PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Models of Cooperativity in Protein Folding

Hue Sun Chan, Sarina Bromberg and Ken A. Dill

Phil. Trans. R. Soc. Lond. B 1995 **348**, 61-70
doi: 10.1098/rstb.1995.0046

Models of cooperativity in protein folding

HUE SUN CHAN, SARINA BROMBERG AND KEN A. DILL

Department of Pharmaceutical Chemistry, Box 1204, University of California, San Francisco, California 94143–1204, U.S.A.

SUMMARY

What is the basis for the two-state cooperativity of protein folding? Since the 1950s, three main models have been put forward.

1. In 'helix-coil' theory, cooperativity is due to local interactions among near neighbours in the sequence. Helix-coil cooperativity is probably not the principal basis for the folding of globular proteins because it is not two-state, the forces are weak, it does not account for sheet proteins, and there is no evidence that helix formation precedes the formation of a hydrophobic core in the folding pathways.

2. In the 'sidechain packing' model, cooperativity is attributed to the jigsaw-puzzle-like complementary fits of sidechains. This too is probably not the basis of folding cooperativity because exact models and experiments on homopolymers with sidechains give no evidence that sidechain freezing is two-state, sidechain complementarities in proteins are only weak trends, and the molten globule model predicted by this model is far more native-like than experiments indicate.

3. In the 'hydrophobic core collapse' model, cooperativity is due to the assembly of non-polar residues into a good core. Exact model studies show that this model gives two-state behaviour for some sequences of hydrophobic and polar monomers. It is based on strong forces. There is considerable experimental evidence for the kinetics this model predicts: the development of hydrophobic clusters and cores is concurrent with secondary structure formation. It predicts compact denatured states with sizes and degrees of disorder that are in reasonable agreement with experiments.

1. INTRODUCTION

What makes protein folding cooperative? Protein folding, at least for small single-domain proteins, is known to involve two-state transitions (Lumry *et al.* 1966; Privalov 1979). The purpose of this review is to describe one-state and two-state transitions, to review various models for the physical basis of the two-state protein folding process, and to describe how the nature of folding kinetics and denatured states must follow directly from the physical basis for the cooperativity. We review three models for cooperativity: the helix-coil model, the sidechain packing model, and the hydrophobic collapse model. We conclude that protein folding cooperativity is modelled most simply as a heteropolymer hydrophobic collapse process.

What is cooperativity? Figure 1*a* shows a non-cooperative process in which a system gradually changes from state A to B as a function of temperature or denaturant. Figure 1*b*, *c* shows the experimental signature of a cooperative process, namely sigmoidal behaviour. Cooperativity may be of two different types: two-state (first order) or one-state (higher order). Neither the observation of a sigmoidal transition curve, nor the observation of a peak of heat absorption by scanning calorimetry can distinguish a two-state from a one-state cooperative transition. The definitive way to distinguish experimentally between these two types of transition is by an analysis of the populations. At the midpoint of a two-state transition, two identifiable states will be populated. In contrast, one-state behaviour is defined by a single broad population peak at the midpoint. It follows that two-state cooperativity implies a free energy surface with two minima separated by a barrier. One-state cooperativity implies a single free energy minimum. Whether the cooperativity of a process is one- or two-state does not necessarily depend on the steepness of its sigmoidal curve. Figure 2 shows examples from exact model studies indicating that the steepness of a sigmoidal curve in a one-state process can be just as great as in a two-state process. Cooperativity depends, rather, on a molecular property that is more difficult to measure: the distribution of the underlying populations.

The discrimination of one-state from two-state cooperativity determines fundamental information about the free energy surface, and the underlying molecular mechanism of the cooperative process. How is two-state behaviour determined experimentally? There are various methods.

1. As noted above, the most definitive determination of two-state behaviour is to observe two distinct populations near the midpoint of the transition. For slowly exchanging systems, this can be done by transport methods (Cann 1970) such as size exclusion chromatography (Uversky 1993) or gel electrophoresis (Creighton 1986). Sometimes a spectroscopic method will give distinct signatures for the native molecule in native conditions and the highly unfolded molecule in

Phil. Trans. R. Soc. Lond. B (1995) **348**, 61–70
Printed in Great Britain

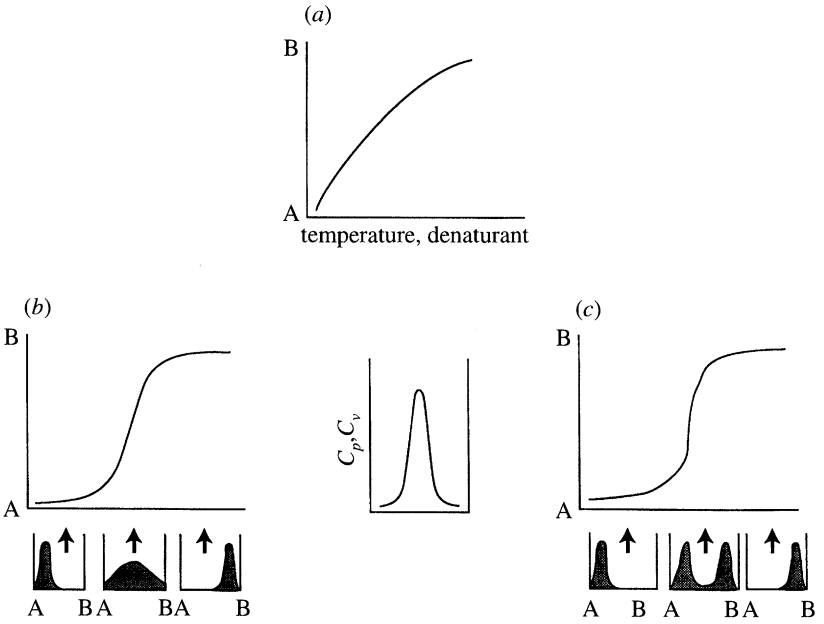


Figure 1. What is cooperativity? Two states, A and B, such as native and denatured states, can change populations with temperature, denaturant, pH, salt, etc. (*a*) Gradual change, no cooperativity. (*b*) Cooperative transition of the one-state type. (*c*) Cooperative transition of the two-state type. Both one-state and two-state transitions can have sigmoidal behaviour and heat absorption (a peak in the C_p or C_v plot); they cannot be distinguished on these bases, or from the steepness of the sigmoidal curve. The main distinction is whether there is one broad peak involving high populations of 'intermediates' near the denaturant midpoint (one-state) or whether there are two populated states and less intermediate population (small plots at the bottom of (*b*) and (*c*)).

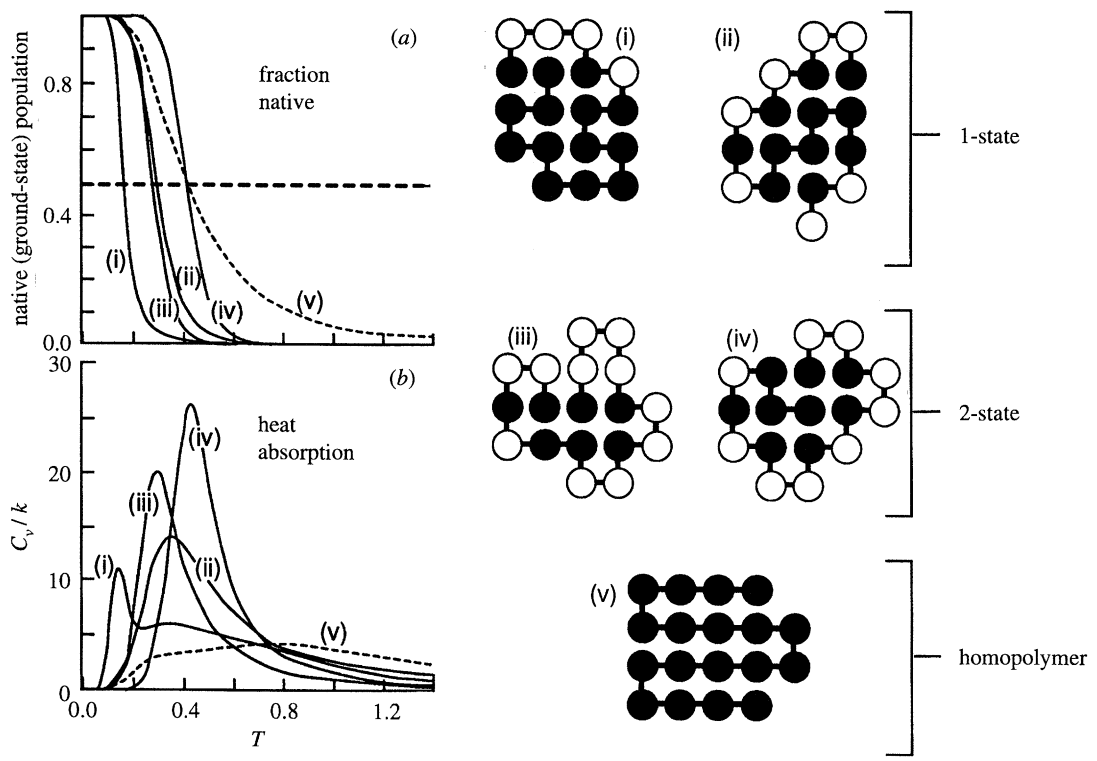


Figure 2. Cooperativity among HP lattice model sequences (H, hydrophobic, black beads; P, polar, white beads). All five of these sequences collapse cooperatively, i.e. have sigmoidal changes in the native population with temperature T, and have heat absorption (C_v) peaks. Sequence (v) is a homopolymer shown in one of its 1673 maximally compact states; all four other sequences are shown in their unique native states. Only an analysis of the underlying populations can show that the cooperativity for sequences (i), (ii), and (v) is one-state, and for sequences (iii) and (iv) is two-state. Figure 3 demonstrates this by showing the underlying populations for sequences (i) and (iv).

strongly denaturing conditions. Spectroscopic signal changes through the denaturation midpoint that can be modelled as a linear combination of the signals from the two limiting states are an indication of two-state behaviour. On the other hand, if the spectroscopic signal near the denaturation midpoint cannot be fitted

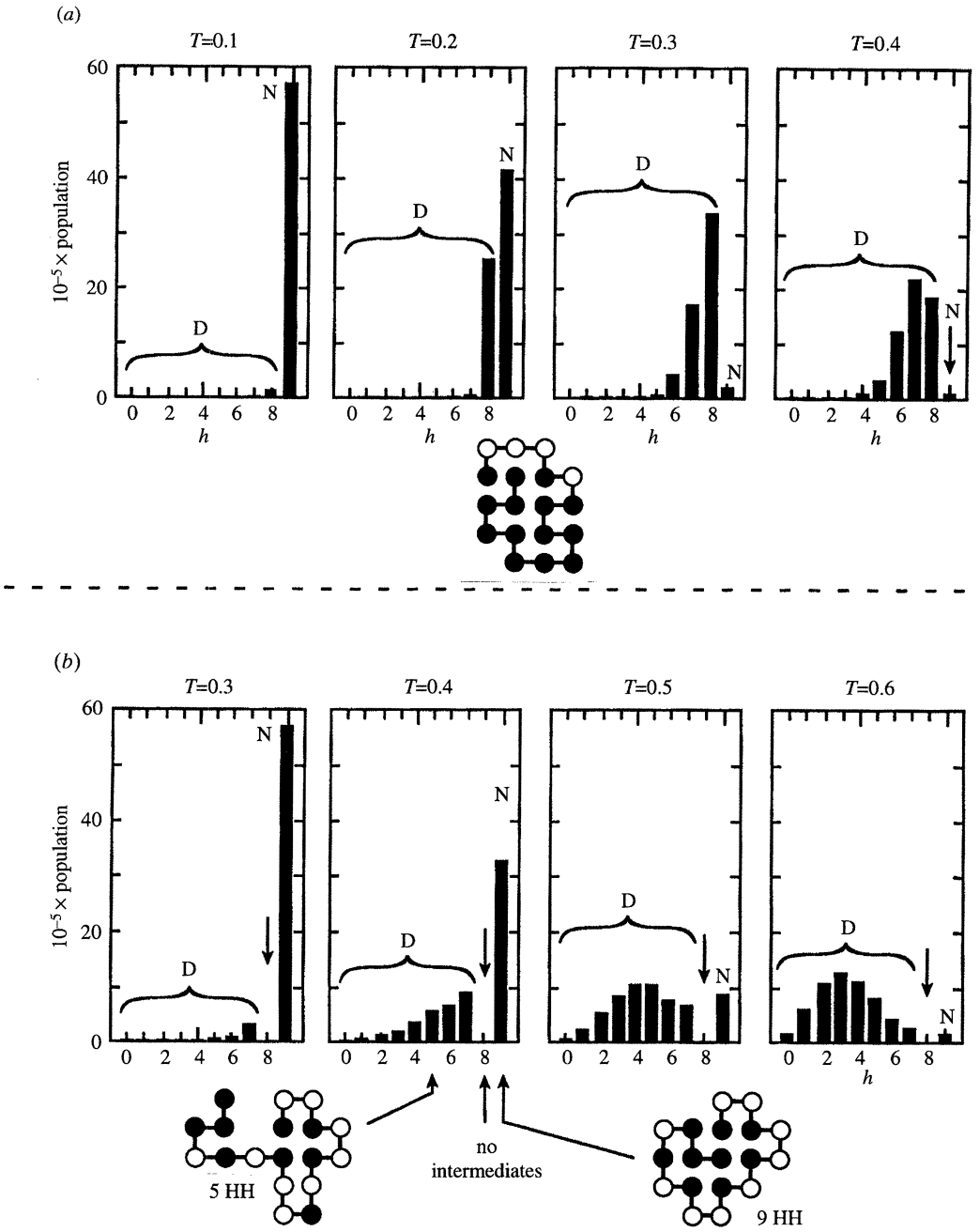


Figure 3. Population analysis of conformations for the HP lattice model sequences in figure 2. (*a*) Sequence (i) shows one-state behaviour. (*b*) Sequence (iv) shows two-state behaviour. On the horizontal axis, h is the number of HH contacts ($h = 9$ is native for both sequences); the vertical axis shows the populations. An example denatured conformation with $h = 5$ is shown in (*b*). At low temperature most molecules are in their native states. At higher temperature most molecules are denatured. Near the denaturation midpoint (*a*) shows a high population of intermediates, while (*b*) shows no population of $h = 8$ intermediate conformations. Notice also that the denatured state shifts population with temperature: at low temperatures the denatured molecules are very compact (high h), whereas at higher temperatures the denatured molecules are more unfolded.

as a simple linear combination of the two limiting state signals, it is evidence for one-state behaviour. Now a significant population of 'intermediate' molecules, which may have different spectral signatures, may contribute to the total signal.

2. Another measure of two-state behaviour is the ratio of the calorimetric enthalpy (a model-independent quantity, determined directly from experiments), to the van't Hoff enthalpy (a model-dependent quantity based on certain thermodynamic assumptions) (Privalov 1979; Privalov & Gill 1988).

3. A less definitive measure of two-state behaviour is the observation that different experimental measures show coincident sigmoidal curves.

The reason that (3) (and possibly (2)) are less definitive measures of two-state behaviour is that both are based on the assumption, not always correct, that the two states of proteins are as fixed and simple as the two different states observed in many small-molecule processes. However, the second state of proteins, the denatured state, is more complex. The radius and other properties of the denatured molecules can change

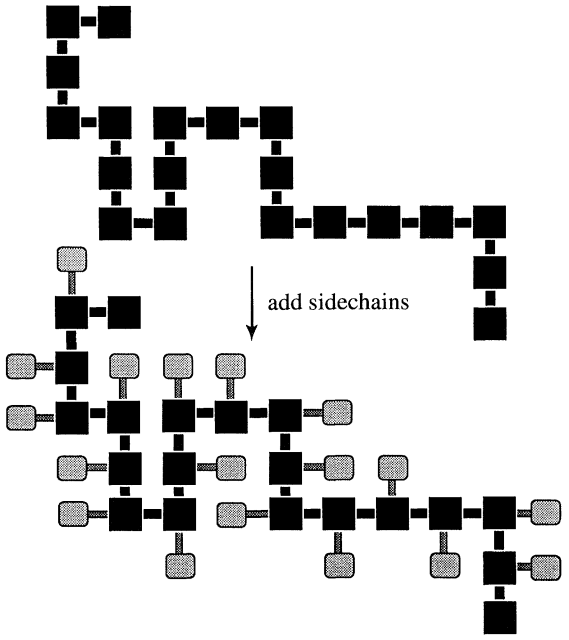


Figure 4. Simple exact side-chain model. Taking the linear chain lattice model to represent the main chain (upper figure), a side-chain model is created by attaching a single side chain unit to each main chain monomer. To represent side chain rotameric degrees of freedom, each side chain unit has the freedom to occupy any one empty lattice site adjacent to its corresponding main chain monomer (lower figure); see Bromberg & Dill (1994).

with external conditions. Two-state protein unfolding behaviour can lead to situations in which different experimental measures will not have coincident sigmoidal curves (see appendix, Dill & Shortle 1991). Coincident curves give evidence supporting two-state behaviour, but non-coincident curves are not definitive evidence against two-state behaviour.

What is the physical basis for cooperativity in protein folding? The first model for cooperativity in biopolymer conformational transitions is the helix-coil theory (Schellman 1958; Zimm & Bragg 1959; Lifson & Roig 1961; Poland & Scheraga 1970). This model represents the cooperativity that comes from the influences of nearest neighbours along the chain. A residue has more propensity to be in a helical configuration if its neighbour in the sequence is helical than if its neighbour is in a coil configuration. Helix-coil theories are based on the one-dimensional Ising model, for which the cooperativity has been proven to be one-state (Stanley 1971). Thus local interactions along the chain are not sufficient to account for the two-state cooperativity of protein folding.

Another possibility is that protein cooperativity originates in the non-local interactions that cause the sharp collapse of hydrophobic polymers in water (Fujishige *et al.* 1989; Ricka *et al.* 1990; Meewes *et al.* 1991; Tiktopulo *et al.* 1994). Ptitsyn and coworkers were the first to develop a model showing that homopolymer collapse processes would be cooperative (Ptitsyn *et al.* 1968). However, many theoretical model studies beginning in the 1970s (de Gennes 1975; Moore 1977; Post & Zimm 1979; Sanchez 1979; Grosberg & Khokhlov 1987) determined that homopolymer collapse is only one-state, unless the chain stiffness is high. Thus the collapse of flexible homopolymers could not account for the two-state cooperativity of protein folding. Recent experiments show that the collapse of homopolymers of poly-*N*(isopropylacrylamide) (PNIPAM) is likely to be a one-state process (Tiktopulo *et al.* 1994).

As homopolymer collapse and helix-coil models could not account for the two-state cooperativity of globular proteins, two different models for protein folding cooperativity were developed during the 1980s. A mean-field theory was developed, based on the assumption that protein (heteropolymer) collapse differs from homopolymer collapse in that some sequences of hydrophobic and polar monomers can cooperatively form good hydrophobic cores (Dill 1985; reviewed in Chan *et al.* 1992; Dill & Stigter 1994). We call this the hydrophobic core (HC) model. On the other hand, Shakhnovich and Finkelstein (1989) regarded cooperativity as arising from jigsaw-puzzle-like packing and interdigitation among sidechains. They modelled unfolding cooperativity arising as sidechains suddenly become freed from their local packing constraints at a critical disjuncture point in the chain expansion. We call this the sidechain packing (SP) model. The HC model assumed two-state behaviour, rather than proving it, and was based on two mean-field approximations. The SP model was based on many different assumptions and approximations. It was not clear whether the two-state cooperativity assumed and implied in those models could be derived on more rigorous grounds. To resolve the basis for cooperativity, we have recently resorted to simplified models in which the partition function can be known exactly and without approximation. This review summarizes those results.

2. HYDROPHOBIC CORE MODEL

By exact enumeration, we have explored the full conformational space of some HP lattice model chains on two-dimensional square lattices (Lau & Dill 1989, 1990; Chan & Dill 1991; Chan *et al.* 1992). Figure 2 shows the native population at various temperatures, for each of four different unique sequences, and a homopolymer. Each unique sequence folds only to a single lowest energy native state, whereas the ground ('native') state of the homopolymer is the full ensemble of maximally compact conformations (Chan & Dill 1989). Figure 2 shows that all these sequences have a sigmoidal variation of the fraction of molecules that are native versus temperature, and that all these sequences have peaks of heat absorption upon folding; observations generally taken to signal cooperative behaviour. But these simple tests (sigmoidal behaviour, heat absorption) alone do not tell us whether the transitions are one-state or two-state. Such information is revealed most definitively by a full population analysis; see figure 3. Exact calculations show rigorously that two-state behaviour is observed in two of the sequences, and one-state behaviour is found for the other two unique sequences. Thus it is clear that the collapse of heteropolymers and the formation of a good hydro-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

phobic core is a sufficient physical basis to obtain the type of two-state behaviour observed in the folding of globular proteins. But within the HP model, not all sequences will do this. For example, the homopolymer sequence HHHH . . . . H collapses with a one-state transition.

3. SIDECHAIN PACKING MODEL

Figure 4 shows the simple exact model we have used to study sidechain packing (Bromberg & Dill 1994) by exhaustive enumeration for short chains, and by Monte Carlo sampling for longer chains. In this simple model, sidechains are represented as monomers of uniform size and shape.

Figure 5*b* shows the computed conformational entropies due to sidechain degrees of freedom and excluded volume. As longer chains expand in denaturation from the native state, sidechain freedom and conformational entropy increase rapidly at first, then more slowly while the volume explored by the backbone continues to expand. The sharp gain in freedom near the native state is like an 'unfreezing' of the sidechains. In contrast, figure 5*a* shows the model of Shakhnovich & Finkelstein (1989), which assumes a 'critical disjuncture point' at which the sidechains suddenly become unlocked, not at the earliest stages of chain expansion, but at a critical point of expansion, around a 25% volume increase. This disjuncture point is the basis for the cooperativity in their model, but such a critical disjuncture point is not found in the simple exact model study.

In light of the fact that neither the model of Shakhnovich & Finkelstein, nor our simple exact sidechain model accurately represents real protein sidechains, what conclusions can be drawn? We believe the exact sidechain model,in its simplicity, bears closer resemblance to real proteins, insofar as sidechains do not need to be precisely matched in shape and locked together for a protein to achieve either (i) tight packing, or (ii) native topology. The following data support this view.

Sidechains can achieve native packing densities without locking together according to a precise stereochemical code:

1. In the crystal structures of the Protein Data Bank (PDB) (Bernstein *et al.* 1977; Abola *et al.* 1987), sidechains show little preference for particular partners to mutually bury surface area (Behe *et al.* 1991).
2. Pairs of hydrophobic core sidechains in the PDB show little preference for particular mutual orientations (Singh & Thornton 1990, 1992).
3. Native packing density in computer simulated protein structures is readily achieved even by misfolded sequences (Novotny *et al.* 1984).
4. Protein interiors are dynamic: small molecules rapidly diffuse to buried residues (Gurd & Rothgeb 1979; Karplus & McCammon 1981); and aromatic

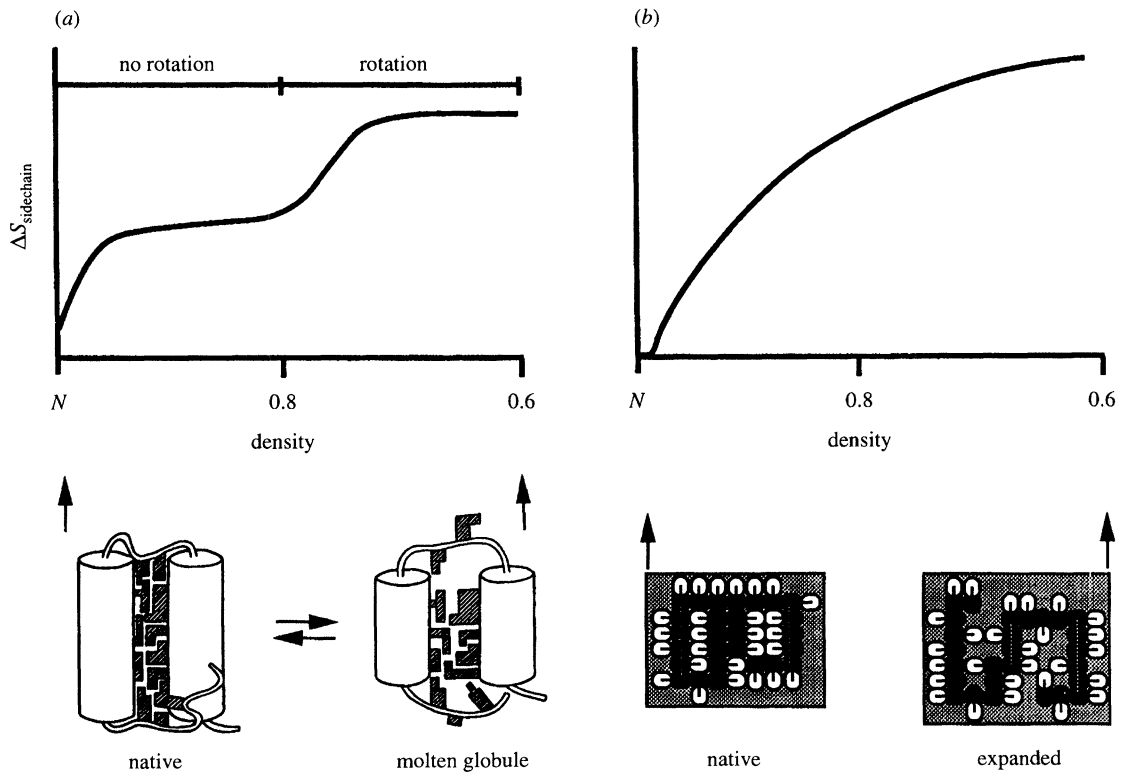


Figure 5. (*a*) Sidechain molten globule model: jig saw puzzle model. A first-order transition is proposed to arise from the native to molten globule state because of a sharp increase in sidechain rotational entropy at a critical disjuncture point. Backbone and secondary structures are assumed fixed in native-like conformations, and thus assumed to be independent of sidechain freedom (Shakhnovich & Finkelstein 1989; Ptitsyn 1992). (*b*) Nuts and bolts model. By exact enumeration in a simplified model of sidechains, the sidechain rotational entropy is shown to increase most sharply even at the earliest expansions from the native state, implying no critical disjuncture point (Bromberg & Dill 1994). Sidechains and backbone are found to be strongly coupled. It is proposed that a critical disjuncture point, corresponding to sidechain unlocking, is not the defining characteristic of compact denatured states.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

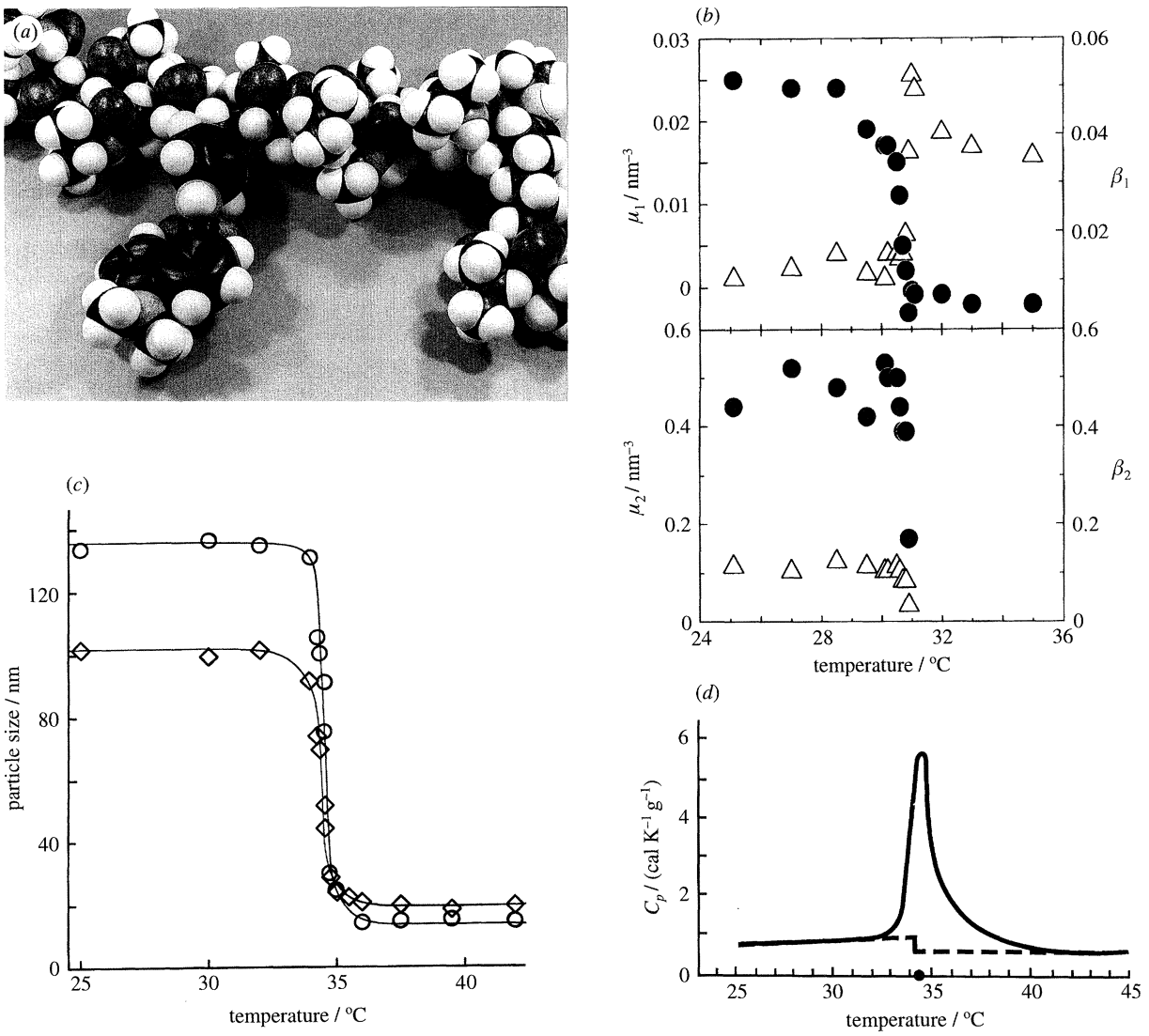


Figure 6. Collapse of poly(*N*-isopropylacrylamide) (PNIPAM) homopolymers in aqueous solutions with small amount of surfactant to suppress aggregration. (*a*) Space-filling model of a section of the polymer. (*b*) Homopolymer collapse freezes out the motions of the sidechains. Time-resolved measurements of fluorescence polarization anisotropy are used to monitor sidechain motions due to the temperature-induced conformational transition. The plots show reduced re-orientational relaxation rates μ_i and their amplitudes β_i versus temperature: (left, filled circles) μ_1, (right, open triangles) β_1, reflecting backbone motions; (left, filled circles) μ_2, (right, open triangles) β_2, reflecting mainly local sidechain motions. Both types of motions undergo a 'freezing' transition at the same temperature around 31 °C. The chains have approximately 3100 monomer units. (*a*) and (*b*) are reproduced from Binkert *et al.* (1991). (*c*) Hydrodynamic radius (open diamonds) and radius of gyration (open circles) of PNIPAM as a function of temperature (data from Meewes *et al.* 1991). (*d*) Temperature dependence of the partial specific heat capacity of PNIPAM. Dashed lines show the extrapolations of heat capacity curves from low and high temperatures to the middle of the transition region. Reproduced from Tiktopulo *et al.* (1994). The number of monomers per chain in experiments (*c*) and (*d*) are approximately 62000.

rings rotate with low activation energy barriers (Wuthrich & Wagner 1978). The temperature factors of protein crystals (Artymiuk *et al.* 1979; Frauenfelder *et al.* 1979) show that the backbone atoms of an amino acid residue in a protein core are more rigidly immobilized than its side-chain atoms.

5. Cavities created by mutations in protein cores are not rigid like jigsaw puzzles lacking a piece. (McRee *et al.* 1990; Eriksson *et al.* 1992; Varadarajan & Richards 1992). Main chain readjustments accomodate many mutations that change side chain volumes (Baldwin *et al.* 1993).

6. The kinetic bottleneck to protein folding is likely to be the breaking of non-native bonds or contacts, rather than side chain locking. It has been shown that a single non-native heme interaction slows the folding of cytochrome *c* by orders of magnitude (Sosnick *et al.* 1994), and non-native disulphide bonds or aromatic interactions occur in folding intermediates of hen lysozyme (Chaffotte *et al.* 1992; Hooke *et al.* 1994). Other examples of non-native contacts are reviewed by Creighton (1994).

Proteins can achieve native topologies without unique sidechain packing:

1. Topologically similar proteins can have differently packed cores (Swindells & Thornton 1993).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

2. Proteins can maintain native-like topology in states that lack native tight packing (Feng *et al.* 1994; Hughson *et al.* 1994; Peng & Kim 1994).

3. The fold of some proteins, such as globins, can be achieved by sequences that are less than 20% identical (Bashford *et al.* 1987).

4. Proteins show considerable structural tolerance for mutations that change core side chain size and shape (Matthews 1987, 1993; Lim & Sauer 1991; Keefe *et al.* 1993, Richards & Lim 1994).

Like proteins (Richards 1974, 1977; Richards & Lim 1994), the simple exact sidechain model fills space without large voids. The configurational freedom of the sidechains is linked to the freedom of the backbone. These model chains collapse to configurations with distributions of contacts among and between the main chain and the sidechains similar to the distributions found in the proteins of the PDB (Bromberg & Dill 1994). Nevertheless, because this model does not treat sidechains of different sizes or shapes, we may rigorously conclude only that this model, by itself, provides no basis for two-state cooperativity.

Experiments confirm this conclusion. Experiments with a protein-model polymer PNIPAM indicate that sidechain packing does not cause two-state behaviour. PNIPAM collapses with a very sharp sigmoidal cooperativity, and has an associated peak in heat absorption; see figure 6. The very steepness of the transition (the full collapse process takes place over only about 1 °C) is not a basis for concluding the transition is two-state. Experiments of Tiktopulo *et al.*(1994) show that the ratio of calorimetric to van't Hoff enthalpies equals 120, which is strong evidence that this collapse process is, as predicted by homopolymer theories, only one-state.

Experiments by Binkert *et al.*(1991) show that the rates of sidechain motions measured by fluorescence anisotropy slow dramatically at the transition (see figure 6). Taken together, these experiments indicate that molecules can have sidechains that 'freeze' upon collapse, as they do in proteins and in the exact sidechain model, without causing the transition to be two-state. These experiments and the simple exact model imply that sidechain freezing does occur in polymer collapse, but that it need not cause two-state behaviour. These are only experimental and theoretical models, not real proteins, but they give no evidence that more realistic models would show that sidechain freezing is the principal cause of two-state cooperativity in proteins.

4. IMPLICATIONS FOR PROTEIN FOLDING KINETICS AND COMPACT DENATURED STATES

How we view the forces causing cooperativity determines how we understand the folding kinetics and the denatured states of proteins. For example, the SP model (Shakhnovich & Finkelstein 1989; Ptitsyn 1987; Karplus & Shakhnovich 1992) predicts that the native state goes to a denatured state that maintains the secondary structures of the native molecule, with only the sidechains freed. This was the basis for the 'molten globule' model of compact denatured states; see figure 5. On the other hand, the HC model transition goes from native to compact denatured states with a much broader ensemble of backbone and sidechain conformations, ensembles of secondary structures and hydrophobic clusters, radii slightly larger than native, and bimodal light scattering $P(r)$ curves (Lattman *et al.* 1994), all of which are strongly dependent on the amino acid sequence and external conditions; see figure 7.

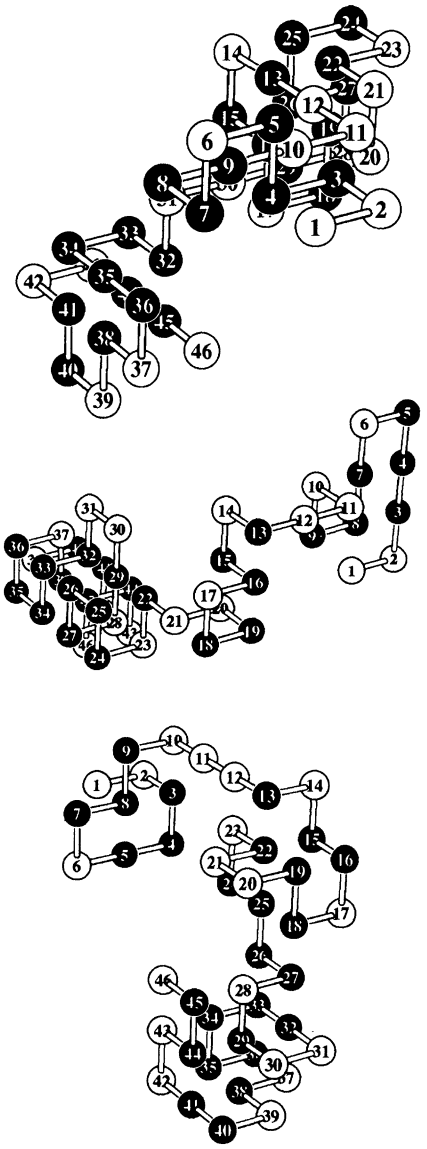


Figure 7. The hydrophobic core model implies that compact denatured states are sometimes broad ensembles of different backbone configurations, depending on sequence and external conditions. These three-dimensional HP lattice model conformations represent three different chain configurations in the ensemble of a compact denatured state of a protein. (From Lattman *et al.* 1994.)

Similarly, folding kinetics is dependent on the nature of cooperativity. Whereas models based on helix-coil-like cooperativity lead to kinetics in which isolated secondary structures are predicted to form as the earliest steps (Ptitsyn *et al.* 1972; Karplus & Weaver, 1976, 1994; Kim & Baldwin 1982; Baldwin 1989),

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

hydrophobic core cooperativity models imply that folding is driven by the hydrophobic zipping and assembly process, which causes concurrent secondary structure formation (Dill *et al.* 1993; Fiebig & Dill 1993; Lattman *et al.* 1994). This is consistent with data showing concurrent collapse and some secondary structure formation in early stages of folding (Gilmanshin & Ptitsyn 1987; Semisotnov *et al.* 1987; Briggs & Roder 1992; Elöve *et al.* 1992; Serrano *et al.* 1992; Jennings & Wright 1993; Itzhaki *et al.* 1994; Nishii *et al.* 1994, reviewed in Barrick & Baldwin 1993).

5. CONCLUSIONS

We have discussed three possible models for the two-state cooperativity in the folding of globular proteins. The limitations of the helix-coil model of biopolymer cooperativity for the basis of globular protein folding are that: (i) helical propensities are weak in aqueous solution (Chakrabartty *et al.* 1991, 1994; Scholtz *et al.* 1991; Scholtz & Baldwin 1992); (ii) helical propensities cannot account for sheet proteins; (iii) helix-coil transitions give only one-state cooperativity; and (iv) evidence is mounting for helix formation concurrent with hydrophobic clustering in the first stages of folding kinetics (Dill *et al.* 1995). The limitations of the sidechain packing model are that: (i) sidechains do not have strong packing preferences; (ii) the transition is probably one-state; and (iii) the predicted molten globule is more rigid and native-like, and less dependent on external conditions than are indicated by experiments (reviewed in Dill & Shortle 1991; Shortle *et al.* 1992; Shortle 1993). We believe the simplest basis for understanding protein cooperativity is in terms of hydrophobic and polar sequences that can form good hydrophobic cores. The driving force (non-polar association in water) is strong, and the folding kinetics of collapse accompanied by secondary structure formation is consistent with experiments and the secondary structure length distribution of proteins in the PDB (Chan & Dill 1990).

We thank the National Institutes of Health for financial support.

REFERENCES

Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. & Weng, J. 1987 Protein data bank. In *Crystallographic databases – information content, software systems, scientific applications* (ed. F.H. Allen, G. Bergerhoff & R. Seivers), pp.107–132. Bonn, Cambridge and Chester: Data Commission of the International Union of Crystallography.

Artymiuk, P.J., Blake, C.C., Grace, D.E., Oatley, S.J., Phillips, D.C. & Sternberg, M.J. 1979 Crystallographic studies of the dynamic properties of lysozyme. *Nature, Lond.* **280**, 563–568.

Baldwin, R.L. 1989 How does protein folding get started? *Trends Biochem. Sci.* **14**, 291–294.

Baldwin, E.P., Hajiseyedjavadi, O., Baase, W.A. & Matthews, B.W. 1993 The role of backbone flexibility in the accomodations of variants that repack the core of T4 lysozyme. *Science, Wash.* **262**, 1715–1718.

Barrick, D. & Baldwin, R.L. 1993 Stein and Moore Award address. The molten globule intermediate of apomyoglobin and the process of protein folding. *Protein Sci.* **2**, 869–876.

Bashford, D., Chothia, C. & Lesk, A.M. 1987 Determinants of a protein fold. Unique features of the globin amino acid sequences. *J. molec. Biol.* **196**, 199–216.

Behe, M.J., Lattman, E.E. & Rose, G.D. 1991 The protein-folding problem: the native fold determines packing, but does packing determine the native fold? *Proc. natn. Acad. Sci. U.S.A.* **88**, 4195–4199.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F. Jr, Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasumi, M. 1977 The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. molec. Biol.* **112**, 535–542.

Binkert, Th., Oberreich, J., Meewes, M., Nyffenegger, R. & Ricka, J. 1991 Coil-globule trasition of poly(*N*-isopropylacrylamide): a study of segment mobility by fluorescence depolarization. *Macromolecules* **24**, 5806–5810.

Briggs, M.S. & Roder, H. 1992 Early hydrogen-bonding events in the folding reaction of ubiquitin. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2017–2021.

Bromberg, S. & Dill, K.A. 1994 Sidechain entropy and packing in proteins. *Protein Sci.* **3**, 997–1009.

Cann, J.R. 1970 *Interacting macromolecules: the theory and practice of their electrophoresis, ultracentrifugation, and chromatography*. New York: Academic Press.

Chaffotte, A.F., Guillou, Y. & Goldberg, M.E. 1992 Kinetic resolution of peptide bond and side chain far-UV circular dichroism during the folding of hen egg white lysozyme. *Biochemistry* **31**, 9694–9702.

Chakrabartty, A., Schellman, J.A. & Baldwin, R.L. 1991 Large differences in the helix propensities of alanine and glycine. *Nature, Lond.* **351**, 586–588.

Chakrabartty, A., Kortemme, T. & Baldwin, R.L. 1994 Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3**, 843–852.

Chan, H.S. & Dill, K.A. 1989 Compact polymers. *Macromolecules* **22**, 4559–4573.

Chan, H.S. & Dill, K.A. 1990 Origins of structure in globular proteins. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6388–6392.

Chan, H.S. & Dill, K.A. 1991 'Sequence space soup' of proteins and copolymers. *J. chem. Phys.* **95**, 3775–3787.

Chan, H.S., Dill, K.A. & Shortle, D. 1992 Statistical mechanics and protein folding. In *Princeton lectures on biophysics* (ed. W. Bialek), pp. 69–173. Singapore: World Scientific.

Creighton, T.E. 1986 Detection of folding intermediates using urea-gradient electrophoresis. *Meth. Enzymol.* **131**, 156–172.

Creighton, T.E. 1994 The energetic ups and downs of protein folding. *Nature Struct. Biol.* **1**, 135–138.

de Gennes, P.-G. 1975 Collapse of a polymer chain in poor solvents. *J. Phys. Lett., Paris* **36**, L55–L57.

Dill, K.A. 1985 Theory for the folding and stability of globular proteins. *Biochemistry* **24**, 1501–1509.

Dill, K.A. & Shortle, D. 1991 Denatured states of proteins. *A. Rev. Biochem.* **60**, 795–825.

Dill, K.A., Fiebig, K.M. & Chan, H.S. 1993 Cooperativity in protein folding kinetics. *Proc. natn. Acad. Sci. U.S.A.* **90**, 748–52.

Dill, K.A. & Stigter, D. 1994 Modeling protein stability as heteropolymer collapse. *Adv. Protein Chem.* (In the press.)

Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D. & Chan, H.S. 1995 Principles of protein folding – a perspective from simple exact models. *Protein Sci.* (In the press.)

Elöve, G.A., Chaffotte, A.F., Roder, H. & Goldberg, M.E.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

1992 Early steps in cytochrome *c* folding probed by time-resolved circular dichroism and fluorescence spectroscopy. *Biochemistry* **31**, 6876–6883.

Eriksson, A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P. & Matthews, B.W. 1992 Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science, Wash.* **255**, 178–183.

Feng, Y.Q., Wand, A.J. & Sligar, S.G. 1994 Solution structure of apocytochrome b562. *Nature Struct. Biol.* **1**, 30–35.

Fiebig, K.M. & Dill, K.A. 1993 Protein core assembly processes. *J. chem. Phys.* **98**, 3475–3487.

Frauenfelder, H., Petsko, G.A. & Tsernoglou, D. 1979 Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature, Lond.* **280**, 558–563.

Fujishige, S., Kubota, K. & Anto, I. 1989 Phase transition of aqueous solutions of poly(*N*-isopropylacrylamide) and poly(*N*-isopropylmethacrylamide). *J. phys. Chem.* **93**, 3311–3313.

Gilmanshin, R.I. & Ptitsyn, O.B. 1987 An early intermediate of refolding α-lactalbumin forms within 20 ms. *FEBS Lett.* **223**, 327–329.

Grosberg, A.Yu & Khokhlov, A.R. 1987 Physics of phase transitions in solutions of macromolecules. *Soviet Sci. Rev.* **A8**, 147–258.

Gurd, F.R. & Rothgeb, T.M. 1979 Motions in proteins. *Adv. Protein Chem.* **33**, 73–165.

Hooke, S.D., Radford, S.E. & Dobson, C.M. 1994 The refolding of human lysozyme: a comparison with the structurally homologous hen lysozyme. *Biochemistry* **33**, 5867–5876.

Hughson, F.M., Barrick, D. & Baldwin, R.L. 1991 Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry* **30**, 4113–4118.

Itzhaki, L.S., Evans, P.A., Dobson, C.M. & Radford, S.E. 1994 Tertiary interactions in the folding pathway of hen lysozyme – kinetic studies using fluorescent probes. *Biochemistry* **33**, 5212–5220.

Jennings, P.A. & Wright, P.E. 1993 Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science, Wash.* **262**, 892–896.

Karplus, M. & McCammon, J.A. 1981 The internal dynamics of globular proteins. *CRC Crit. Rev. Biochem.* **9**, 293–349.

Karplus, M. & Shakhnovich, E. 1992 Protein folding: theoretical studies of thermodynamics and dynamics. In *Protein folding* (ed. T. Creighton), pp. 127–195. New York: Freeman.

Karplus, M. & Weaver, D.L. 1976 Protein-folding dynamics. *Nature, Lond.* **260**, 404–406.

Karplus, M. & Weaver, D.L. 1994. Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci.* **3**, 650–668.

Keefe, L.J., Sondek, J., Shortle, D. & Lattman, E.E. 1993 The alpha aneurism: a structural motif revealed in an insertion mutant of staphylococcal nuclease. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3275–3279.

Lattman, E.E., Fiebig, K.M. & Dill, K.A. 1994 Modeling compact denatured states of proteins. *Biochemistry* **33**, 6158–6166.

Lau, K.F. & Dill, K.A. 1989 A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules* **22**, 3986–3997.

Lau, K.F. & Dill, K.A. 1990 Theory for protein mutability and biogenesis. *Proc. natn. Acad. Sci. U.S.A.* **87**, 638–642.

Lifson, S. & Roig, A. 1961 On the theory of helix-coil transition in polypeptides. *J. chem. Phys.* **34**, 1963–1974.

Lim, W.A. & Sauer, R.T. 1991 The role of internal packing interactions in determining the structure and stability of a protein. *J. molec. Biol.* **219**, 359–376.

Lumry, R., Biltonen, R. & Brandts, J.F. 1966 Validity of the 'two-state' hypothesis for conformational transitions of proteins. *Biopolymers* **4**, 917–944.

Matthews, B.W. 1987 Genetic and structural analysis of the protein stability problem. *Biochemistry* **26**, 6885–6888.

Matthews, B.W. 1993 Structural and genetic analysis of protein stability. *A. Rev. Biochem.* **62**, 139–160.

McRee, D.E., Redford, S.M., Getzoff, E.D., Lepock, J.R., Hallewell, R.A. & Tainer, J.A. 1990 Changes in crystallographic structure and thermostability of a Cu, Zn superoxide dismutase mutant resulting from the removal of a buried cysteine. *J. biol. Chem.* **265**, 14234–14241.

Meewes, M., Ricka, J., de Silva, R., Nyffenegger, R. & Binkert, Th. 1991 Coil-globule transition of poly(*N*-isopropylacrylamide). A study of surfactant effects by light scattering. *Macromolecules* **24**, 5811–5816.

Moore, M.A. 1977 Theory of the polymer coil-globule transition. J. Phys. A **10**, 305–314.

Nishii, I., Kataoka, M., Tokunaga, F. & Goto, Y. 1994 Cold denaturation of the molten globule states of apomyoglobin and a profile for protein folding. *Biochemistry* **33**, 4903–4909.

Novotny, J., Bruccoleri, R., Karplus, M. 1984 An analysis of incorrectly folded protein models. Implications for structure predictions. *J. molec. Biol.* **177**, 787–818.

Peng, Z. & Kim, P.S. 1994 A protein dissection study of a molten globule. *Biochemistry* **33**, 2136–2141.

Poland, D.C. & Scheraga, H.A. 1970 *Theory of the helix-coil transition*. New York: Academic Press.

Post, C.B. & Zimm, B.H. 1979 Internal condensation of a single DNA molecule. *Biopolymers* 18, 1487–1501.

Privalov, P.L. 1979 Stability of proteins: small globular proteins. *Adv. Protein Chem.* **33**, 167–241.

Privalov, P.L. & Gill, S.J. 1988 Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* **39**, 191–234.

Ptitsyn, O.B., Kron, A.K. & Eizner, Yu. Ye. 1968 The models of the denaturation of globular proteins. I. Theory of globule-coil transitions in macromolecules. *J. Polymer Sci.* C **16**, 3509–3517.

Ptitsyn, O.B., Lim, V.I. & Finkelstein, A.V. 1972 *Analysis and simulation of biochemical systems. Proceedings of the Eighth FEBS Meeting* (ed. B. Hess & H.C. Hemker), pp. 421–431. Amsterdam: North-Holland.

Ptitsyn, O.B. 1987 Protein folding: hypotheses and experiments. *J. Protein Chem.* **6**, 273–293.

Richards, F.M. 1974 The interpretation of protein structures: total volume, group volume distributions and packing density. *J. molec. Biol.* **82**, 1–14.

Richards, F.M. 1977 Areas, volumes, packing and protein structure. *A. Rev. Biophys. Bioeng.* **6**, 151–176.

Richards, F.M. & Lim, W.A. 1994 An analysis of packing in the protein folding problem. *Q. Rev. Biophys.* **26**, 423–498.

Ricka, J., Meewes, M., Nyffenegger, R. & Binkert, Th. 1990 Intermolecular and intramolecular solubilization: collapse and expansion of a polymer chain in surfactant solutions. *Phys. Rev. Lett.* **65**, 657–660.

Sanchez, I.C. 1979 Phase transition behaviour of the isolated polymer chain. *Macromolecules* **12**, 980–988.

Schellman, J.A. 1958 The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. *J. phys. Chem.* **62**, 1485–1494.

Scholtz, J.M., Qian, H., York, E.J., Stewart, J.M. & Baldwin, R.L. 1991 Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers* **31**, 1463–1470.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Scholtz, J.M. & Baldwin, R.L. 1992. The mechanism of α-helix formation by peptides. *A. Rev. Biophys. biomolec. Struct.* **21**, 95–118.

Semisotnov, G.V., Rodionova, N.A., Kutyshenko, V.P., Ebert, B., Blanck, J. & Ptitsyn, O.B. 1987 Sequential mechanism of refolding of carbonic anhydrase B. *FEBS Lett.* **224**, 9–13.

Serrano, L., Matouschek, A. & Fersht, A.R. 1992 The folding of an enzyme. III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure. *J. molec. Biol.* **224**, 805–818.

Shakhnovich, E.I. & Finkelstein, A.V. 1989 Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers* **28**, 1667–1680.

Shortle, D., Chan, H.S. & Dill, K.A. 1992 Modeling the effects of mutations on the denatured states of proteins. *Protein Sci.* **1**, 201–215.

Shortle, D. 1993 Denatured states of proteins and their roles in folding and stability. *Curr. Opin. Struct. Biol.* **3**, 66–74.

Singh, J. & Thornton, J.M. 1990 SIRIUS. An automated method for the analysis of the preferred packing arrangements between protein groups. *J. molec. Biol.* **211**, 595–615.

Singh, J. & Thornton, J.M. 1992 *Atlas of protein side-chain interactions*. New York: Oxford University Press.

Sosnick, T.R., Mayne, L., Hiller, R. & Englander, S.W. 1994 The barriers in protein folding. *Nature Struct. Biol.* **1**, 149–156.

Stanley, H.E. 1971, 1987 *Introduction to phase transitions and critical phenomena*. New York: Oxford University Press.

Swindells, M.B. & Thornton, J.M. 1993 A study of structural determinants in the interleukin-1 fold. *Protein Engng* **6**, 711 –715.

Tiktopulo, E.I., Bychkova, V.E., Ricka, J. & Ptitsyn, O.B. 1994 Cooperativity of the coil-globule transition in a homopolymer – microcalorimetric study of poly(*N*-isopropylacrylamide). *Macromolecules* **27**, 2879–2882.

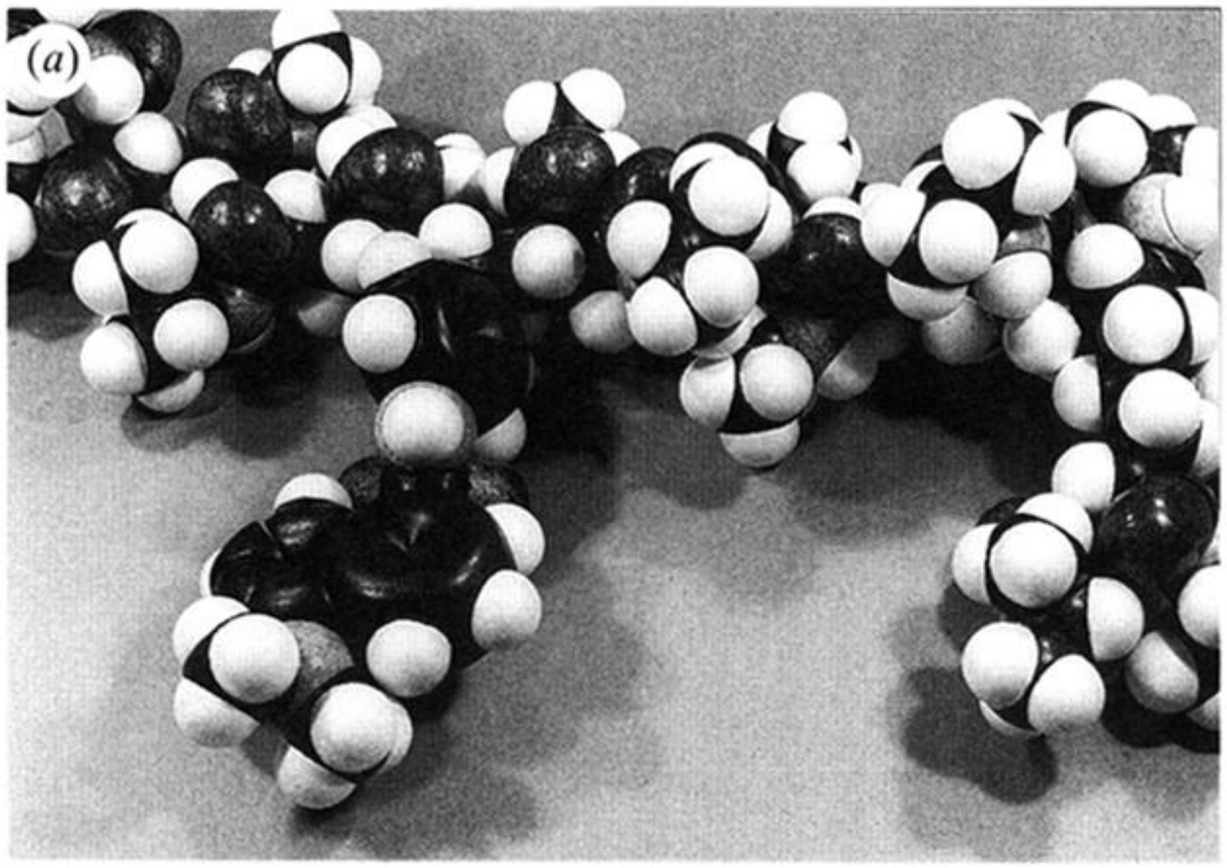
Uversky, V.N. 1993 Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry* **32**, 13288–13298.

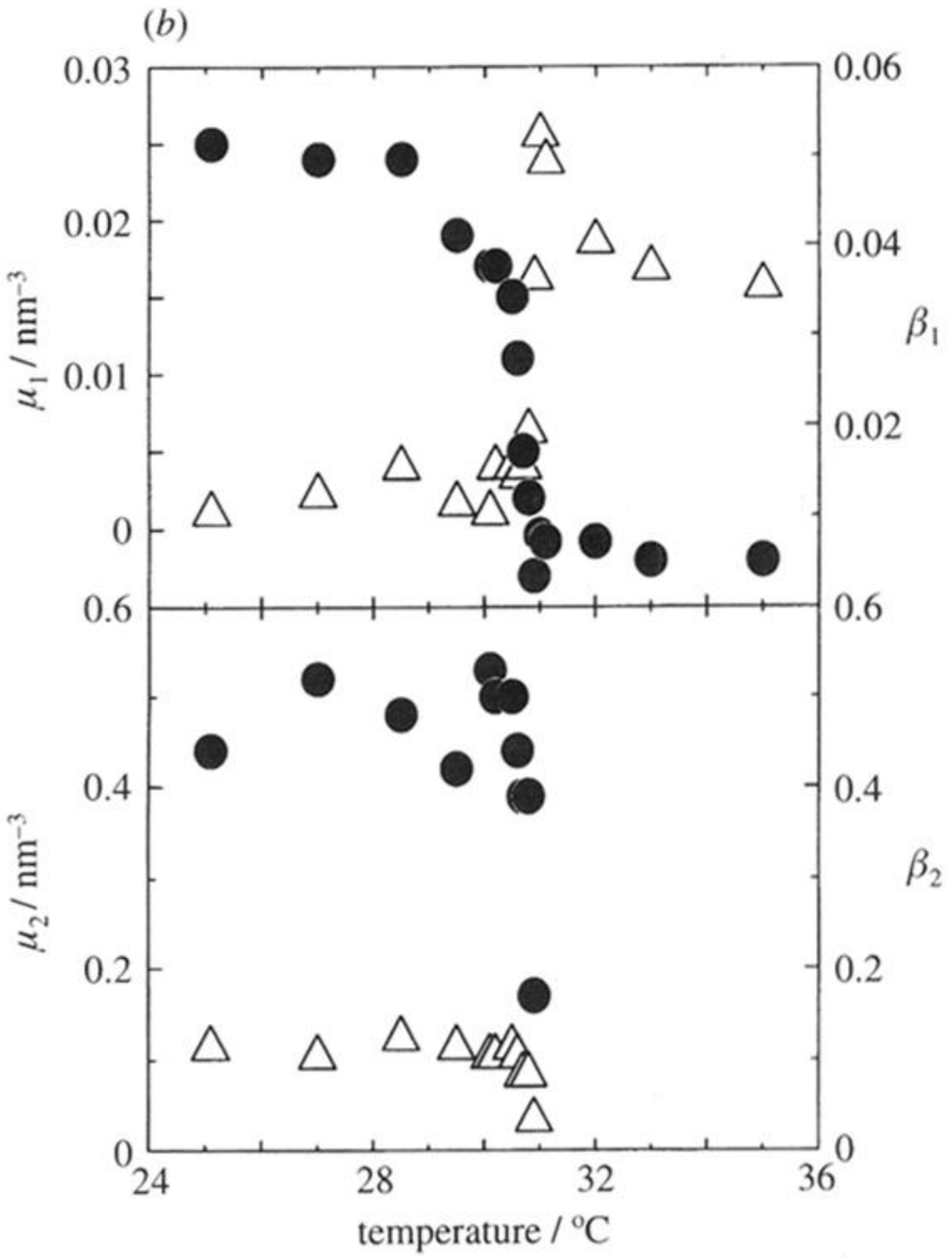
Varadarajan, R. & Richards, F.M. 1992 Crystallographic structures of ribonuclease S variants with nonpolar substitution at position 13: packing and cavities. *Biochemistry* **31**, 12315–12327.

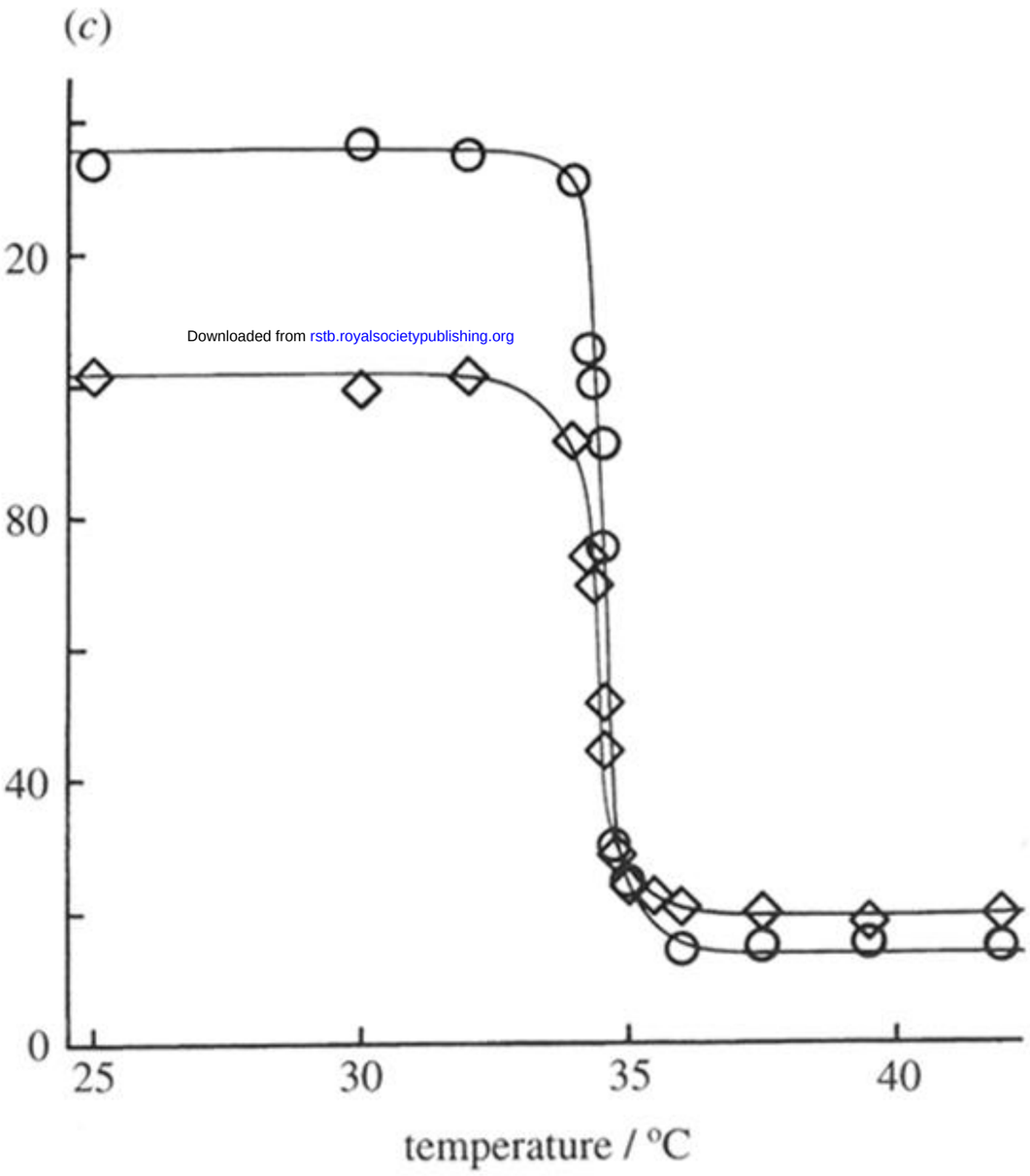
Wuthrich, K. & Wagner, G. 1978 Internal motions in globular proteins. *Trends Biochem. Sci.* **3**, 227–230.

Zimm, B.H. & Bragg, J.K. 1959 Theory of the phase transition between helix and random coil in polypeptide chains. *J. chem. Phys.* **31**, 526–535.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES







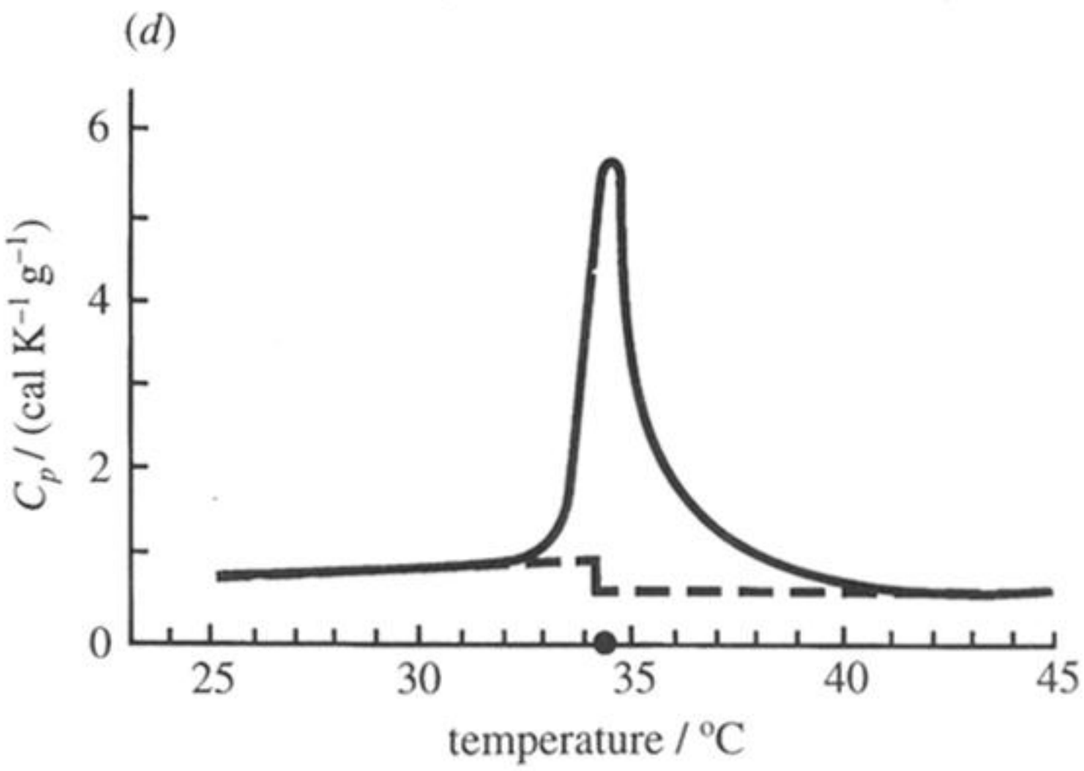


Figure 6. Collapse of poly(*N*-isopropylacrylamide) (PNIPAM) homopolymers in aqueous solutions with small amount of surfactant to suppress aggregration. (*a*) Space-filling model of a section of the polymer. (*b*) Homopolymer collapse freezes out the motions of the sidechains. Time-resolved measurements of fluorescence polarization anisotropy are used to monitor sidechain motions due to the temperature-induced conformational transition. The plots show reduced re-orientational relaxation rates μ_i and their amplitudes β_i versus temperature: (left, filled circles) μ_1, (right, open triangles) β_1, reflecting backbone motions; (left, filled circles) μ_2, (right, open triangles) β_2, reflecting mainly local sidechain motions. Both types of motions undergo a 'freezing' transition at the same temperature around 31 °C. The chains have approximately 3100 monomer units. (*a*) and (*b*) are reproduced from Binkert *et al.* (1991). (*c*) Hydrodynamic radius (open diamonds) and radius of gyration (open circles) of PNIPAM as a function of temperature (data from Meewes *et al.* 1991). (*d*) Temperature dependence of the partial specific heat capacity of PNIPAM. Dashed lines show the extrapolations of heat capacity curves from low and high temperatures to the middle of the transition region. Reproduced from Tiktopulo *et al.* (1994). The number of monomers per chain in experiments (*c*) and (*d*) are approximately 62000.