



An empirical relationship between rotational correlation time and solvent accessible surface area

V.V. Krishnan* and Monique Cosman

Biology and Biotechnology Research Program, L-452 Lawrence Livermore National Laboratory, Livermore, CA 94551, U.S.A.

Received 8 December 1997; Accepted 4 February 1998

Key words: rotational correlation time, solvent accessible surface area, structure and dynamics

Abstract

Structure–dynamics interrelationships are important in understanding protein function. We have explored the empirical relationship between rotational correlation times (τ_c) and the solvent accessible surface areas (SASA) of 75 proteins with known structures. The theoretical correlation between SASA and τ_c through the equation $\text{SASA} = K_r \tau_c^{(2/3)}$ is also considered. SASA was determined from the structure, τ_c^{calc} was determined from diffusion tensor calculations, and τ_c^{expt} was determined from NMR backbone ^{13}C or ^{15}N relaxation rate measurements. The theoretical and experimental values of τ_c correlate with SASA with regression analyses values of K_r as 1696 and 1896 $\text{m}^2\text{s}^{-(2/3)}$, respectively, and with corresponding correlation coefficients of 0.92 and 0.70.

The rotational correlation time (τ_c) of a protein is a measurement of the time the molecule rotates through an angle of one radian, and is dependent on the size, shape, and dynamics of the molecule, as well as the bulk physical characteristics of the solvent (Tanford, 1963; Cantor and Schimmel, 1980; van Holde, 1985; Primrose, 1993). It is important to have a reliable method for estimating τ_c in order to calculate accurate three-dimensional solution structures (Wüthrich, 1986; Güntert, 1997) and derive information on dynamics for a molecule from experimental NMR data (Tycho, 1994; Delpuech, 1995; Palmer et al., 1996). Several physical-chemical techniques, such as fluorescence polarization, dichroism, sedimentation coefficients, and NMR, can be used to estimate the rotational correlation times of biological molecules. (Tanford, 1963; Cantor and Schimmel, 1980; van Holde, 1985). In particular, recent advances in NMR

spectroscopy in studying the dynamics of biomacromolecules at time scales ranging from approximately a picosecond to a second have provided powerful techniques for measuring the global correlation times of proteins, as well as localized intramolecular motions (Nirmala and Wagner, 1988; Kay et al., 1989; Palmer et al., 1991; Jones et al., 1994; Tycho, 1994; Delpuech, 1995; Dayie et al., 1996; Palmer et al., 1996). One way that this can be accomplished is by determining the ratios of the spin-lattice (R_1) and spin-spin (R_2) relaxation rates of the backbone ^{15}N or ^{13}C atoms (Kay et al., 1989; Palmer et al., 1991; Jones et al., 1994; Dayie et al., 1996).

The rotational correlation time of a protein is directly related to its volume and molecular weight (Tanford, 1963; Cantor and Schimmel, 1980). However, it is not always easy to determine the volume of an irregularly shaped protein, and the literature provides many choices as to how to define volume (Paci and Velikson, 1997). Although the molecular weight of a protein with known amino acid sequence can be easily calculated, the effect of molecular weight on τ_c is modulated by the shape and the packing configuration of the molecule (Squire and Himmel, 1979; Eimer

*To whom correspondence should be addressed.

Supplementary Information Available: A table containing τ_c^{calc} , τ_c^{expt} , SASA, corresponding PDB access code and references as well as the related parameters for the set of 75 proteins. An updated list and the correlation coefficients are available as part of the Protein Dynamics Data Base at Lawrence Livermore National Laboratory; <http://www-structure.llnl.gov/NMR/pddb/>.

et al., 1990; Barbato et al., 1992; MacKay et al., 1996). Both volume and molecular weight do not readily take into account effects induced by the shell of water and ions located on the surface of the protein (Venable and Pastor, 1988; Müller, 1991; Smith and van Gunsteren, 1994). Thus, an alternative way to relate the dynamic parameter τ_c to structure is to investigate another measurable parameter of the protein, namely the solvent accessible surface area (SASA) (Lee and Richards, 1971; Richards, 1977).

The advantage of relating the τ_c to the SASA is evident in that it automatically takes into account the molecular shape and hydration shell and is explicitly proportional to the volume and molecular weight as well. In addition, SASA has a unique definition, as proposed by Gerstein and Lynden-Bell (1993), as being the natural boundary of the protein surface with the solvent and the locus of the second-shell water molecules. According to the algorithm of Lee and Richards (1971; Richards, 1977), the SASA can be determined by rolling a sphere, typically having a 'Richards radius' of 1.4 Å similar to that of a water molecule, over the van der Waals surface of the protein. Thus, for a known three-dimensional structure, the SASA can be calculated in a straightforward manner. In addition, the SASA is an important property of the protein that can define how soluble it is in water, how it will interact with other molecules, and how it folds to form a stable active form (Richmond, 1984; Eisenberg and McLachlan, 1986; Creamer et al., 1997). Thus, SASA values by themselves are useful to know in order to characterize a protein as completely as possible. In this communication, the empirical interrelationship between τ_c and the solvent accessible surface area of a protein is examined in an unbiased manner.

In the conventional method of determining τ_c for a hydrated globular protein, an estimated value is given by the Stokes–Einstein–Debye equation (Tanford, 1963; Koenig, 1975; Cantor and Schimmel, 1980):

$$\tau_c = (4\pi\eta r_h^3)\rho/k_B T \quad (1)$$

where η is the solvent viscosity in Nsm^{-2} , k_B is the Boltzmann constant ($1.3806 \times 10^{-23} \text{ m}^2\text{kg s}^{-2}\text{K}^{-1}$), and T is the temperature in K. ρ is the shape factor that relates the frictional coefficient of a sphere to that of an ellipsoid, similar to Perrin's shape factor. Expressions of ρ for a general ellipsoid (Koenig, 1975) as well as for several complex shapes are known (Robert, 1995; Zhou, 1995a,b). The hydrated radius (r_h) of the pro-

tein is given by (Tanford, 1963; Cantor and Schimmel, 1980):

$$r_h = \{[3M_w(V_2 + \delta_1 V_1)]/[4N_0]\}^{1/3} \quad (2)$$

where V_1 and V_2 are the partial specific volumes of the protein and solvent molecules, respectively, in m^3kg^{-1} , δ_1 is the fractional amount of solvent bound to the protein (grams of water per gram of protein), M_w is the molecular weight in kg and N_0 is Avogadro's number ($6.02217 \times 10^{23} \text{ mol}^{-1}$).

Thus, τ_c is inversely proportional to temperature and directly proportional to the molecular weight which can be determined by the amino acid sequence, the volume and shape which can be determined from the three-dimensional structure, and by the amount of water bound, which is estimated to be 0.328 g of water per gram of protein (Tanford, 1963; Squire and Himmel, 1979; Cantor and Schimmel, 1980). Since the shape factor and the amount of water bound are based on estimated values, τ_c is also an estimate and often additional hydration shells need to be considered in order to fit the experimental results (Squire and Himmel, 1979; Barbato et al., 1992; MacKay et al., 1996).

Rearranging Equation (1) in terms of the surface area ($4\pi r^2$) instead of volume ($4\pi r^3/3$) of the hydrated protein yields:

$$SA = K_r(\rho)\tau_c^{2/3} \quad (3)$$

where $K_r(\rho)$ is given by:

$$K_r(\rho) = (1/4\pi)^{1/3} \times (k_B T/\eta\rho)^{2/3} \quad (4)$$

The left-hand side of Equation (3) can be rigorously calculated as the solvent accessible surface area (SASA), without having to produce an equation for a complicated shape factor or determining the viscosity of the solvent. In this calculation, the three dimensional structure of the protein obtained from the Protein Data Bank (PDB; Brookhaven National Laboratory) and the program Quanta (Molecular Simulations Inc.) are used. A sphere with a probe ratio of 1.4 Å and a surface point density of 20 is rolled along the surface of the protein, directly yielding a SASA value.

Calculated values of the rotational correlation times, τ_c^{calc} , can be determined from the rotational diffusion tensor, D , based on the beads model approximation of García de la Torre and Bloomfield (1981; García de la Torre et al., 1994). This method has been used successfully by several groups to calculate translational as well as rotational diffusion tensors of

proteins (Venable and Pastor, 1988; Eimer et al., 1990; Yu-Orekhov et al., 1995; Krishnan, 1997). In this method, the protein is modeled as a collection of point sources of friction (denoted as beads) with hydrodynamic Oseen tensor interactions between them. The rotational diffusion tensor is calculated from a set of linear equations solved by integrating the $3N \times 3N$ matrix, where N is the number of atoms determined from the structure of the protein. The program DIFFC based on the beads theory (Yu-Orekhov et al., 1995) is used in the present work. All backbone atoms are considered as beads of equal size, $\sigma = 5.0 \text{ \AA}$ (García de la Torre and Bloomfield, 1981; García de la Torre et al., 1994), the temperature value equals that at which the experimental relaxation values were determined, and the viscosity (in Nsm^{-2}) of pure water (Viswanth and Natarajan, 1989) was used. In these calculations, the protein can be modeled using exact representations of the atoms instead of beads of equal radii. The former approach is computationally demanding, but allows a detailed description of the surface. When adopted for a subset of six proteins, the τ_c values decreased by less than 5% of the values obtained by using the bead approximation (data not shown). These values were within the experimental error obtained for the data for the complete set of 75 proteins using beads instead of atoms.

The isotropic overall rotational correlation time was calculated from the principal values of the diffusion tensor:

$$\tau_c^{\text{calc}} = (6D_{\text{iso}})^{-1} \quad (5)$$

where D_{iso} is the isotropic value of the diffusion tensor, $(D_{xx} + D_{yy} + D_{zz})/3$.

Experimentally determined τ_c values can be obtained from the published values of R_1 and R_2 relaxation rates for a particular protein. Specifically, the ratio of R_1/R_2 , which is a polynomial of 5th order in τ_c (Carper and Keller, 1997) is solved using numerical algorithms (Kay et al., 1989; Palmer et al., 1991; Carper and Keller, 1997).

Assuming that slight variations in experimental temperatures will not drastically change the SASA or the structure of a protein, the SASA for each protein versus τ_c^{calc} and τ_c^{expt} can then be plotted to determine whether a good correlation between surface area and correlation time can be found. We can then compare $K_r(\rho)^{\text{calc}}$ and $K_r(\rho)^{\text{expt}}$ values derived from τ_c^{calc} and τ_c^{expt} , respectively, using Equation (3) and regression analysis. A comparison of these coefficients yields an impartial consideration as to whether an em-

pirical correlation between SASA and τ_c is a valid hypothesis.

A set of 75 polypeptides and proteins (Table 1, Supplementary material) which fit the following criteria have been selected for this study: (1) most of the molecule is in the folded state; (2) the overall motion of the molecule is dominated largely by isotropic rotational tumbling without significant anisotropic contributions ($D_z/D_{xy} > 2$; D_z and D_{xy} are the calculated rotational diffusion tensors along the major and minor axis respectively); and (3) the molecules are predominantly in a single oligomeric state (i.e., monomer or dimer). Of the 75 proteins, 20 have been determined by X-ray diffraction and the remainder by NMR spectroscopy. The molecular weights, SASA, τ_c^{expt} , and τ_c^{calc} values, as well as the PDB access identification number of the structure for each of the 75 proteins are tabulated in Table 1 (Supplementary material).

Using the computed SASA and τ_c^{calc} values for each protein, $K_r(\rho)^{\text{calc}}$ can be estimated from Equation (3) by either nonlinear least square fit or logarithmic regression analysis (Press et al., 1988). The plot of SASA as a function of τ_c^{calc} is shown in Figure 1A. The solid line indicates the best nonlinear least square fit and yields a value for $K_r(\rho)^{\text{calc}}$ of $2031.8 \pm 26.5 \text{ m}^2\text{s}^{-(2/3)}$ for $\tau_c^{2/3}$ values. The dashed line indicates the best fit to the logarithmic regression analysis and yields a value of $1695.9 \text{ m}^2\text{s}^{-(2/3)}$ with $\tau_c^{0.72}$ values and a correlation factor of 0.92.

Analogously, a plot of SASA and τ_c^{expt} values using Equation (3) is plotted in Figure 1B and yields values of $K_r(\rho)^{\text{expt}}$. The solid line indicates the best fit to the nonlinear least square fit and yields a value for $K_r(\rho)^{\text{expt}}$ of $1775.9 \pm 45.9 \text{ m}^2\text{s}^{-(2/3)}$ for $\tau_c^{2/3}$ values. The dashed line indicates the best fit to the logarithmic regression analysis and yields a value of $1896.5 \text{ m}^2\text{s}^{-(2/3)}$ with $\tau_c^{0.61}$ values and a correlation factor of 0.70. In these analyses, SASA and τ_c are in units of \AA^2 and ns, respectively.

Figure 1 shows that the experimentally determined data points have a wider dispersion and a lower correlation factor than the calculated data points. In addition, although the $K_r(\rho)^{\text{calc}}$ (2032 and 1696) and $K_r(\rho)^{\text{expt}}$ (1776 and 1897) are of comparable magnitude, they do not agree with one another. There are several possible reasons for these differences. One of the assumptions used in the hydrodynamic theory states that the proteins are in infinite dilution and all intermolecular interactions are absent, but this would not be the case in a crystal or in a high-concentration

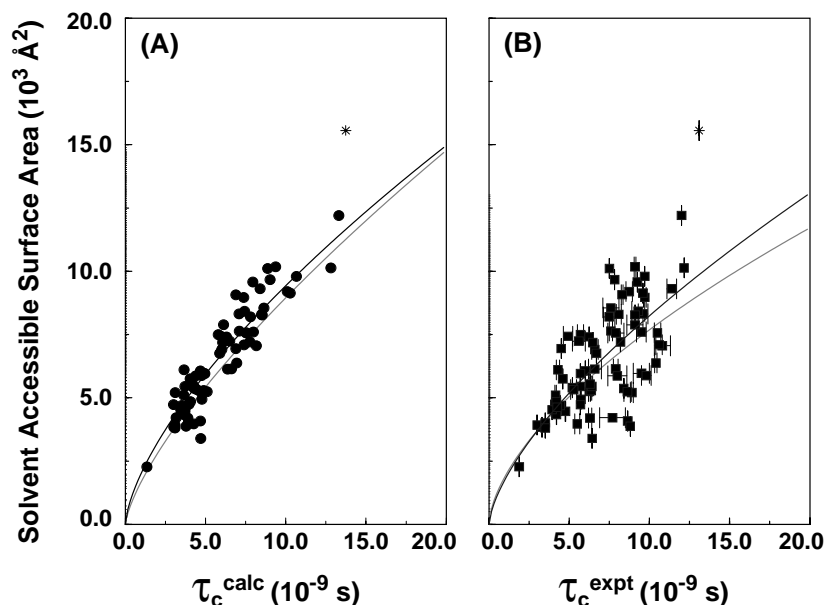


Figure 1. Plots of the solvent accessible surface area (SASA) versus (A) the calculated rotational correlation times (τ_c^{calc}) and (B) the corresponding experimental values (τ_c^{expt}) for a set of 75 proteins. The error bars in τ_c^{expt} indicate the error in the measurement. The nonlinear least square fit and the logarithmic regression analyses are shown as solid and dashed lines, respectively. The symbol * corresponds to the values obtained for the protein phosphotransferase enzyme I (PDB code 2eza), which was not considered in the analyses because of its highly anisotropic motion.

NMR sample. It has also been noted by several groups that R_1/R_2 ratios estimated from trimmed (within a standard deviation of variation) (Kay et al., 1989; Palmer et al., 1991; Jones et al., 1994) values are generally lower than untrimmed values. This indicates that experimental τ_c values are often underestimated. This may provide one reason why $K_r(\rho)^{\text{expt}}$ was $255 \text{ m}^2\text{s}^{-(2/3)}$ units lower than $K_r(\rho)^{\text{calc}}$ in the nonlinear least square analysis using Equation (3). When measurements expressed as a function of magnetic field strengths are used in a simultaneous fitting algorithm, the quality of τ_c^{expt} improves (Buck et al., 1995; Peng and Wagner, 1995; Carper and Keller, 1997; Luginbühl et al., 1997). As additional τ_c^{expt} values of more proteins become available in the literature, the nonlinear fit of Equation (3) can be refined to statistically more significant values of K_r .

Another reason for a difference between the calculated and experimental values may be the effect of hydrophobic versus hydrophilic atoms located on the surface and interacting with the hydration shell. It would be expected that the properties of the hydration layer would be varied on an inhomogeneous surface. Differences in the interactions between the solvent and

the surface would in turn affect the time it takes for the molecule to rotate.

It is also important to remember that the shape of the molecule has an impact on τ_c . For example, a cylindrical-shaped molecule would have an identical volume and surface area if it were to be bent into a donut shape, but would no longer be expected to have the same τ_c . Although a shape factor is included explicitly in Equation (4) for $K_r(\rho)$, it is still an estimated value. The experimentally determined τ_c value would thus be based on the true shape of the protein, while the corresponding calculated value, which is based on the SASA, would not.

Our choice of proteins used in this study was driven by the fact that the majority of proteins for which the structures have been determined met the three criteria outlined above: the molecule is mostly in the folded state, the overall motion of the molecule is dominated by isotropic rotational tumbling, and the molecules are predominantly in a single oligomeric state. For example, experimental relaxation time measurements of lysozyme denatured in trifluoroethanol (Buck et al., 1996) and partially folded basic pancreatic trypsin inhibitor (BPTI) (van Mierlo et al., 1993) have yielded rotational correlation times of 12.2 ± 0.1

and 10.7 ± 1.0 ns, respectively, while the correlation times of the folded states of these proteins are 5.7 ± 0.2 and 3.5 ns, respectively. The PDB structures for the folded forms of lysozyme and BPTI are available and the τ_c values correlate well with the SASA. It would not be possible to determine this correlation for the unfolded states, since the structures are not available. Proteins that undergo highly anisotropic rotation (such as phosphotransferase enzyme I; Tjandra et al., 1997) (Figure 1, data point indicated by *), or multi-domain motions (such as calmodulin, Barbato et al., 1992; Bruschiweiler et al., 1995; Tjandra et al., 1995; Walters et al., 1997) also do not fall within the distribution shown in Figure 1. Tjandra et al. (1997) have shown that for proteins that tumble anisotropically ($D_{\text{major}}/D_{\text{minor}} > 2$), the anisotropic information can be used effectively to refine the structure determination of the protein.

The relationship between structure and dynamics is becoming increasingly important in understanding how proteins function. Without introducing bias into our analysis, we have examined the empirical correlation between the solvent accessible surface area of a protein and the global rotational correlation time. The structural parameter SASA was chosen as the focus of this study because it directly incorporates volume, molecular weight, tertiary packing conformation, and interactions with solvent into one criterion that can influence τ_c . We found a weak correlation of 0.70 in the regression analysis of SASA vs. τ_c^{expt} plot and a stronger correlation of 0.92 for the SASA vs. τ_c^{calc} plot. It is to be expected that the correlation factor for the experimental value of $K_r(\rho)$ will improve significantly as more structures become available in the Protein Data Bank and as measurements of τ_c become increasingly easier and more accurate. However, there are several reasons for why the fit will never be perfect. Proteins that are not predominantly in the folded state, or are dominated largely by anisotropic rotational tumbling, or are not in a single oligomeric state can be easily identified as extreme deviators in the SASA vs. τ_c plot. Researchers can then decide whether specialized refinement protocols are needed for accurate structural determinations. Although the dynamic properties of the smaller domains within a large protein are important in understanding the function of a protein, the average motion estimated by a global τ_c of the whole molecule is often used to calculate the protein structure. By using an iterative method in which the SASA of the initial structure can be used to estimate τ_c^{expt} , one can further refine a structure.

Researchers who study proteins will increasingly require a reliable means of relating function, structure and dynamics to one another. Alternate ways of relating structure and dynamics need to be continuously explored in order to fulfill this need because currently, there are no real methods available in the literature to assess the accuracy of τ_c values used in structure calculations. The method proposed here of estimating τ_c from the SASA of the protein can potentially fill this gap.

Acknowledgements

The authors thank Drs. R. Balhorn, N. Murali and K. Thornton for critical reading of the manuscript. Thanks also go to Professor Arseniev and co-workers for a copy of the program DIFFC. This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract W-7405-ENG-48 and supported by the Lawrence Livermore National Laboratory LDRD Director's Initiative grant 96-DI-010.

References

- Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W. and Bax, A. (1992) *Biochemistry*, **31**, 5269–5278.
- Bruschweiler, R., Liao, X.B. and Wright, P.E. (1995) *Science*, **268**, 886–889.
- Buck, M., Boyd, J., Redfield, C., MacKenzie, D.A., Jeenes, D.J., Archer, D.B., and Dobson, C.M. (1995) *Biochemistry*, **34**, 4041–4055.
- Buck, M., Schwalbe, H. and Dobson, C.M. (1996) *J. Mol. Biol.*, **257**, 669–683.
- Cantor, C.R. and Schimmel, P.R. (1980) *Biophysical Chemistry, part. II, Techniques for the Study of Biological Structure and Function*, W.H. Freeman, San Francisco, CA, pp. 539–590.
- Carper, W.P. and Keller, C.E. (1997) *J. Phys. Chem. Ser A*, **101**, 3246–3250.
- Creamer T.P., Srinivasan, R. and Rose, G.D. (1997) *Biochemistry*, **36**, 2832–2835.
- Dayie, K.T., Wagner, G. and Lefèvre, J.-F. (1996) *Annu. Rev. Phys. Chem.*, **47**, 243–282.
- Delpuech, J. (Ed.) (1995) *Dynamics of Solutions and Fluid Mixtures.*, Wiley, New York, NY, Chapters I, II, and VII.
- Eimer, W., Williamson, J.R., Boxer, S.G. and Pecora, R. (1990) *Biochemistry*, **29**, 799–811.
- Eisenberg, D. and McLachlan, A.D. (1986) *Nature*, **319**, 199–203.
- García de la Torre, J. and Bloomfield, N.A. (1981) *Quart. Rev. Biophys.*, **14**, 81–139.
- García de la Torre, J., Navarro, S., Lopez Martinez, M.C., Diaz, F.G. and Lopez Cascales, J.J. (1994) *Biophys. J.*, **67**, 530–531.
- Gerstein, M. and Lynden-Bell, R.M. (1993) *J. Mol. Biol.*, **230**, 641–650.
- Güntert, P. (1997) in *Protein NMR techniques*, (Ed., Reid, B.R.), Humana Press, NJ, pp. 157–194.

- Jones, D.N.M., Searles, M.A., Shaw, G.L., Churchill, M.E.A., Ner, S.S., Keller, J., Travers, A.A. and Neuhaus, D. (1994) *Structure*, **2**, 609–627.
- Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.
- Koenig, S.H. (1975) *Biopolymers*, **14**, 2421–2423.
- Krishnan, V.V. (1997) *J. Magn. Reson.*, **124**, 468–473.
- Lee, B. and Richards, F.M. (1971) *J. Mol. Biol.*, **55**, 379–400.
- Luginbühl, P., Pervushin, K.V., Iwai, H. and Wüthrich, K. (1997) *Biochemistry*, **36**, 7305–7312.
- MacKay, J.P., Shaw, G.L. and King, G.F. (1996) *Biochemistry*, **35**, 4867–4877.
- Müller, J.J. (1991) *Biopolymers*, **31**, 149–160.
- Nirmala, N.R. and Wagner, G. (1988) *J. Am. Chem. Soc.*, **110**, 7557–7558.
- Paci, E. and Velikson, B. (1997) *Biopolymers*, **41**, 785–797.
- Palmer, A.G., Rance, M. and Wright, P.E. (1991) *J. Am. Chem. Soc.*, **113**, 4371–4380.
- Palmer, A.G., Williams, J. and McDermott, A. (1996) *J. Phys. Chem.*, **100**, 13293–13310.
- Peng, J.W. and Wagner, G. (1995) *Biochemistry*, **34**, 16733–16752.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988) *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York, NY, pp. 498–546.
- Primrose, W.U. (1993) In *NMR of Macromolecules, A Practical Approach* (Ed., Roberts, G.C.K.) Oxford University Press, Oxford, p. 9.
- Richards, F.M. (1977) *Annu. Rev. Biophys. Bioeng.*, **6**, 151–176.
- Richmond, T.J. (1984) *J. Mol. Biol.*, **178**, 2133–2141.
- Robert, C.H. (1995) *Biophys. J.*, **69**, 840–848.
- Smith, P.E. and van Gunsteren, W.F. (1994) *J. Mol. Biol.*, **236**, 629–635.
- Squire, P.G. and Himmel, M.E. (1979) *Arch. Biochem. Biophys.*, **196**, 165–177.
- Tanford, C. (1963) *Physical Chemistry of Macromolecules*, Wiley, New York, NY, pp. 317–456.
- Tjandra, N., Kuboniwa, H., Ren, H. and Bax, A. (1995) *Eur. J. Biochem.*, **230**, 1014–1024.
- Tjandra, N., Garrett, D.S., Gronenborn, A.M., Bax, A. and Clore, G.M. (1997) *Nat. Struct. Biol.*, **4**, 443–449.
- Tycko, R. (Ed.) (1994) *Nuclear Magnetic Resonance Probes of Molecular Dynamics*, Kluwer Academic Publishers, Dordrecht.
- van Holde, K.E., (1985) *Physical Chemistry*, 2nd ed., Prentice-Hall Inc., Englewood Cliffs, NJ.
- van Mierlo, C.P.M., Darby, N.J., Keller, J., Neuhaus, D. and Creighton, T.E. (1993) *J. Mol. Biol.*, **229**, 1125–1146.
- Venable, R.M. and Pastor, R.W. (1988) *Biopolymers*, **27**, 1001–1014.
- Viswanath, D.S. and Natarajan, G. (1989) *Data Book on the Viscosity of Liquids*, Hemisphere, New York, NY.
- Walters, K.J., Dayie, K.T., Reece, R.J., Ptashne, M. and Wagner, G. (1997) *Nat. Struct. Biol.*, **4**, 744–750.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.
- Yu-Orekhov, V., Nolde, D.E., Golovanov, A.P., Dorzhnev, D.M. and Arseniev, A.S. (1995) *Appl. Magn. Reson.*, **9**, 581–588.
- Zhou, H.-X. (1995a) *Biophys. J.*, **69**, 2286–2297.
- Zhou, H.-X. (1995b) *Biophys. J.*, **69**, 2298–2303.