Journal of Biomolecular NMR, 10 (1997) 199–203
KLUWER/ESCOM

J-Bio NMR 482

Characterisation of protein unfolding by NMR diffusion measurements

Jonathan A. Jones, Deborah K. Wilkins, Lorna J. Smith and Christopher M. Dobson*

Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, U.K.

Received 30 June 1997
Accepted 31 July 1997

Keywords: Protein folding; Non-native states; Hydrodynamic radius; Lysozyme

Summary

The characterisation of non-native states of proteins is a key problem in studies of protein folding. Complete characterisation of these states requires a description of both local and global properties, including molecular dimensions. Here we present results from pulsed field gradient experiments designed to compare the effective hydrodynamic radii of a protein in native and non-native states. Measurements performed on lysozyme indicate that the effective hydrodynamic radius increases by 38±1% on unfolding in urea, a result completely consistent with a recent study by small-angle X-ray scattering.

Protein folding involves the conversion of an unfolded polypeptide chain to a compact native state (Dobson and Ptitsyn, 1997). While the structures of many native state proteins are well defined, the characterisation of non-native states remains a key problem in studies of protein folding. The study of such states is challenging as they inhabit a rapidly interconverting conformational ensemble, and a complete characterisation requires descriptions of both their local and global properties, together with an understanding of how these are related (Smith et al., 1996). While the recent application of heteronuclear multidimensional NMR techniques has led to considerable advances in our understanding of the local properties of these states (Shortle, 1996), another important aspect is the measurement of molecular dimensions, described in the simplest case by an effective radius. Some values for the effective radii of unfolded proteins have been measured by small-angle X-ray scattering (SAXS) using synchrotron radiation (Lattman, 1994), but these experiments are difficult to perform, and the high radiation intensity can result in substantial sample degradation. An attractive alternative is to measure the diffusion coefficient of the protein, as this is inversely proportional to the effective hydrodynamic radius. Furthermore, for rapidly interconverting species the observed diffusion coefficient is a time average over the ensemble. Here we demonstrate the use of gradient NMR techniques to compare the hydrodynamic radii of native and non-native states of lysozyme.

The diffusion coefficient of a spherical body in a continuous fluid at temperature T is given by the Stokes–Einstein equation

$$D = \frac{kT}{6\pi a\eta} \tag{1}$$

where a is the radius of the body and η is the viscosity of the fluid. This equation can also be applied to more complex systems if a is replaced by the effective hydrodynamic radius, R_H. (The effective hydrodynamic radius may be defined as the radius of a sphere with the same diffusion coefficient; for simple shapes, such as ellipsoids, the effective radius can be calculated from the molecular dimensions.) Clearly, absolute values of D can only be interpreted if the temperature and viscosity of the solution are known (Gast et al., 1997), and these values may be inconvenient to obtain. Instead, we choose to use another molecule as an internal radius standard (the use of a small molecule as an internal viscosity standard has been reported previously (Chen et al., 1995a)). For protein and reference molecules in the same solution,

$$R_H^{protein} = \frac{D_{ref}}{D_{protein}} \times R_H^{ref} \tag{2}$$

*To whom correspondence should be addressed.

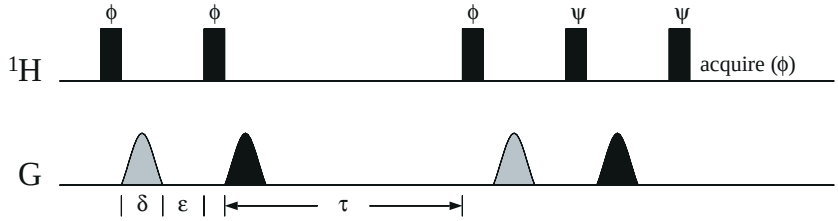

Fig. 1. The PG-SLED sequence. The phase cycle (ϕ = x,x,y,y,–x,–x,–y,–y; ψ = x,–x,y,–y,–x,x,–y,y) combines the standard CYCLOPS cycle, to remove acquisition artefacts, with phase alternation of the longitudinal storage sequence, which causes spin–lattice relaxation to appear as an exponential decay (Sklenář al., 1987). The diffusion labelling gradients (shown in grey) were varied, while the crush gradients (shown in black) were applied at full strength. All gradient pulses were shaped as sine waves, and the two crush gradients were applied perpendicular to one another and to the diffusion gradients. For the experiments in this paper, δ = 6.3 ms, ε = 0.7 ms and τ = 100 ms.

and so it should be possible to obtain absolute values of protein radii once the system has been calibrated. We have applied this approach to the characterisation of the native and urea denatured states of lysozyme, using dioxane as a radius standard. Our measurements indicate that the hydrodynamic radius of lysozyme increases by

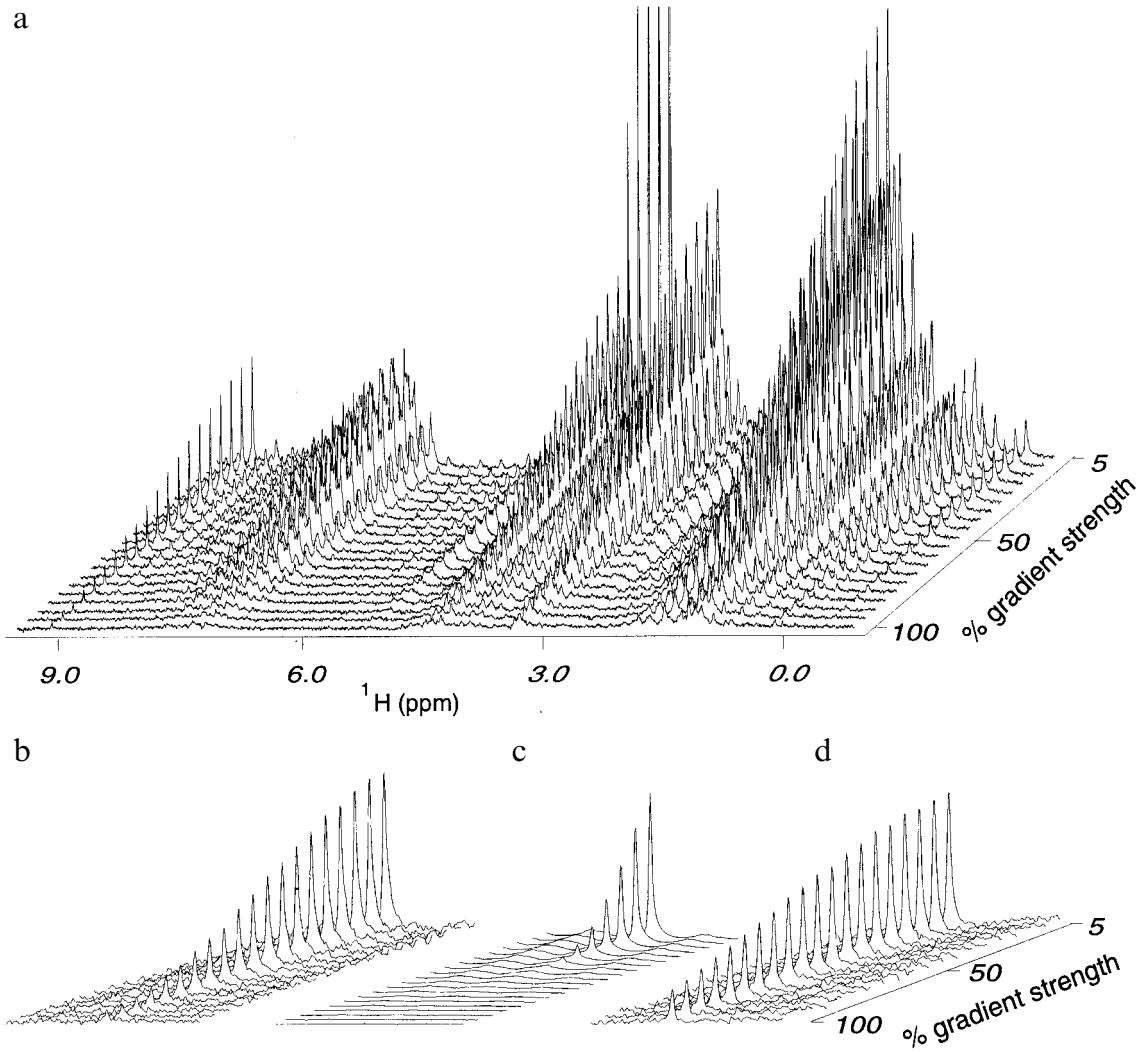

Fig. 2. PG-SLED spectra obtained from solutions of HEWL. (a) The entire spectrum from HEWL in D_2O. The residual signal from HOD was suppressed by saturating the resonance during the relaxation and τ periods, while the large signal from dioxane near 3.7 ppm has been truncated. The remaining signals arise from HEWL and all decay at the same rate. The diffusion gradients were varied between 5% and 100% of their maximum strength (approximately 60 G cm^{-1}). The total acquisition time was about 1 h. (b) Expansion of the aromatic signal from His^{15} near 9.0 ppm. (c) Expansion of the dioxane signal; note that this signal decays much more rapidly than the signals from HEWL. (d) Expansion of the signal from His^{15} taken from the spectrum of HEWL in 8 M urea; the signal decays much more slowly as a result of the lower diffusion coefficient.

38±1% on unfolding, in complete agreement with recent SAXS studies (Chen et al., 1996).

NMR has been used to study diffusion for many years. The original pulse gradient spin echo (PGSE) sequence (Stejskal and Tanner, 1965) is well suited to small molecules, and has been extensively applied to the measurement of self-diffusion coefficients in pure liquids and gases. PGSE is not, however, appropriate for larger molecules such as proteins in which the spin–spin relaxation time (T_2) is much shorter than the spin–lattice relaxation time (T_1). In this case it is better to replace the simple spin echo with a two-pulse stimulated echo, in which the magnetisation is stored along the z-axis during the diffusion period. Better still is to use the PG-SLED sequence (Gibbs and Johnson, 1991), shown in Fig. 1. This sequence, which incorporates a 'longitudinal echo' period before detection, was originally devised to allow diffusion measurements to be performed using gradient systems with long recovery times. With modern systems, incorporating actively shielded gradient coils, shaped pulses and gradient preemphasis, recovery times are less than 1 ms and this is not a serious problem. However, this sequence, with the addition of a crush gradient pulse within the longitudinal echo, can also be used to remove antiphase signals arising from J-modulation due to homonuclear scalar couplings, which are not refocused by the spin echo. It might seem that spin–lattice relaxation during the final longitudinal echo period would distort measurements, but this can be avoided by alternately storing magnetisation along the ±z-axes, in which case spin–lattice relaxation simply results in signal attenuation (Sklenář al., 1987; Gibbs and Johnson, 1991). This sequence has been used recently to study protein aggregation (Altieri et al., 1995; Lin and Larive, 1995). The closely related diffusion ordered spectroscopy (DOSY) family of experiments (Morris and Johnson, 1992; Wu et al., 1996) has also been used to study a wide variety of problems, including characterising the molecular weight distribution of a disperse homopolymer (Chen et al., 1995b).

The signal intensity observed in a PG-SLED sequence depends in a complex manner on T_1, T_2 and the diffusion coefficient, D. If, however, all delays are held constant, and only the gradient strength, g, is varied, then the signal intensity depends only on g and D:

$$s(g) = Ae^{-dg^2} \qquad (3)$$

where the observed decay rate, d, is proportional to D. As discussed above, absolute values of D are not necessary for our purposes, and it is sufficient to obtain a value of d.

Figure 2 shows PG-SLED spectra obtained from a 1.4 mM solution of hen egg white lysozyme (HEWL) in D_2O at pH 2.0 (uncorrected meter reading) and 20 °C; a small amount of 1,4-dioxane was added as the radius standard. All NMR experiments were carried out on a home-built NMR spectrometer with a proton frequency of 600 MHz. The loss of intensity of the HEWL resonances due to diffusion is clear. The diffusion coefficient for the protein was obtained by integrating the aromatic region (6.5–8.5 ppm), and fitting the resulting intensities to Eq. 3. It is not possible, however, to use this simple method for the dioxane resonance as it overlies signals from HEWL. Instead, the region around the dioxane peak was integrated and the resulting intensities were fitted to the sum of two decaying Gaussian terms, one of which was constrained to have the same decay rate as that observed for the aromatic resonances of HEWL. The decay rate of the second Gaussian was then taken as an estimate of d for dioxane. In both cases the experimental data fit the theoretical curves extremely well, and the residuals show no evidence of any additional structure.

The experiment described above was repeated for a series of HEWL solutions at a pH value of 2.0 containing varying concentrations of urea. In these solutions HEWL will exist as a mixture of native and denatured forms, with different diffusion coefficients, and so the observed signal might be described by the sum of two Gaussians, corresponding to the native and denatured states. In practice, however, it is not possible to distinguish such a bi-Gaussian decay from a single Gaussian unless the two decay rates are very different (as is the case for HEWL and dioxane) or the signal-to-noise ratio is very high; this is related to the well-known difficulty of characterising bi-

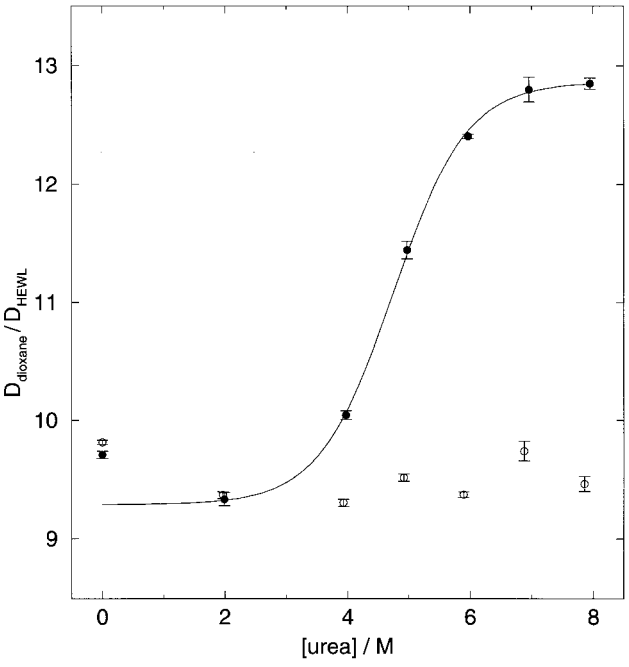

Fig. 3. Diffusion coefficients for 1.4 mM HEWL solutions as a function of urea concentration. The filled circles and the solid line show data at pH 2.0, while open circles show data at pH 5.5 (uncorrected meter readings). Data points are shown as mean±standard error; the solid line was fitted to the data at pH 2.0 as described in the text.

exponential relaxation (Jones et al., 1996). Instead, we choose to fit the decay to a single Gaussian, resulting in an apparent decay rate which is a weighted average of the different decay rates contributing to the signal. Hence, changes in apparent diffusion coefficients can be used to characterise changes in the relative populations of different folding states. Note that it is important to integrate the entire aromatic region, thus ensuring that signals from both folded and unfolded states are fully represented (the aromatic region is used in preference to the aliphatic region as many common contaminants in protein solutions give rise to resonances in the aliphatic region of the spectrum).

The observed diffusion coefficients of both dioxane and HEWL decreased with increasing urea concentration, as a result of the increased viscosity, but the ratio $D_{dioxane}/D_{HEWL}$ increased, as shown in Fig. 3. In each case the experiment was repeated several times, and the figure shows the mean and standard error for each set of observations. Clearly, the apparent hydrodynamic radius of HEWL increases substantially at high concentrations of urea, consistent with protein unfolding. The solid line shows the results of fitting the data for urea concentrations between 2 and 8 M to a simplified model. This model assumes that unfolding is a two-state process, and that the free energy of unfolding decreases linearly with increasing urea concentration (Pace, 1986). In fact, the unfolding of HEWL by urea may involve at least three states (Chen et al., 1996). Despite these simplifying assumptions, the agreement between the data and the model is remarkably good. The initial decrease in apparent radius between 0 and 2 M urea may simply indicate that HEWL undergoes slight aggregation in pure water; similar results were seen in SAXS measurements (Chen et al., 1996). Values of $D_{dioxane}/D_{HEWL}$ were obtained for the folded and unfolded states from the extreme values of the fitted model, and their ratio indicates that the effective hydrodynamic radius increases by 38±1% on unfolding (the error was estimated by Cramér–Rao theory (van den Bos, 1982; Jones et al., 1996) using parameters from the fitted model and experimental error values). The experiment was then repeated with HEWL solutions at pH 5.5, under which conditions the folded state is stable even in 8 M urea (Steiner, 1964). In this case the initial decrease in apparent radius was again observed, but no subsequent increase was seen. This further confirms our interpretation that the increase in $D_{dioxane}/D_{HEWL}$ observed at pH 2 is due to unfolding. The substantial scatter observed in these measurements may be a result of aggregation, which is known to be more serious for HEWL at higher pH values.

In order to investigate the possibility of aggregation, the experiments were repeated for a variety of protein concentrations (0.4–1.4 mM) with urea concentrations of 0, 2 and 8 M at pH 2.0. The measured diffusion coefficients in 2 and 8 M urea solutions were consistent with one another, and showed no evidence of any dependence on protein concentration. Measurements in the absence of urea indicated that the diffusion coefficient was reduced in solutions with very high protein concentrations (7 mM), but the measurements were extremely variable, even for two samples with nominally identical conditions. This suggests that the extent of aggregation under these conditions depends strongly on details of sample preparation and storage, as well as on protein concentration.

The apparent increase in the hydrodynamic radius of HEWL on unfolding in urea (38±1%) is identical to the increase in radius observed by SAXS (Chen et al., 1996). This suggests that the two experiments are measuring the molecular dimensions in a similar way, even though diffusion is sensitive to the time-averaged dimensions while SAXS determines an ensemble averaged radius (Smith et al., 1996). The SAXS radii measured for HEWL were 16 Å for the folded state and 22 Å for the unfolded state; using Eq. 2 this gives a fitted hydrodynamic radius (R_H) of 1.7 Å for dioxane. This is similar to, but slightly larger than, the radius of gyration (1.5 Å) calculated from its crystal structure (Buschmann, 1986); this may reflect the consequences of a hydration shell around the dioxane molecule.

The analysis above rests on the assumption that the effective hydrodynamic radius of the radius standard (dioxane) is unchanged by the addition of urea. This might seem unlikely as the effective hydrodynamic radius will include the effects of any hydration shells, and urea is believed to affect the hydrogen bonding potential of water (Arêas et al., 1995; Plaxco et al., 1997). Furthermore, the method also assumes that dioxane does not interact with HEWL, even in its unfolded state. These assumptions were investigated by repeating the measurements in 2 and 8 M urea at pH 2.0 with a mixture of five different reference molecules (alanine, acetic acid, dioxane, dimethylsulphoxide (DMSO) and 2,2-dimethyl-2-silapentane-5-sulphonate (DSS)) in the same solution. The diffusion rates of these reference molecules were also studied in pure D_2O, with neither urea nor protein present. The results from the three species which are uncharged at pH 2.0 were very similar: dioxane, acetic acid and DMSO all gave apparent increases in the hydrodynamic radius of HEWL which were statistically indistinguishable from that measured using dioxane alone. Furthermore, the relative diffusion rates of these molecules were the same in all three solutions. These results strongly suggest that these three species do not interact significantly with either HEWL or urea, and so they are appropriate reference molecules. The results for the two charged species, however, were significantly different: DSS gave an apparent increase of 52±2%, while alanine indicated an increase of only 29±1%, and in both cases the relative diffusion rates (compared to the three species above) were significantly

different in the three solutions, indicating that DSS and alanine interact either with HEWL or with urea. Thus, these two species are poor choices as reference molecules.

In conclusion, we have demonstrated that NMR diffusion measurements may be used to detect and characterise protein folding transitions through measurements of effective hydrodynamic radii. The use of a radius standard greatly simplifies the application of the method to systems of varying viscosity, such as denaturant solutions, and the results obtained for HEWL are completely consistent with those obtained by small-angle X-ray scattering. It may be possible to use the method to determine absolute values for molecular sizes. In conjunction with rapid mixing techniques (Balbach et al., 1995) it should be possible to extend this experiment to allow the kinetics of slow folding and unfolding transitions to be studied in real time.

Acknowledgements

We thank J. Balbach, N. Soffe, A. Miranker, S. Grimshaw and H. Lu for helpful discussions. This is a contribution from the Oxford Centre for Molecular Sciences, which is supported by the U.K. EPSRC, BBSRC and MRC. J.A.J. thanks Merton College, Oxford, for a Junior Research Fellowship. L.J.S. is a Royal Society University Research Fellow. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Foundation.

References

Altieri, A.S., Hinton, D.P. and Byrd, R.A. (1995) *J. Am. Chem. Soc.*, **117**, 7566–7567.

Arêas, E.P.G., Mauro, C.C. and Santos, P.S. (1995) *J. Mol. Struct.*, **378**, 111–119.

Balbach, J., Forge, V., Van Nuland, N.A.J., Winder, S., Hore, P.J. and Dobson, C.M. (1995) *Nat. Struct. Biol.*, **2**, 865–870.

Buschmann, J., Muller, E. and Luger, P. (1986) *Acta Crystallogr.*, **C42**, 873–876.

Chen, A.D., Wu, D.H. and Johnson Jr., C.S. (1995a) *J. Phys. Chem.*, **99**, 828–834.

Chen, A., Wu, D. and Johnson Jr., C.S. (1995b) *J. Am. Chem. Soc.*, **117**, 7965–7970.

Chen, L., Hodgson, K.O. and Doniach, S. (1996) *J. Mol. Biol.*, **261**, 658–671.

Dobson, C.M. and Ptitsyn, O.B. (1997) *Curr. Opin. Struct. Biol.*, **7**, 1–2.

Gast, K., Nöppert, A., Müller-Frohne, M., Zirwer, D. and Damaschun, G. (1997) *Eur. Biophys. J.*, **25**, 211–219.

Gibbs, S.J. and Johnson Jr., C.S. (1991) *J. Magn. Reson.*, **93**, 395–402.

Jones, J.A., Hodgkinson, P.J., Barker, A.L. and Hore, P.J. (1996) *J. Magn. Reson.*, **B113**, 25–34.

Lattman, E.E. (1994) *Curr. Opin. Struct. Biol.*, **4**, 87–92.

Lin, M. and Larive, C.K. (1995) *Anal. Biochem.*, **229**, 214–220.

Morris, K.F. and Johnson Jr., C.S. (1992) *J. Am. Chem. Soc.*, **114**, 3139–3141.

Pace, C.N. (1986) *Methods Enzymol.*, **131**, 266–280.

Plaxco, K.W., Morton, C.J., Grimshaw, S.B., Jones, J.A., Pitkeathly, M., Campbell, I.D. and Dobson, C.M. (1997) *J. Biomol. NMR*, in press.

Shortle, D.R. (1996) *Curr. Opin. Struct. Biol.*, **6**, 24–30.

Sklenář, V., Torchia, D. and Bax, A. (1987) *J. Magn. Reson.*, **73**, 375–379.

Smith, L.J., Fiebig, K.M., Schwalbe, H. and Dobson, C.M. (1996) *Folding Design*, **1**, R95–R106.

Steiner, R.F. (1964) *Biochim. Biophys. Acta*, **79**, 51–63.

Stejskal, E.O. and Tanner, J.E. (1965) *J. Chem. Phys.*, **42**, 288–292.

Van den Bos, A. (1982) In *Handbook of Measurement Science*, Vol. 2 (Ed., Sydenham, P.H.), Wiley, New York, NY, U.S.A., pp. 331–377.

Wu, D., Chen, A. and Johnson Jr., C.S. (1996) *J. Magn. Reson.*, **A123**, 215–218.