enhance the specificity of TRK binding and this could be tested by performing binding studies in the LNGFR-mice. One can perhaps also speculate that NGF effects are independent of LNGFR (consistent with the normal sympathetic phenotye in the LNGFR-mice) but that BDNF which supports a subpopulation of DRG neurons may require LNGFR expression. Examining the TRK subtype expression on the surviving sensory neurons would answer the question. These and other experiments may provide a clearer understanding of the role of *trk* and LNGFR *in vivo*.

1. Chao, M., *Neuron*, 1992, 9, 583–593.
2. Lee, K. *et al.*, *Cell*, 1992, 69, 737–749.
3. Carroll, S. *et al.*, *Neuron*, 1992, 9, 779–788.
4. Rodriguez-Tebar, A., Denchant, G., Gotz, R. and Barde, Y., *EMBO J.*, 1992, 11, 917–922.

*Mahendra Rao is in the Howard Hughes Medical Institute Laboratories, Division of Biology, Caltech, Pasadena, CA 91125, USA*

# Folding in an unfolded protein

*Jayant B. Udgaonkar*

For more than 25 years, it has virtually been a dogma to most biochemists, that a protein dissolved in a high concentration of denaturant, for instance, 6 molar guanidine hydrochloride or 8 molar urea, loses its structure so completely that for all purposes it exists in a random coiled conformation. The work leading to this wide-spread belief was elegantly summarized in two classic reviews on protein denaturation[1,2]. The experimental methods used at that time, however, lacked structural resolution. For instance, viscosity and light scattering experiments could show that the radius of the polypeptide chain was independent of sequence for several proteins and dependent only on the number of residues, as expected for a chain with unrestricted conformational freedom, but such measurements could not rule out the possibility that residual structure still existed.

In the last few years, the application of newer experimental methodologies has provided an increasing number of indications (summarized in ref. 3) that the notion of a completely unstructured polypeptide chain needs to be re-examined. Perhaps the most convincing of the recent experiments have been those of David Shortle and his coworkers, whose studies on the effects of certain site-directed mutations on the stability of staphylococcal nuclease could be rationalized only on the basis of the mutations affecting residual structure in the urea-denatured protein. Nevertheless, residual structure in any urea-denatured or guanidine hydrochloride-denatured protein has escaped structural characterization so far. Now Kurt Wüthrich and his coworkers have succeeded in doing just that. In two recent papers[4,5], they report the use of modern two-dimensional NMR techniques to characterize, for the first time, structure in an urea-denatured protein.

The success of two dimensional NMR spectroscopy in the determination of the solution structure of a small protein in solution, relies on the protein possessing a well-defined structure: the chemical shifts of the resonances in the NMR spectrum are then well-dispersed, which permits their sequence-specific assignments. Moreover, the presence of a unique structure usually allows unambiguous use of the NOE (Nuclear Overhauser Effect) data to arrive at a solution structure. The situation for an unfolded protein is very different. The absence of a defined structure leads to poor dispersion of chemical shifts. The presence of multiple conformations in rapid exchange not only means that observed chemical shifts are the averages of chemical shifts for all the conformations (and there will be many for a protein unfolded by high concentrations of denaturant) present, but it also obfuscates interpretation of NOE data. Neri *et al.* have now managed to overcome these difficulties and their work with the 63-residue amino-terminal domain of the phage 434 repressor is the first instance of the determination by NMR of structure in the urea-unfolded form of a protein.

Their success hinged on several factors. The availability of a good bacterial expression system for the 434 repressor made it possible to obtain protein that was selectively labelled with $^{15}N$ and $^{13}C$. Good dispersion in the $^{15}N$ chemical shifts compensated for poor dispersion in the $^{1}H$ NMR chemical shifts. The 434 repressor is fully unfolded in 7 molar urea, but under suitable conditions of pH, temperature and urea concentration, the native and fully unfolded forms could be made to coexist in equal amounts. The chemical exchange rate between the two forms was suitable for exchange-relayed NMR experiments: the NMR assignments of the fully unfolded form[4] could be obtained using the assignments for the fully folded native form[6].

Neri *et al.*[4–6] observed a large set of NOE distance constraints for the polypeptide segment 53 to 60, which could all be satisfied by a single well-defined structure obtained using distance geometry calculations and the NOE data. They could not estimate the fraction of the number of protein molecules that possess this segmental structure, but they point out that the calculation of the structure is unaffected by the presence of protein molecules that are completely unstructured. This is important because the chemical shift data indicate that completely unstructured molecules coexist with the partly-structured ones.

The structure that is present in the polypeptide segment 53 to 60 for a significant fraction of the protein molecules is, however, the most definitive evidence so far for the clustering of hydrophobic residues in a polypeptide unfolded by high concentrations of denaturant. This polypeptide segment has four apolar residues, and these residues form a local hydrophobic cluster which is structurally related to a larger hydrophobic core seen in the fully folded protein[5].

There has been considerable recent interest in the proposal[7] that a random hydrophobic collapse is the first step in the folding process, and experimental evidence for such a collapse has been mounting[8]. The detection of a hydro-

phobic cluster in urea-unfolded 434 receptor could mean that, during folding, a hydrophobic collapse of the polypeptide chain may be foreordained, with the hydrophobic cluster acting as a nucleation site. Thus, if a hydrophobic collapse does occur as the first step, it would not be expected to be completely random. In fact, Neri et al.[5] have used their data to propose a model for the first events that occur on the folding pathway of the 434 repressor.

The concept of a solvent-denatured protein possessing a random coiled conformation is also under attack from another quarter. Peter Privalov and his group have used differential scanning calorimetry to study the thermodynamics of the interaction of urea and guanidine hydrochloride with several proteins[9]. They could explain their results with a simple binding model with independent and equivalent binding sites, each formed by several hydrogen bonding groups. Such binding, which is observed to be accompanied by a substantial decrease in enthalpy and entropy, would significantly restrict the conformational freedom within a polypeptide chain.

The data of Privalov and his co-workers have also resulted in a resurrection of the binding model for denaturation[2]. In recent years, this model had lost ground to other models according to which urea and guanidine hydrochloride act indirectly through their interaction with water (after all, the molar concentration of water in a 6 M guanidine hydrochloride solution is only about half that of pure water), either through hydrophobic interactions or through other weak interactions[10]. It is too early, however, to tell whether the

simple binding model is the only one that can account for the calorimetric data, but what is particularly attractive about it is that it accounts for the necessity of using molar concentrations of denaturant to unfold micromolar concentrations of protein.

The NMR data reveal the presence of specific residual structure in the unfolded form of one protein, the 434 repressor. If this phenomenon is general, as suggested by the calorimetric data which point to the presence of non-random structure in several unfolded proteins previously thought to be in random coiled conformation, it has important ramifications for experimental studies on protein folding. The method of choice for studying how proteins fold is to identify and characterize structural intermediates on the kinetic folding pathway. In the last few years, a method to determine the structures of such transient intermediates has been developed, which exploits the capabilities of two-dimensional NMR and amide hydrogen exchange methods[11,12], so that it is becoming possible to determine the temporal and spatial development of structure in folding protein molecules. In such kinetic studies, there has always been an implicit assumption that the starting state in a high concentration of denaturant is completely unstructured, and it has therefore served as a standard reference state in folding studies. The absence of such a reference state may lead to the results of kinetic studies being misconstrued, particularly with regard to the earliest formation of structure.

No less important is the bearing on computational methods to study protein

folding. A major hurdle in the application of such methods has been that although the structure of the fully folded product of the folding reaction is very well-characterized for hundreds of proteins, very little is known about the structure of the fully unfolded starting state. Better structural information on fully unfolded proteins will undoubtedly lead to more realistic computer modelling studies of protein folding reactions.

1. Tanford, C., Adv. Prot. Chem., 1968, 23, 121–282.
2. Tanford, C., Adv. Prot. Chem., 1970, 24, 1–95.
3. Dill, K. A. and Shortle, D., Annu. Rev. Biochem., 1991, 60, 795–825.
4. Neri, D., Wider, G. and Wüthrich, K., Proc. Natl. Acad. Sci. USA, 1992, 89, 4397–4401.
5. Neri, D., Billeter, M., Wider, G. and Wüthrich, K., Science, 1992, 257, 1559–1563.
6. Neri, D., Billeter, M. and Wüthrich, K., J. Mol. Biol., 1992a, 223, 743–753.
7. Dill, K. A., Biochemistry, 1985, 24, 1501–1506.
8. Baldwin, R. L., Chemtracts, 1991, 2, 379–389.
9. Makhatadze, G. I. and Privalov, P. L., J. Mol. Biol., 1992, 226, 491–505.
10. Schellman, J. A., Annu. Rev. Biophys. Biophys. Chem., 1987, 16, 115–137.
11. Roder, H., Elöve, G. A. and Englander, S. W., Nature, 1988, 335, 700–704.
12. Udgaonkar, J. B. and Baldwin, R. L., Nature, 1988, 335, 694–699.

*Jayant B. Udgaonkar is in the National Centre for Biological Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India*