Proc. Int. Symposium on Advances in Neutron Scattering Research, Tokai, 2000
J. Phys. Soc. Jpn. **70** (2001) Suppl. A pp. 392–395

Protein Folding and Dynamics - New Insights from Computer Simulation and Scattering Experiments

Nicoleta BONDAR[1], Roy DANIEL[2], John L. FINNEY[3], Stefan FISCHER[1], Mikio KATAOKA[4], Andrei PETRESCU[5] and Jeremy C. SMITH[1]

[1] *Lehrstuhl für Biocomputing, IWR, Universität Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany*
[2] *Department of Biochemistry, University of Waikato, Hamilton, New Zealand*
[3] *Department of Physics and Astronomy, University College London, Gower Street, London WCIE, 6BT, England*
[4] *Laboratory of Bioenergetics and Biophysics, Nara Institute of Science and Technology (NAIST), Nara, Japan*
[5] *Institute of Biochemistry of the Romanian Academy, 77700 Bucharest 17, Romania*

Neutron and X-ray scattering, when combined with computer simulation, are useful probes of protein structural dynamics. Here we review some recent work using this approach to probe motions and configurational distributions in native (folded) and denatured (unfolded) proteins. Some new results on high-resolution protein vibrational spectroscopy are also presented.

KEYWORDS: protein folding and dynamics, dynamical transition

§1. Introduction

In the post-genomic era attention is being turned to 'proteomics' i.e. an understanding of the products of genes (protein sequences) and how they determine biological function. The information present in a protein sequence is translated into three-dimensional protein structure. The structure itself leads to function of the protein. However, it is now realised that to understand how proteins fold up into functional architectures and how this architecture functions requires that knowledge of structures of folded (native) and unfolded (denatured) proteins be complemented with information on their internal dynamics. Here we describe recent work on the dynamics and configurational distribution of native and denatured proteins, obtained by combining neutron and X-ray scattering with molecular simulation.

§2. Protein Folding

Protein folding kinetics and thermodynamics is determined by the configurational statistics of proteins at various stages in folding. To understand this requires the development and combination of experimental and theoretical tools for characterising the large number of possible configurations that can coexist in any given state along the folding pathway.

Information on both local (short-range) and nonlocal (long-range) interactions is required so as to obtain a description of the accessible configurational space of the chain. Moreover, it is also important to examine possible interdependence of the nonlocal and local configurational properties. To do this techniques giving access to appropriate statistical configuration distribution functions are required. In this regard solution scattering of neutrons and X-rays is particularly promising, as the scattered intensity is determined by pair distance distribution functions of the system under study.

Small-angle scattering is commonly used to determine the size (radius of gyration) of biological macromolecules. Although this is valuable information for denatured states,[1,2] there exists the additional possibility that this technique can be extended to provide a more detailed picture, involving the determination of the chain flexibility at higher resolution. With this aim we have pursued a program to characterise the configurational statistics of a strongly-denatured protein.[3–8] The system investigated is phosphoglycerate kinase (PGK) in 4M guanidinium chloride, chosen as a model of an unfolded protein without disulphide bonds. Small-angle neutron scattering (SANS) was chosen over its X-ray counterpart as it allows the difference in sign of the scattering lengths of hydrogen and deuterium to be exploited so as to optimise the contrast between the protein and solvent at high concentrations of denaturant.

The radius of gyration of the denatured PGK, extrapolated to zero protein concentration, was found to be 90.0 ± 4.5 Å, largely increased from that of the native state (23.8 ± 0.3 Å). At higher resolution, to examine the configurational statistics it is necessary to construct theoretical models of the protein chain, to calculate their scattering properties, and compare with experiment. In this way it was shown that the SANS data can be reproduced by the scattering of a random flexible chain of spheres of excess scattering density, with 5 residues per sphere, subject to excluded volume interactions.[7] This coarse-grained modelling quantifies the increase on unfolding of the long-range ($> \sim$20Å) chain flexibility.

Clearly, the ultimate aim of experiments on unfolded states is to obtain an atomic-detail description of the configurational distribution. Small-angle scattering experiments are limited in this endeavour insofar as they yield principally low-resolution, spherically-averaged information. However, the possibility does exist that the effects of statistical local conformational preferences might

propagate into the small-angle region. This was examined in recent work, using atomistic models of unfolded PGK.[8] Atomic-detail ensembles of the unfolded protein chain were modelled and their scattering profiles compared with experiment. Three of these models are shown in Fig. 1. The local conformational statistics were found to strongly influence the experimental intensity at scattering vectors between 0.05 and 0.3 $Å^{-1}$. Denaturation leads to a reduction in the protein atom-pair distance distribution function over the $\sim$ 3-15 Å region that is associated with a quantifiable shift in the backbone torsional angle (ϕ,ψ) distribution. This work has demonstrated that combination of molecular simulation with SANS can give unique information on configurational distributions of denatured proteins, down even to the local level.

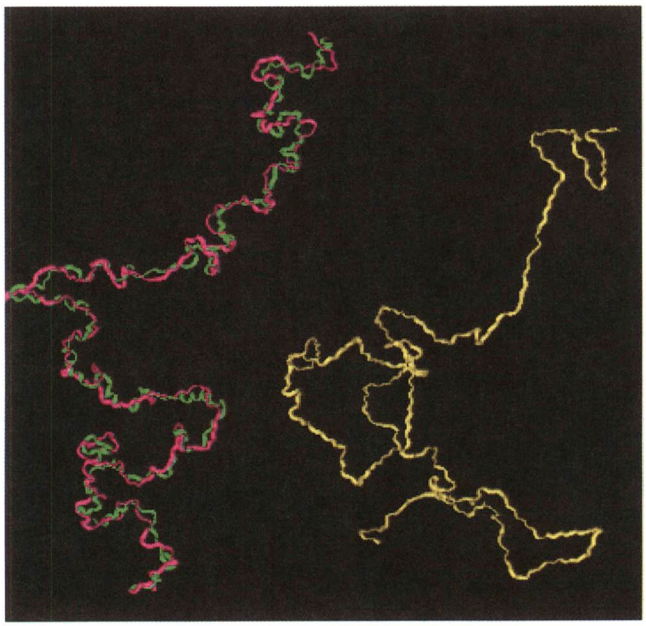

Fig. 1. Snapshots of statistical models of strongly-denatured PGK. Comparison of these models with SANS experiments has provided new information on the configurational statistics of non-native protein states. Full details in Petrescu *et al.*, 2000.

§3. Protein Dynamics

Just as conformational flexibility must be understood in protein folding, so must it also be understood in protein function. It is generally accepted that enzymes require internal flexibility for catalytic activity. However, which motions are required is not yet well understood. Of particular interest is the role in activity of the fast (e.g., picosecond (ps) and nanosecond (ns) timescale) structural fluctuations, that are probed by molecular dynamics simulations. The picosecond motions in proteins are a combination of vibrations and conformational transitions. The vibrational part of the spectrum can be usefully probed with neutron scattering.[9] In recent work, the high-resolution TFXA spectrometer at the Rutherford-Appleton Lab (ISIS), Oxford, was used to determine the vibrational spectrum of a protein, Staphylococcal nuclease, at 25 K.[10] We present here preliminary results obtained from the higher-resolution spectrometer, TOSCA, which has replaced TFXA. In Fig. 2 the experimental data are compared with a spectrum calculated from a normal mode analysis of the protein, calculated using the CHARMM program.[11] The basic features of the spectrum are reproduced by the calculation, attesting to the quality of the force field. Further details on these calculations will be published later.

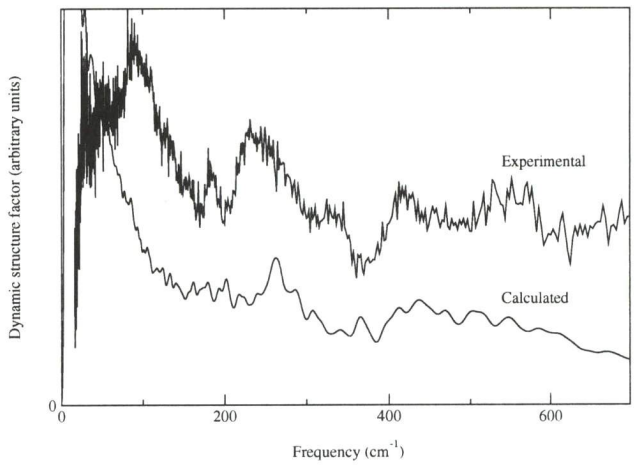

Fig. 2. Experimental low-frequency neutron vibrational spectrum from the TOSCA instrument at the Rutherford-Appleton Lab, in comparison with spectrum calculated from a normal mode analysis of the protein. The two spectra have been separated on the vertical scale for clarity.

An important question is whether the fast (ps-ns) motions in proteins are coupled to the structural changes associated with the catalytic rate-limiting step, which itself may take place on timescales several orders of magnitude slower. Do the fast motions need to be anharmonic for protein function? This might be the case if, for example, ps-ns motions involve rearrangements of the protein that are required to permit slow dynamics across the highest-energy reaction barrier.

X-ray diffraction, dynamic neutron scattering and various spectroscopies have demonstrated a quantitative change in the nature of internal motions of proteins, at $\sim$200-220 K.[12] Below this transition the internal motions are essentially harmonic whereas above it anharmonic dynamics contribute and, at physiological temperatures, dominate the internal fluctuations. The anharmonic motions may involve confined continuous diffusion[13] and/or jump diffusion between potential energy wells associated with 'conformational substates' of slightly different structure in which proteins are trapped below the transition.[14]

Correlations have been made between some protein functions, such as ligand binding or proton pumping, and the presence of equilibrium anharmonic motion.[15,16] The increased flexibility conferred by anharmonic dynamics may indeed be required for some proteins to rearrange their structures to achieve functional configurations. However, the forms and timescales of the an-

harmonic motions required for function are in general unknown.

The first parallel studies of enzyme activity and dynamics, using glutamate dehydrogenase (GDH) in a cryosolution, probed picosecond timescale motions. The results showed no deviation from Arrhenius behaviour through the dynamical transition.[17] These results indicate that there is a range of temperatures (190-220 K) at which the enzyme rate limiting step does not require, and is not affected by, anharmonic motions taking place on the picosecond timescale.

Motions in proteins are known to exist over a range of timescales. An important remaining issue then is the timescale dependence of the dynamical transition. Whether there is a timescale dependence depends on the nature of the underlying potential surface. Motions in a cryosolution of the enzyme glutamate dehydrogenase (GDH) were examined and compared with activity.[18,19] Dynamic neutron scattering was performed with two instruments of different energy resolution, permitting the separate determination of the average dynamical mean-square displacements on the sub ~100 ps and sub ~5 ns timescales. The results demonstrate a marked dependence on timescale of the temperature profile of the mean-square displacement (Fig. 3). The lowest temperature at which anharmonic motion is observed is heavily dependent on the time window of the instrument used to observe the dynamics. Several dynamical transitions (inflexions of the mean-squared displacement) are observed in the slower dynamics. Comparison with the temperature profile of the activity of the enzyme in the same solvent reveals dynamical transitions that have no effect on GDH function. As the temperature is lowered, the anharmonic motions slow down, moving from fast to slow time windows. This qualitative effect is consistent with a description of the dynamics associated with the lowest transition as essentially activated i.e. involving energy-barrier crossing.

To further understand the nature of dynamical transitions in proteins it is particularly important to characterize solvent effects. Solvent can in principle affect protein dynamics by modifying the effective potential surface of the protein and/or by frictional damping. Therefore, a solvent dependence of the dynamical transition might be expected. In very recent work a dynamic neutron scattering analysis of the solvent dependence of the picosecond-timescale dynamical transition behaviour of solutions of a simple single-subunit enzyme, xylanase was performed.[20,21] The protein was examined in powder form, in D_2O and in four two-component perdeuterated single-phase cryosolvents in which it is active and stable. The scattering profiles of the mixed solvent systems in the absence of protein were also determined. The general features of the dynamical transition behaviour of the protein solutions were found to follow those of the solvents. The dynamical transition in all the mixed cryosolvent: protein systems was observed to be much more gradual than in pure D_2O, consistent with a distribution of energy barriers. The differences between the dynamical behaviours of the various cryosolvent protein solutions themselves are remarkably small. These results are consistent with a picture in which the picosecond-timescale atomic dynamics responds strongly to melting of pure water solvent but is relatively invariant in cryosolvents of differing composition and melting points.

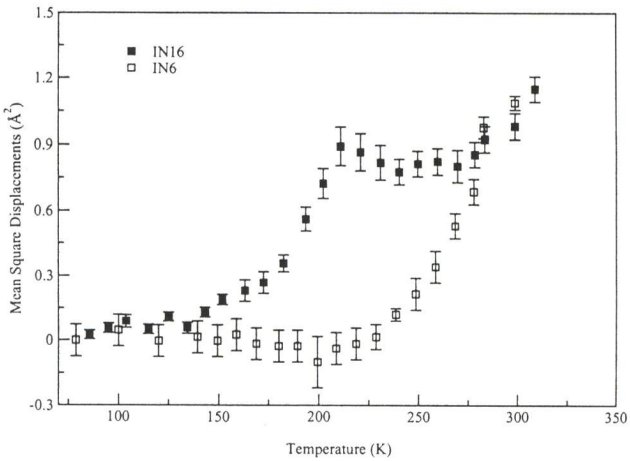

Fig. 3. Effect of temperature on the dynamics as measured by neutron scattering, of glutamate dehydrogenase in CD_3OD/D_2O (70:30). From Daniel *et al.*, 1999. The open squares give the fast motions, measured on IN6, and the filled squares the slower motions, measured on IN16.

§4. Discussion

The above-presented results indicate that neutron scattering can be combined with molecular simulation to provide useful information on a range of protein motions, from internal vibrations via anharmonic transitions up to protein folding configurational distributions.

In protein folding different points along the folding pathway must be characterised. Strongly denatured proteins are often considered to be completely unfolded i.e., in a 'random coil' state with no residual structure. But even in such a state the polypeptide chain will have local conformational preferences. This leads to fundamental problems in protein folding concerning the determination of these preferences, environmental effects on them, their reconciliation with non-local configurational statistics and possible deviations from native state local conformational distributions. Neutron scattering presents a new approach that promises to be of use in addressing these problems. The results so far obtained demonstrate how the combination of atomic-detail simulation with small-angle neutron scattering can be used to provide information on both the global and local configurational statistics of a strongly denatured protein.

The SANS approach allows the long- and short-range residue-averaged properties of the chain to be combined, via the radial distribution function. In a 'folding funnel' parlance, the width of the funnel is related to the accessible configurational space of the protein chain. A strongly denatured protein can be thought of as representing the unfolded end state of protein folding, such that the

present results provide information on the widest rim of the folding funnel, and which regions of conformational space it contains. This information can be directly incorporated in statistical mechanical analyses of protein folding. With further experimental and theoretical work, the present combination of experiment and theory may also prove fruitful in determining the configurational space of different proteins and at other points in the funnel i.e., at various stages in the folding process.

Concerning the dynamics of native protein states, clearly more needs to be done in combining techniques with which the dynamical transitions can be studied. Further work is required to examine the intramolecular localisation of functionally important motions. It is encouraging in this respect to see that site-specific labelling studies have been recently shown to be successful in dynamic neutron scattering studies on proteins.[16] The results of further investigations can be expected to improve our understanding of the relationships between protein activity, flexibility and stability. Clearly, a complete understanding of these and related problems will require continued investigation into the forms, amplitudes and timescales of motions involved in protein folding and function.

1) M. Kataoka, I. Nishii, T. Fijisawa, T. Ukey, F. Tokunaga and Y. Goto: J. Mol. Biol **249** (1995) 215.
2) M. Kataoka and Y. Goto: Folding & Design **1** (1997) 107.
3) P. Calmettes, B. Roux, D. Durand, M. Desmadril and J. C. Smith: J. Mol. Biol **231** (1993) 840.
4) P. Calmettes, D. Durand, M. Desmadril, P. Minard, V. Receveur and J. C. Smith: Biophys. Chem **53** (1994) 105.
5) V. Receveur, P. Calmettes, J. C. Smith, M. Desmadril, G. Coddens and D. Durand: Proteins Struct. Funct. Genetics **28** (1997) 380.
6) A. J. Petrescu, V. Receveur, P. Calmettes, D. Durand, N. Desmadril, B. Roux and J. C. Smith: Biophys. J **72** (1997) 335.
7) A. J. Petrescu, V. Receveur, P. Calmettes, D. Durand and J. C. Smith: Prot. Sci. **7** (1998) 1396.
8) A. J. Petrescu, P. Calmettes, D. Durand, V. Receveur and Smith: Prot. Sci. **9** (2000) 1129.
9) J. C. Smith: Q. Rev. Biophys. **25** (1991) 227.
10) A. Lamy, J. C. Smith, J. Yunoki, S. F. Parker and M. Kataoka: J. Am. Chem. Soc. **119** (1997) 9268.
11) B. R. Brooks, R. E. Bruccoleri, B. Olafson, D. J. States, S. Swaminathan, and M. Karplus: J. Comput. Chem. **4** (1983) 187.
12) W. Doster, S. C. Cusack and W. Petry: Nature **337** (1989) 754.
13) G. R. Kneller and J. C. Smith: J. Mol. Biol. **242** (1994) 181.
14) H. Frauenfelder, S. Sligar and P. Wolynes: Science **254** (1991) 1598.
15) M. Ferrand, A. J. Dianoux, W. Petry, and G. Zaccai: Proc. Natl. Acad. Sci. U.S.A. **90** (1993) 9668.
16) V. Réat, H. Patzelt, M. Ferrand, C. Pfister, D. Oesterhelt and G. Zaccai: Proc. Natl. Acad. Sci. USA. **95** (1998) 4970.
17) R. M. Daniel, J. C. Smith, M. Ferrand, S. Hery, R. Dunn and J.L. Finney: Biophys. J. **75** (1998) 2504.
18) R. M. Daniel, J. L. Finney, V. Réat, R. Dunn, M. Ferrand and J. C. Smith: Biophys. J. **77** (1999): 2184.
19) R. V. Dunn, V. Réat, J. L. Finney, M. Ferrand, J. C. Smith and R. M. Daniel: Biochem. J. **346** (2000) 355.
20) V. Réat, J. L. Finney, A. Steer, M. A. Roberts, J. C. Smith, R. Dunn, M. Petersen and R. M. Daniel: J. Biochem. Biophys. Methods **42** (2000) 97.
21) V. Réat, R. Dunn, M. Ferrand, J. L. Finney, R. M. Daniel and J. C. Smith: Proc. Nat. Acad. Sci.(U.S.A.) **97** (2000) 9961.