Looking toward the future

The current round of structural data has provided many answers but also raised important questions. The field faces clear challenges of both a functional and a structural nature.

From a structural standpoint, some questions are basic. What are the important considerations in designing a rational screen for crystallization of membrane proteins (Michael Weiner, University of Virginia)? What is the nature of the intricate latticework in the vestibule of the Ca^{2+}-release channel (Irina Seryshev, Baylor College of Medicine)?

Many questions discussed at this meeting remain largely unaddressed using structural techniques. How is a signal transduced from sensing domains to the pore (Horn; Richard Aldrich, Stanford University)? How do auxiliary proteins interact with the channels (John Hanrahan, McGill University)? What determines the rate of ion throughput (David Clapham, Children's Hospital of Boston)? How will the ClC Cl^- channels, which contain two subunits and two pores, be organized (Merritt Maduke, Brandeis University)?

From a functional perspective, the emerging structures provoke new questions. How do caps on channels affect channel behavior? What role do cofactors play? How do associated subunits and proteins modulate function? And finally, with sufficient structural information, how can ion channels be engineered to work for us (Hagan Bayley, Texas A&M Health Science Center)?

Perhaps because it has waited so long, ion channel biophysics is enjoying the new complexities that structural knowledge brings. As was clear throughout the conference, the hard-sought integration of structural and functional approaches in this field greatly enhances our understanding of how channels work at a molecular level.

Lise Heginbotham is in the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, USA. email: lise.heginbotham@yale.edu

*Ion Channel Structures & Function, July 16–18, 1999; Yale University, New Haven, Connecticut, USA. Organized by Steve A.N. Goldstein.

1. Gulbis, J.M., Mann, S. & Mackinnon, R. *Cell* **97**, 943–952 (1999).
2. Armstrong, C. & Bezanilla, F. *J. Gen. Physiol.* **70**, 567–590 (1977).
3. Miyazawa, A., Fujiyoshi, Y., Stowell, M. & Unwin, N. *J. Mol. Biol.* **288**, 765–786 (1999).
4. Chang, G., Spencer, R.H., Lee, A.T., Barchay, M.T. & Rees, D.C. *Science* **282**, 2220–2226 (1998).
5. Doyle, D.A. *et al. Science* **280**, 69–77 (1998).
6. Kreusch, A., Pfaffinger, P.J., Stevens, C.F. & Choe, S. *Nature* **392**, 945–948 (1998).
7. Bixby, K.A. *et al. Nature Struct. Biol.* **6**, 38–43 (1999).

Protein folding from 1961 to 1982

Robert L. Baldwin

It was 1961. Anfinsen[1] had shown that the correct disulfide bonds of ribonuclease A (RNaseA) form spontaneously upon oxidation in air, driven by the free energy of folding. The central dogma of molecular biology, DNA→RNA→Proteins, was being analyzed step by step and the factors (especially enzymes) needed to make each step take place *in vitro* were being discovered. Anfinsen's work illuminated the point that the chemical process of RNA translation yields an unfolded polypeptide chain and the chain must fold before it becomes a biologically active protein. Previously some proteins, such as RNase A and chymotrypsinogen had been shown to denature and renature reversibly when their disulfide bonds were left intact. But how are the disulfide bonds formed? Most proteins containing disulfide bonds unfold when these bonds are reduced to the form present in the nascent polypeptide chain. Straightforward calculations had shown that the probability of forming the correct bonds in the fully unfolded state is close to zero for a protein such as RNase A, whose eight Cys residues form four disulfide bonds. In addition to demonstrating that folding drives disulfide bond formation, Anfinsen found that some incorrect disulfide bonds are formed initially but are reduced and reshuffled, provided a suitable reducing agent is present; moreover, an enzyme, protein disulfide isomerase, achieves this reshuffling *in vivo*[2,3]. Through this work, Anfinsen connected the new science of molecular biology with the established field of protein chemistry, and he focused the attention of scientists in both fields on solving the mechanism of protein folding.

In 1961 the structure of one protein, myoglobin, was known at 6 Å resolution and its 2 Å structure would soon be published. The structural determinations of hemoglobin and hen lysozyme were well under way. Regarding the energetics of folding, Kauzmann had shown in 1959 that the hydrophobic interaction would be a dominant factor, but in 1961 this was a revolutionary concept, not yet widely accepted. Schellman in 1955 and later also Klotz had shown that model structures formed in water by amide hydrogen bonds have at best marginal stability. The time was right to begin the task of isolating and characterizing folding intermediates: but did they exist?

The two-state model of folding

The concept that protein denaturation is simply unfolding of a giant semi-crystalline molecule had been suggested by Wu in 1931, but it remained for Kauzmann and his students to validate the concept in the early 1950s with kinetic studies of the irreversible denaturation of several proteins. Brandts proceeded to characterize the thermodynamics of the reversible thermal unfolding of chymotrypsinogen[4]. After obtaining very precise data for the equilibrium between the native and unfolded forms as a function of temperature, pH, and chloride ion concentration, he interpreted the data using a two-state model, without folding intermediates. He obtained self-consistent results although the resulting thermodynamics (which evidently were dominated by the hydrophobic interaction) had some curious features, such as a temperature of maximum stability at 10 °C. His success in applying the two-state model led to a careful analysis of its meaning, and also to development of thermodynamic criteria for testing it, which were outlined by Lumry, Biltonen and Brandts[5]. A different test of the two-state model was applied to the thermal unfolding of RNase A by monitoring the unfolding process using four different probes[6] that responded to widely different molecular properties,

such as secondary *versus* tertiary structure and overall dimensions. The four unfolding curves could be superimposed on each other as would be expected if there were no populated folding intermediates, whose properties could in principal be measured differently by the four probes.

Tanford[7,8] characterized the guanidinium chloride (GdmCl)-induced unfolding reactions of several small proteins, found they were two-state reactions, and demonstrated the usefulness of GdmCl in studying protein folding reactions. A concentration of 6 M GdmCl thoroughly unfolds water soluble proteins, and it provides a defined starting point for analyzing refolding reactions. Privalov[9] undertook the monumental task of building a sensitive and accurate scanning microcalorimeter for determining the thermodynamics of protein folding reactions and of testing the two-state model.

The domain as a unit of folding

In 1967 Edelman observed domain structure in the first amino acid sequence of an antibody[10]. The observation of intracistronic complementation, in which inactive mutants at opposite ends of the β-galactosidase gene complement each other to restore activity, led Goldberg to study its molecular mechanism in 1969. He deduced that segments of the giant polypeptide chain fold independently of each other to form 'globules'[11]. In the following year Taniuchi demonstrated that RNase A, which has only 124 amino acids, is unable to fold correctly and form the native disulfide bonds when the four C-terminal amino acids are deleted[12]. Thus, units smaller than the entire RNase A domain do not achieve stably the native fold. Then in 1973 Wetlaufer observed domain structures, namely contiguous segments of polypeptide chain folded on themselves as separate units and apparently connected by flexible hinges, in the X-ray structures of proteins[13]. These observations led to a general acceptance of the concept that domains act as units of folding. In 1981 Lesk and Rose[14] noted that domains can often be divided into subdomains (contiguous folded segments that are similar domains but are not separated from each other by hinge-like connections) and these can be further subdivided.

Levinthal's thought experiment

If there are no folding intermediates, as implied by the success of the two-state model, and if instead a protein must fold by a random search of all possible conformations, how long should the folding process take? Levinthal's answer[15] in 1968, longer than the lifetime of the universe, had a stunning impact. It led many scientists to believe there must be folding pathways and folding intermediates. Levinthal himself began the attempt to fold proteins with the aid of computer modeling.

A feature of Levinthal's calculation that attracted little attention in 1968 but became important later was his conclusion that a protein should fold by the fastest pathway and this need not lead to the thermodynamically most stable structure. He argued that evolution produces proteins that commonly do fold to their most stable structures but that evolution might also produce other types of proteins in response to a suitable biological stress. His reasoning was validated by later studies of some serpins, which exist in two structural forms with different thermodynamic stabilities, and of some extracellular proteases (such as α-lytic protease and subtilisin) which can fold only with the aid of a prosequence that is cleaved off to form the mature protease after the proenzyme has folded.

The search for folding intermediates

Equilibrium measurements of unfolding reactions failed to detect the intermediates suggested by Levinthal's calculation. Was there a missing variable that could be manipulated to make them appear? In 1971 two laboratories reported that fast reaction kinetics revealed folding intermediates. Ikai and Tanford[16] reported biphasic kinetics of unfolding and refolding cytochrome *c* in a stopped-flow study of its GdmCl-induced unfolding transition. Tsong and Baldwin[17] reported a fast phase (on the order of 1–10 ms) in addition to the major slow unfolding reaction (on the order of s), seen in a temperature-jump study of the thermal unfolding of RNase A at low pH. Pohl[18] had earlier reported kinetic data on the folding and unfolding reactions of some small proteins, and he found only two-state behavior. He used a slow temperature-jump method, however, whose long dead time did not permit detection of minor fast reactions.

These initial results demonstrated kinetic complexity in the folding and unfolding reactions of two proteins (more followed) but they left unresolved some basic questions. Were the intermediates on-pathway or abortive? Ikai and Tanford tackled the problem by analyzing quantitatively the kinetic folding and unfolding amplitudes fitted to simple models with one or two kinetic intermediates, and they applied their analysis to hen lysozyme[19] as well as cytochrome *c*[20]. Their analysis suggested off-pathway intermediates, but other models were able to produce alternative interpretations[21]. After 1973 Tanford left the protein folding problem in order to study membrane proteins.

Why should fast kinetics provide evidence for folding intermediates when equilibrium tests showed only two-state folding? A possible answer was provided by a nucleation dependent folding model[22], similar to the one used by Pörschke and Eigen[23] to describe the unfolding and refolding kinetics of short RNA double helices. In this model local unfolding events occur in a fast time range while overall unfolding is a steady state reaction occurring in a much slower time range.

Are the kinetic folding intermediates genuine structural intermediates or do they arise only from kinetic complexity already present in the unfolded state? This last question rapidly became the most pressing because Garel and Baldwin[24] reported in 1973 that unfolded RNase A contains two kinetically distinct forms in slow equilibrium with each other: a minor (~20%) fast folding form and a major (~80%) slow folding form. Both unfolded forms refold to produce the same native protein.

Proline isomerization

Brandts and coworkers[25] proposed in 1975 that the fast folding and slow folding forms of RNase A are produced by slow *cis-trans* isomerization about proline peptide bonds occurring in the unfolded state. This was plausible for three reasons: (i) proline isomerization was known to be a slow reaction, occurring in a suitable time range; (ii) the *cis* proline isomer is moderately populated in peptides (10–20%, depending on sequence), unlike the *cis* isomers of other peptide bonds; and (iii) native RNase A has two *cis* proline residues that would be expected to isomerize chiefly to the *trans* isomer in the unfolded state, thus producing slow folding species.

It proved difficult to decide by refolding experiments whether or not proline isomerization explains the two unfolded forms of RNase A because the effects of proline isomerization and structural folding intermediates are intermingled. Experiments in the unfolded state pro-

vided a clear answer, however. In 1978 Schmid and Baldwin found that interconversion between the two unfolded forms is catalyzed by strong acid[26] as expected for proline isomerization, and other kinetic properties of the two reactions also matched.

Structural folding intermediates

Although fast kinetics demonstrated that protein folding is not a simple two-state reaction, these experiments did not solve the problem of isolating and characterizing the structures of folding intermediates. Several approaches were tested in the 1970s and three of them were successful in either isolating or labeling folding intermediates. In 1974 Creighton[27] used the coupling between folding and disulfide bond formation to isolate covalently trapped disulfide intermediates in the refolding/reoxidation of bovine pancreatic trypsin inhibitor (BPTI). This small protein has three disulfide bonds and only 58 residues. Specific intermediates could be isolated by chromatographic separation after blocking the unreacted Cys residues with iodoacetate at intermediate stages in reoxidation.

Proline isomerization was used to trap a folding intermediate in RNase A refolding at 0 °C[28]. Proline isomerization has a large activation enthalpy (20 kcal mol^{-1}) and the reaction becomes quite slow (~10^3 s) at 0 °C. The question was: how far can folding proceed before it is blocked by the presence of a non-native proline isomer? The astonishing answer was that a native-like intermediate (I_N) is formed that even has RNase catalytic activity[29], although it still contains at least one non-native proline isomer. Also surprising was the rate of proline isomerization in I_N (which converts I_N to native RNase A): it is 30-fold faster in I_N than in the unfolded state.

Hydrogen exchange was used to label a different folding intermediate[30] of RNase A. Hvidt[31] had shown in 1964 that hydrogen exchange in native proteins commonly occurs by an interconversion between closed (exchange-resistant) and open (exchange-susceptible) forms, and the rate of exchange depends directly on the equilibrium constant of the interconversion reaction. By using ^{3}H-labeled unfolded RNase A as the starting material for refolding, and by setting up a competition between exchange-out ($^3H \rightarrow {}^1H$) and refolding, it was possible to detect an early folding intermediate. In unfolded peptides the rate of exchange-out is base catalyzed, and so one can choose a pH where exchange-out is much faster than refolding; consequently ^{3}H label is retained in the folded protein only if an early folding intermediate is formed rapidly that protects some label against exchange-out. The same type of experiment is easier to interpret when pulse labeling[32] is used. Much later, when 2D-NMR became available, it was possible to make this experiment using $^1H \rightarrow {}^2H$ exchange to determine which peptide NH protons are protected in a folding intermediate.

Monitoring protein refolding by real time NMR was also tested in 1978 with the refolding reaction of RNase S at 10 °C, pH 2 (ref. 33). One result of interest was found although most of the data could be fit to two-state folding, presumably because the probe was sensitive chiefly to tertiary rather than secondary structure. The chemical shifts of the ε1 CH protons of the four histidine residues were measured; they are well-resolved in native RNase S even at 270 MHz, which was the field strength available at that time. RNase S contains two polypeptide segments that separate upon unfolding; the S-protein (104 residues) and the S-peptide (20 residues). The interesting result was that the S-peptide is partly folded at 10 °C, pH 2, according to the chemical shift of His 12.

How does protein folding get started?

Two extreme views have held the imagination of workers in this field. The first is that folding is hierarchic and begins by forming local backbone structures that persist until the native structure emerges. The other is that folding begins by forming a tertiary interaction, a cluster of side chains drawn together from residues distant in sequence. In 1965 Perutz, Kendrew and Watson[34] postulated that proline is a helix termination signal, on the basis of the X-ray structures of myoglobin and hemoglobin and of the amino acid sequences of numerous globins. If helix termination signals are encoded locally rather than by tertiary interactions, then folding should begin locally by a hierarchic mechanism. Later Schellman[35] observed that two sequence motifs are found commonly at the C-termini of helices, and she proposed that such motifs also act as helix termination signals.

If folding is hierarchic, then fragments of proteins should fold detectably. The simplest experiment tests for helix formation using circular dichroism (CD), and this was tried in three laboratories: with a helix-containing fragment of staphylococcal nuclease (Snase)[36], with the entire sequence of myoglobin contained in three cyanogen bromide fragments[37], and with N-terminal peptides of RNase A[38], either the C-peptide (residues 1–13) or S-peptide (residues 1–20). These peptides include residues 3–12, which code for a protein helix. Only the RNase A peptides showed any CD signal and this occurred only at low temperatures, near 0 °C. The myoglobin and Snase experiments were performed at 25 °C.

The positive C-peptide result was confirmed and extended by Bierzynski *et al.*[39] 11 years later, after an NMR study[33] had shown that S-peptide is partly folded at 10 °C. One reason for the long lag was that helix propensities had been measured in random copolymers by the host-guest technique[40] and the results indicated that all short peptides should fail to show detectable helix formation in water if the helices are stabilized only by helix propensities. The later C-peptide results[39] showed that short peptide helices can be stable in water because specific side chain interactions, especially ones involving charged side chains, can help to stabilize them in addition to helix propensities. The widespread belief that short helices are undetectable in water began to be replaced by the view that helices exist on the borderline of being detectable.

Anfinsen thought that hierarchic folding was highly probable[3]. After failing to detect helix formation in a Snase fragment by CD[36], he succeeded in developing a more sensitive immunologic method[41] of detecting 'native format' in protein fragments. He fractionated polyclonal antibodies made against native Snase on columns to which specific Snase fragments were bound. Then he developed enzymatic as well as immunologic methods of quantitating complex formation between the Snase polypeptide fragment and the antibody fraction. He determined values of K_{conf}, the equilibrium constant for folding the fragment into the 'native format'; K_{conf} is the fraction of time spent by the fragment in the native format. For a 51-residue Snase fragment, he found $K_{conf} = 2 \times 10^{-4}$, meaning that the fragment has a native conformation only 0.02% of the time.

Protein misfolding

In 1961 the ability of a few proteins to unfold and refold was a mystery, but by the 1970s those proteins that failed to

refold had become the mystery. Often they were large oligomeric proteins, and it was clear that concentration-dependent aggregation was a major problem that reduces the yield of native protein. The subject was studied systematically in Jaenicke's laboratory by analyzing the mechanism of folding coupled to association for representative oligomeric proteins[42]. A finding of particular interest by Goldberg and coworkers[43] was that the yield of correctly folded tryptophanase (a tetramer) plunged toward zero at a critical urea concentration, in a manner dependent on protein concentration. Aggregates consisting of networks of partly folded protein were found at the critical urea concentration and the results were interpreted in terms of networks formed by swapping of subdomains (or even microdomains) between protein subunits. These observations of misfolding are of much interest today in understanding how chaperones prevent misfolding and how inclusion bodies arise.

The molten globule hypothesis

In 1981 Ptitsyn and coworkers[44] reported that the acid form of α-lactalbumin is compact, almost as compact as the native protein. This stable but only partly folded form of a protein had been studied extensively in Sugai's laboratory, especially by Kuwajima[45]. Ptitsyn advanced the hypothesis that this partly folded form is in fact representative of the folding intermediates that occur in the process of forming the native structure. He argued further that these stable folding intermediates shown by a few proteins share basic properties: a high content of native-like secondary structure, compact conformation, little visible tertiary structure, and non-rigid side chains. He began study of these acid forms when he realized they might be the folding intermediates he had hypothesized in 1973 in a specific proposal[46] for hierarchic folding of protein structures. The name 'molten globule' was given in 1983 by Ohgushi and Wada (after discussion with Ptitsyn) for the acid form of cytochrome *c* they had been studying. 'Molten' because the side chains are non-rigid, 'globule' because the conformation is compact. Ptitsyn's molten globule hypothesis would hold center stage in the years immediately following 1981.

In 1979 the first international conference on protein folding was held in Regensburg, Germany, organized by Jaenicke. He edited the book of papers[42] that emerged, which served as a basic handbook of protein folding for years afterward. The field exploded as new techniques of site-directed mutagenesis, 2D NMR, and rapid X-ray structure determination became available. The age of steel had arrived, and the age of silicon would follow, as folding simulations showed the need to mesh theory with experiment.

Scope of this review

Two major subjects, which are central to understanding the mechanism of protein folding, are not included here because of the limitation on length. Each of them deserves its own history. They are: development of the theory of folding, especially for predicting structure from sequence, and analysis of the relation between sequence homology and evolutionary and structural relatedness.

For consistency, I have not included references to work before 1961 or after 1982.

Robert L. Baldwin is in the Department of Biochemistry, Stanford University School of Medicine, Beckman Center, Room B-400, Stanford, California 94305-5307, USA. email: bbaldwin@cmgm.stanford.edu

1. Anfinsen, C.B., Haber, E., Sela, M. & White, Jr., F.W. *Proc. Natl. Acad. Sci. USA* **47**, 1309–1314 (1961).
2. Epstein, C.J., Goldberger, R.F. & Anfinsen, C.B. *Cold Spring Harbor Symp. Quant. Biol.* **28**, 439–449 (1963).
3. Anfinsen, C.B. *Science* **181**, 223–230 (1973).
4. Brandts, J.F. *J. Am. Chem. Soc.* **86**, 4291–4301 (1964).
5. Lumry, R., Biltonen, R. & Brandts, J.F. *Biopolymers* **4**, 917–944 (1966).
6. Ginsburg, A. & Carroll, W.R. *Biochemistry* **4**, 2159–2174 (1965).
7. Tanford, C. *Adv. Prot. Chem.* **23**, 122–282 (1968).
8. Tanford, C. *Adv. Prot. Chem.* **24**, 1–95 (1970).
9. Privalov, P.L., Khechinashvili, N.N. & Atansov, B.P. *Biopolymers* **10**, 1865–1890 (1971).
10. Waxdal, M.J., Konigsberg, W.H. & Edelman, G.M. *Cold Spring Harbor Symp. Quant. Biol.* ***32****, 53–63* (1967).
11. Goldberg, M.E. *J. Mol. Biol.* **46**, 441–446 (1969).
12. Taniuchi, H. *J. Biol. Chem.* **245**, 5459–5468 (1970).
13. Wetlaufer, D.B. *Proc. Natl. Acad. Sci. USA* **70**, 697–701 (1973).
14. Lesk, A.M. & Rose, G.D. *Proc. Natl. Acad. Sci. USA* **78**, 4304–4308 (1981).
15. Levinthal, C. *J. Chim. Phys.* **65**, 44–45 (1968).
16. Ikai, A. & Tanford, C. *Nature* **230**, 100–102 (1971).
17. Tsong, T.Y., Baldwin, R.L. & Elson, E.L. *Proc. Natl. Acad. Sci. USA* **68**, 2712–2715 (1971).
18. Pohl, F.M. *FEBS LETT* **3**, 60–64 (1969).
19. Tanford, C., Aune, K.C. & Ikai, A. *J. Mol. Biol.* **73**, 185–197 (1973).
20. Ikai, A., Fish, W.W. & Tanford, C. *J. Mol. Biol.* **73**, 165–184 (1973).
21. Hagerman, P.J. *Biopolymers* **16**, 731–747 (1977).
22. Tsong, T.Y., Baldwin, R.L. & McPhie, P. *J. Mol. Biol.* **63**, 453–475 (1972).
23. Pörschke, D. & Eigen, M. *J. Mol. Biol.* **62**, 361–381 (1971).
24. Garel, J.-R., Baldwin, R.L. *Proc. Nat. Acad. Sci. USA* **70**, 3347–3351 (1973).
25. Brandts, J.F., Halvorson, H.R. & Brennan, M. *Biochemistry* **14**, 4953–4963 (1975).
26. Schmid, F.X. & Baldwin, R.L. *Proc. Nat. Acad. Sci. USA* **75**, 4764–4768 (1978).
27. Creighton, T.E. *J. Mol. Biol.* **87**, 579–602 (1974).
28. Cook, K.H., Schmid, F.X. & Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **76**, 6157–6161 (1979).
29. Schmid, F.X. & Blaschek, H. *Eur. J. Biochem.* **114**, 111–117 (1981).
30. Schmid, F.X. & Baldwin, R.L. *J. Mol. Biol.* **135**, 199–215 (1979).
31. Hvidt, A. C.R. Trav. Lab. Carlsberg Ser. Chim. **34**, 299–317 (1964).
32. Kim, P.S. & Baldwin, R.L. *Biochemistry* **19**, 6124–6129 (1980).
33. Blum, A.D., Smallcombe, H. & Baldwin, R.L. *J. Mol. Biol.* **118**, 305–316 (1978).
34. Perutz, M.F., Kendrew, J.C. & Watson, H.C. *J. Mol. Biol.* **13**, 669–678 (1965).
35. Schellman, C. In *Protein folding* (ed. Jaenicke, R.) 53–61 (Elsevier, North Holland; 1980).
36. Taniuchi, H. & Anfinsen, C.B. *J. Biol. Chem.* **244**, 3864–3875 (1969).
37. Epand, R.M. & Scheraga, H.A. *Biochemistry* **8**, 2864–2872 (1968).
38. Brown, J.E. & Klee, W.A. *Biochemistry* **10**, 470–476 (1971).
39. Bierzynski, A., Kim, P.S. & Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **79**, 2470–2474 (1982).
40. Chang, M.C., Fredrickson, R.A., Powers, S.P. & Scheraga, H.A. *Macromolecules* **14**, 633–634 (1981).
41. Sachs, D.H., Schechter, A.N., Eastlake, A. & Anfinsen, C.B. *Proc. Natl. Acad. Sci. USA* **69**, 3790–3794 (1972).
42. Jaenicke, R. & Rudolph, R. In *Protein folding* (ed. Jaenicke, R.) 525–548 (Elsevier, North Holland; 1980)
43. London, J., Skrzynia, C & Goldberg, M.E. *Eur. J. Biochem.* **47**, 409–415 (1974).
44. Dolgikh, D.A. *et al. Febs Lett.* **136**, 311–315 (1981).
45. Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S*J. Mol. Biol.* **106**, 359–373 (1976).
46. Ptitsyn, O.B. *Dokl. Akad. Nauk SSSR* **210**, 1213–1215 (1973).