hydration of nonpolar groups is the only negative and temperature-dependent contributor to the entropy of protein unfolding. One has to assume the existence of another large component of entropy that becomes increasingly negative with temperature: this can be associated only with the hydration of polar groups. There might still be room for doubt on the exact magnitude of this effect, but it certainly exists and is large and negative. It is clear that the entropy of hydration of polar groups is a particularly important component for understanding protein folding, as well as complexation with other molecules. It provides about 5 kJ per mol of amino acid residue to the Gibbs energy of stabilization of the native protein structure at room temperature, i.e., almost all of what is usually attributed to hydrogen bonding in aqueous media.
Acknowledgments

We thank Colyn Crane-Robinson and Nick Pace for helpful comments and An-Suei Yang and Barry Honig for communicating their manuscript prior to publication. This study was supported by grants from NIH (GM48036-01) and NSF (MCB 9118687).

References


Spolar RS, Record MT. 1994. Coupling of local folding to site specific binding of proteins to DNA. Science 263:777-784.


