Protein Science (1996), 5:2647–2650. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Insights into the local residual entropy of proteins provided by NMR relaxation

ZHIGANG LI,[1] SOUMYA RAYCHAUDHURI, and A. JOSHUA WAND

Departments of Biological Sciences, Biophysical Sciences and Chemistry and the Center for Structural Biology, 816 Natural Sciences Complex, State University of New York at Buffalo, Buffalo, New York 14260

(RECEIVED June 18, 1996; ACCEPTED September 10, 1996)

Abstract: A simple model is used to illustrate the relationship between the dynamics measured by NMR relaxation methods and the local residual entropy of proteins. The expected local dynamic behavior of well-packed extended amino acid side chains are described by employing a one-dimensional vibrator that encapsulates both the spatial and temporal character of the motion. This model is then related to entropy and to the generalized order parameter of the popular "model-free" treatment often used in the analysis of NMR relaxation data. Simulations indicate that order parameters observed for the methyl symmetry axes in, for example, human ubiquitin correspond to significant local entropies. These observations have obvious significance for the issue of the physical basis of protein structure, dynamics, and stability.

Keywords: protein dynamics; protein stability; protein entropy; NMR relaxation

The physical basis of protein structure, dynamics, and stability has been a subject of intense study and debate for several decades. While the taxonomy of protein structure appears to be well understood, a similar depth of understanding of the existence and character of the dynamics of proteins remains to be developed (for a review, see Frauenfelder et al., 1991). These issues are fundamental to a complete understanding of the origins, temporal behavior, and marginal stability of protein structure. For example, empirical evidence indicates that entropic effects dominate the free energy changes that accompany folding of a solvated "random coil" polypeptide usually to a dominant structure of lower free energy. The observed net increase in system entropy upon folding has been attributed to the dominance of a large increase in solvent entropy upon side-chain dehydration over a putatively smaller decrease in side-chain entropy upon packing in the folded globular state. Atomic scale structural analyses of proteins have shown that they generally have extremely high packing densities, which in turn, suggests that proteins are rigid with residual motion being extremely restricted and local. Accordingly, in discussions of protein stability and related issues, it is commonly assumed that the residual entropy of proteins is negligible. Nevertheless, over the past two decades there has been an accumulation of experimental (vide infra) and computational evidence (e.g., Karplus et al., 1987) that suggests quite a different view—that proteins are quite dynamic on time scales relevant to the question of residual entropy (for a recent review, see Doig and Sternberg, 1995).

The magnitude of the residual entropy of proteins is of critical importance to an understanding of protein structure, stability, and ultimately function. In principle, nuclear magnetic resonance spectroscopy can be employed to estimate the local dynamics throughout a protein. Though comprehensive studies of the backbone dynamics of proteins based on analysis of ^{15}N relaxation have appeared, it is only recently that general techniques have emerged to probe the fast dynamics of amino acid side chains. Earlier approaches employing ^{13}C relaxation in selectively ^{13}C enriched samples (e.g., Nicholson et al., 1992) have recently been supplemented by use of methods relying on general random fractional labeling (Wand et al., 1995). In parallel, techniques using deuterium relaxation in both selective (e.g., Tamura et al., 1996) and randomly fractionally deuterated proteins (Muhandiram et al., 1995) have also been developed to investigate the fast dynamics of proteins.

Views of the magnitudes and time scales of internal motions of proteins derived from heteronuclear relaxation studies are often presented using the "model free" treatment of Lipari and Szabo (1982a, 1982b), which describes internal motion in terms of a generalized order parameter and an effective correlation time, τ_e. A recent study of ^{13}C relaxation in randomly fractionally labeled recombinant human ubiquitin reports an unexpected range of generalized order parameters for motions of hydrophobic side chains that constitute the core of the protein and are, in the static view of the structure provided by crystallography (Vijay-Kumar et al., 1987), completely buried. In addition, it was also noted that the dynamics of an individual side chain was context dependent and may not be estimated on the basis of amino acid type alone (Wand et al., 1996).

Reprint Requests to: A.J. Wand, Department of Chemistry, 816 Natural Sciences Complex, State University of New York at Buffalo, Buffalo, New York 14260; e-mail: wand@jasper.chem.buffalo.edu.

[1]Present address: Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

These data have many implications. First and perhaps foremost is the suggestion that the interior of ubiquitin is not only dynamic but heterogeneously so. The presence of potentially spatially extensive dynamics within the core of the protein points to the existence of considerable side-chain entropy. Side-chain conformational entropy is one of many components of the free energy balance that leads to the marginal stability of proteins. Estimates of changes in side-chain conformational entropy upon folding have often been obtained by using highly (or completely) conformationally restricted models for the native state and various degrees of conformational freedom for the unfolded state. Resulting estimates range between 0 and -2 kcal mol^{-1} at 300 K, depending on amino acid type (Doig and Sternberg, 1995).

In a pioneering effort, Palmer and co-workers have described many of the issues underlying the conversion of changes in model-free order parameters to changes in the Gibbs free energy (Akke et al., 1993). Here we will concentrate on the provision of a reasonable model, which encapsulates *both* the temporal and spatial properties of the system with respect to its entropy. To illustrate the connection, the simple one-dimensional harmonic vibrator will be employed to describe the motion of a heteronuclear vector such as the C–H bond of a methine carbon or the symmetry axis of a methyl group (Fig. 1). The potential energy of this system is $V = \frac{1}{2}\mu\omega_0^2(d\theta)^2$ as long as $P(\theta) \to 0$ as $\theta \to \pi$. The solution for this one-dimensional vibrator has energy levels of

$$E_n = (n+\tfrac{1}{2})\hbar\omega_0$$

and a ground state defined by

$$\psi_0 = \frac{\sqrt{\alpha}}{\pi^{1/4}}\, e^{-(1/2)\alpha^2(d\theta)^2} \tag{1}$$

where

$$\alpha = \sqrt{\frac{\mu\omega_0}{\hbar}}$$

and μ and ω_0 are the reduced mass and fundamental frequency of the vibrator. Accordingly, the number of energy levels between the ground state and the level H above the ground state is

$$N \cong \frac{H}{\hbar\omega_0}.$$

If we have an ensemble of these one-dimensional vibrators, the partition function of the system is given by

$$Z = \sum_{n=0}^{N} e^{-\beta(n+1/2)\hbar\omega_0} = \frac{e^{-\beta(1/2)\hbar\omega_0}(1 - e^{-\beta(N+1)\hbar\omega_0})}{1 - e^{-\beta(1/2)\hbar\omega_0}} \tag{2}$$

where

$$\beta = \frac{1}{k_B T}.$$

The entropy is defined as

$$S = Nk_B\left(\ln Z - \beta\frac{\partial}{\partial\beta}\ln Z\right) \tag{3}$$

or

$$S' = R\left(\ln Z - \beta\frac{\partial}{\partial\beta}\ln Z\right) \tag{4}$$

with units of J mol^{-1} K^{-1}. Substituting Equation 2 into 4 yields

$$S' = R\ln\left(\frac{1 - (1 - e^{-(N+1)\hbar\omega_0\beta})}{1 - e^{-\hbar\omega_0\beta}}\right) + \frac{R\beta\hbar\omega_0 e^{-\hbar\omega_0\beta}}{1 - e^{-\hbar\omega_0\beta}} - \frac{R\beta(N+1)\hbar\omega_0 e^{-(N+1)\hbar\omega_0\beta}}{1 - e^{-(N+1)\hbar\omega_0\beta}}. \tag{5}$$

What remains is to relate the probability distribution of the one-dimensional vector to the definition of the generalized order parameter. Assuming a simple distribution of the form:

$$P(\theta) = Ae^{-C\theta^2} \tag{6}$$

with the following properties

$$P(\pi) \approx 0$$

$$\int P(\theta)\, d\Omega = 1.$$

In the context of Equation 1, the constant C becomes

$$\frac{1}{2}\frac{\mu\omega_0}{\hbar}d^2$$

for the ground state and

$$\frac{\mu\omega}{\hbar}\tanh\left(\frac{\hbar\omega}{2kT}\right)d^2$$

when taken over all states of the vibrator. For the corresponding classical vibrator, C becomes

$$\frac{1}{2}\mu\omega_0^2 d^2/kT.$$

The definition of the generalized order parameter

$$\tilde{S}^2 = \sum_{m=-2}^{2}\left|\left\langle\frac{C_{2m}(\Omega)}{r^3}\right\rangle\right|^2 \tag{7}$$

for a vector of constant length r becomes

$$\tilde{S}^2 = \frac{1}{r^6}\sum_{m=-2}^{2}|\langle C_{2m}(\Omega)\rangle|^2. \tag{8}$$

Thus,

$$\langle C_{2m}(\Omega)\rangle = \int_0^{2\pi} d\phi\int_0^{\pi} d\theta\sin\theta P(\theta)C_{2m}(\Omega) \tag{9}$$

and noting that $\langle C_{2\pm1}\rangle = \langle C_{2\pm2}\rangle = 0$ one finds

$$\tilde{S}^2 = \frac{|\langle C_{20}(\Omega)\rangle|^2}{r^6}$$

where

$$\langle C_{20}(\Omega)\rangle = \int_0^{2\pi} d\varphi\int_0^{\pi}\sin\theta\, d\theta\left[\frac{1}{2}(3\cos^2\theta - 1)\right]Ae^{-C\theta^2}. \tag{10}$$

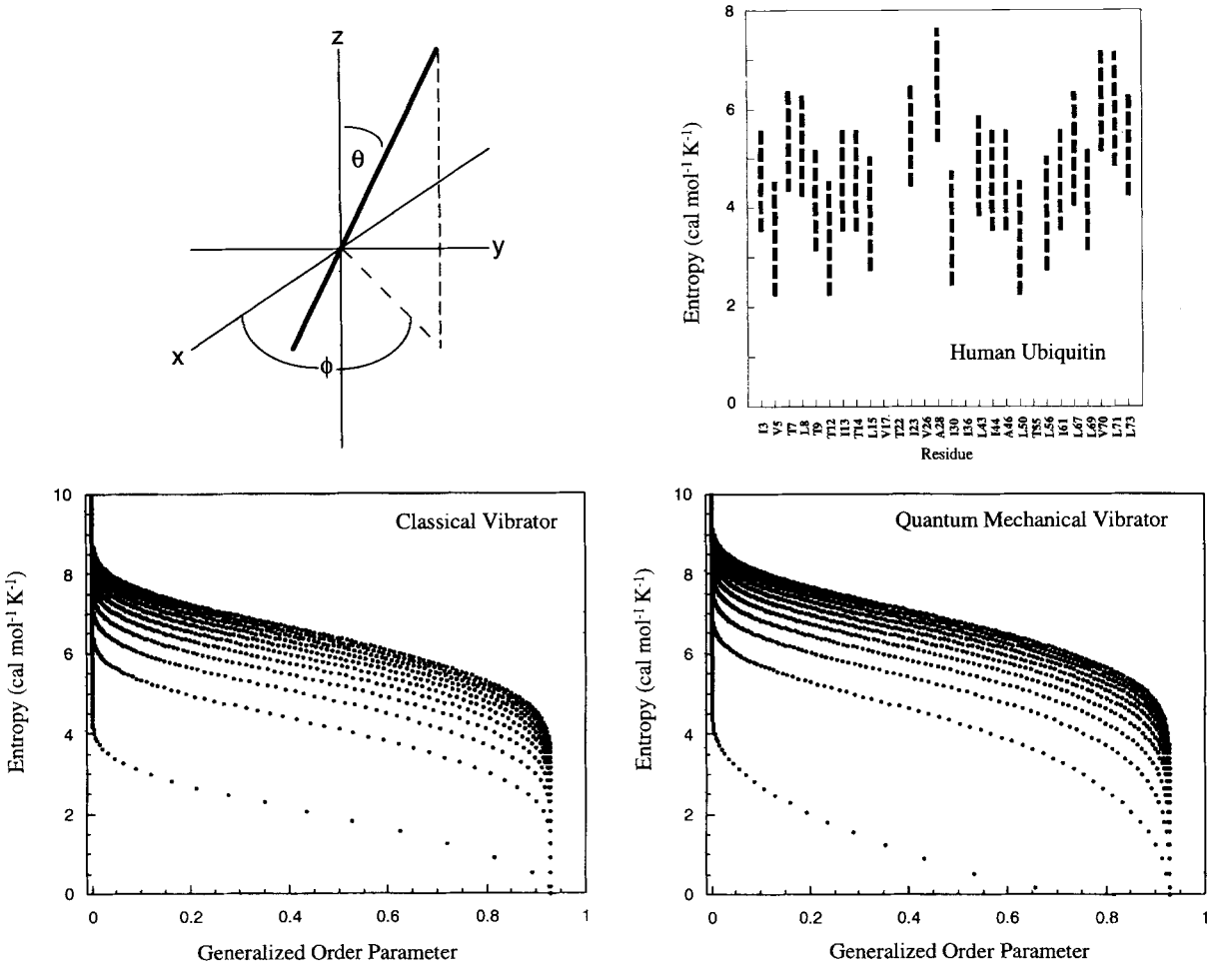

Fig. 1. The model used to describe the dynamic behavior of a C–H vector attached to the end of a long chain amino acid is shown in the top left of the figure. A simple one-dimensional oscillator results by assuming that the behavior of the vibrator is independent of ϕ. The relationship between the entropy of the system and the effective reduced mass (μ) and length of the vibrator (d) is shown for a quantum mechanical vibrator (lower right) and for a classical vibrator (lower left). The effective vector defining the relaxation (e.g., methyl symmetry axis or methine C–H bond) is assumed to be colinear with the oscillator vector. Here a well depth (H) of 1 kcal mol^{-1} and temperature of 300 K have been used. The relationships

$$C = \frac{\mu\omega}{\hbar}\tanh\left(\frac{\hbar\omega}{2kT}\right)d^2 \quad \text{and} \quad C = \frac{1}{2}\mu\omega_0^2 d^2/kT$$

have been used for the quantum mechanical and classical treatments of the vibrator, respectively. A range of fundamental time constants (0.02 ps to 100 ps) has been used to generate the illustrated range of entropies and order parameters using Equations 5 and 10. The lowermost curve of each simulation corresponds to an isolated C–H bond where the oscillator length equals the C–H bond length (r). Subsequent curves correspond to calculations using values of C (see Equation 6) scaled by factors of 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100, respectively, and correspond to systems with increasing effective reduced mass and length. The generalized order parameters obtained for the symmetry axis of methyl groups in human ubiquitin (Wand et al., 1996) are interpreted in terms of the one-dimensional vibrator in the upper right. The parameter C has been scaled by factors of 10 and 100 as above and the entropy determined for a fundamental time constants ranging from 0.02 ps to 100 ps. Uncertainty in C and τ_e and with the uncertainty in the experimental generalized order parameter combine to obscure somewhat the heterogeneous nature of the distribution of dynamics and entropy throughout the protein. Three residues (T22, V26, T55) have order parameters corresponding to no internal motion of the symmetry axis.

There are a number of parameters defining the physical model. C describes the width of the distribution and while H defines the depth of the well. C is proportional to the reduced mass and inversely related to the fundamental time period (τ_0) of the vibrator. The number of levels within the well (N) is defined by the depth of the well (H) and the fundamental frequency (ω_0). Here we will use $H \approx 1$ kcal mol^{-1} (Weber, 1993) and $T = 300$ K. When cast as a one-dimensional system [i.e., $P(\phi)$ = constant], the model resembles, in a spatial sense, the diffusion in the cone model often employed in the analysis of NMR relaxation. Here, however, the model used maintains a direct connection to the temporal character of the motion.

Equation 10 is easily evaluated numerically. The adjustable parameter C encapsulates both the effective reduced mass of the oscillator and its length (see Fig. 1). Figure 1 also illustrates the relationship between the generalized order parameter and funda-

mental frequency of the oscillator for a range of C values. It is important to note that the time scale of side chain motions in human ubiquitin has been found to be on the order of 10^{-12}–10^{-10} seconds for buried side chains. This is consistent with reduced masses and effective oscillator lengths anticipated for large hydrophobic side chains such as isoleucine, leucine, and valine. Figure 1 also shows the connection between the generalized order parameter and the corresponding local entropy in both the quantum mechanical and classical vibrators. These two treatments are essentially equivalent for reduced masses and effective oscillator lengths expected for hydrophobic side chains. One sees immediately that for a wide range of reduced masses and oscillator lengths that the local entropy is decidedly non-zero and significant for order parameters below 0.90. Estimates of order parameters of methyl group symmetry axes (S^2_{axis}) in large hydrophobic side chains of ubiquitin range from ~1 to as low as 0.1 (Wand et al., 1996). This potentially corresponds to a local side chain entropy of up to ~8 cal mol^{-1} K^{-1}. For example, if one interprets the generalized order parameters obtained for methyl groups in human ubiquitin one finds that the methyl containing residues each, on average, contribute about 5 cal mol^{-1} K^{-1} to the residual entropy of the protein for a total of approximately 40 kcal mol^{-1} at 300 K (see Fig. 1). It is also interesting to note that the order parameter is essentially linear with entropy in the range of S^2 values commonly seen for methyl symmetry axes (see Nicholson et al., 1992; Muhandiram et al., 1995; Pascal et al., 1995; Kay et al., 1996; Wand et al., 1996). Finally, notwithstanding the potential uncertainty of the parameter C in the model, the heterogeneous distribution of dynamics seen in the interiors of several proteins corresponds to a significant range of contributions to their residual entropy.

This analysis ignores the obvious influence of conditional probabilities on local entropy in so far as the motions within a protein are presumably correlated to some extent. Nevertheless, the data presented thus far in the literature, when combined with the analysis presented here, clearly indicate that the local residual entropies of proteins are large and arise from states interconverting on the subnanosecond time scale. In addition, the entropic contribution to the stability of the native state of a protein is, of course, the difference between two absolute entropies and the comments made here refer only to the folded state. Indeed, a premise of the derivation presented, that the angular excursions of the vibrator are small, precludes the use of this treatment in the analysis of the highly dynamic unfolded state. It should noted, however, that recursive approaches such as those used to interpret the solvent exposed and highly dynamic C-terminus of ubiquitin may be generally applicable to the unfolded state (Schneider et al., 1992).

Acknowledgments: This work was supported by NIH Grant DK-39806.

References

Akke M, Bruschweiler R, Palmer AG III. 1993. NMR order parameters and free energy: An analytical approach and its application to cooperative Ca^{2+} binding by calbindin D_{9k}. *J Am Chem Soc 115*:9832–9833.

Doig AJ, Sternberg MJE. 1995. Side-chain conformational energy in protein folding. *Protein Sci 4*:2247–2251.

Frauenfelder H, Sligar SG, Wolynes PG. 1991. The energy landscapes and motions of proteins. *Science 254*:1598–1603.

Karplus M, Ichiye T, Pettitt BM. 1987. Configurational entropy of native proteins. *Biophys J 52*:1083–1085.

Kay LE, Muhandiram DR, Farrow NA, Aubin Y, Forman-Kay JD. 1996. Correlation between dynamics and high affinity binding in an SH2 domain interaction. *Biochemistry 35*:361–368.

Lipari G, Szabo A. 1982a. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc 104*:4546–4559.

Lipari G, Szabo A. 1982b. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc 104*:4559–4570.

Muhandiram DR, Yamazaki T, Sykes BD, Kay LE. 1995. *J Am Chem Soc 117*:11536–11544.

Nicholson LK, Kay LE, Baldisseri DM, Arango J, Young, PE, Bax A, Torchia DA. 1992. *Biochemistry 31*:5253–5263.

Pascal SM, Yamazaki T, Singer AU, Kay LE, Forman-Kay JD. 1995. Structural and dynamic characterization of the phosphotyrosine binding region of a Src homology 2 domain–phosphopeptide complex by NMR relaxation, proton exchange, and chemical shift approaches. *Biochemistry 34*:11353–11362.

Schneider DM, Dellwo MJ, Wand AJ. 1992. Fast internal main chain dynamics of human ubiquitin. *Biochemistry 31*:3645–3652.

Tamura A, Matushita M, Naito A, Kojima S, Miura KI, Akasaka K. 1996. Dynamics of the three methionyl side chains of streptomyces subtilisin inhibitor-deuterium NMR studies in solution and in the solid state. *Protein Sci 5*:127–139.

Vijay-Kumar S, Bugg CE, Cook WJ. 1987. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol 194*:531–544.

Wand AJ, Bieber RJ, Urbauer JL, McEvoy RP, Gan Z. 1995. Carbon relaxation in randomly fractionally ^{13}C-enriched proteins. *J Magn Reson Series B*:173–175.

Wand AJ, Urbauer JL, McEvoy RP, Bieber, RJ. 1996. Internal dynamics of human ubiquitin revealed by ^{13}C-relaxation studies of randomly fractionally labeled protein. *Biochemistry 35*:6116–6125.

Weber G. 1993. Thermodynamics of the association and the pressure dissociation of oligomeric proteins. *J Phys Chem 97*:7108–7115.