Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 5082–5086, July 1988
Biophysics

Toward a better understanding of protein folding pathways

(bovine pancreatic trypsin inhibitor/RNase A)

THOMAS E. CREIGHTON

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Communicated by Frederic M. Richards, April 8, 1988 (received for review October 27, 1987)

ABSTRACT **Experimental observations of how unfolded proteins refold to their native three-dimensional structures contrast with many popular theories of protein folding mechanisms. The available experimental evidence (ignoring slow cis–trans peptide bond isomerization) is largely consistent with the following general scheme: under folding conditions, unfolded protein molecules rapidly equilibrate between different conformations prior to complete refolding. This rapid prefolding equilibrium favors certain compact conformations that have somewhat lower free energies than the other unfolded conformations. Some of the favored conformations are important for productive folding. The rate-limiting step occurs late in the pathway and involves a high-energy, distorted form of the native conformation; there appears to be a single transition state through which essentially all molecules refold. Consequently, proteins are not assembled via a large number of independent pathways, nor is folding initiated by a nucleation event in the unfolded protein followed by rapid growth of the folded structure. The known folding pathways involving disulfide bond formation follow the same general principles. An exceptional folding mechanism for reduced ribonuclease A proposed by Scheraga *et al.* (Scheraga, H.A., Konishi, Y., Rothwarf, D.M. & Mui, P.W. (1987) *Proc. Natl. Acad. Sci. USA* 84, 5740–5744) is shown to result from experimental shortcomings, an incorrect kinetic analysis, and a failure to consider the kinetics of unfolding.**

The mechanism by which proteins fold to their complex three-dimensional structures has been the subject of much debate and investigation. After Levinthal (1) pointed out that a protein could not sample all possible conformations in a reasonable period of time, kinetic pathways of folding were envisaged. Nucleation events in the unfolded protein were proposed to act as templates upon which folding would proceed rapidly to completion (2). Alternatively, the rate-limiting step was proposed to be the coalescence of metastable, "flickering" nuclei (3, 4). Virtually all mechanisms envisaged a relatively simple incorporation of elements of native-like conformation to build up the fully folded conformation. If this were to be the case, there could be very many ways of assembling a complex structure, just as there are very many ways of assembling a jigsaw puzzle (5).

Experimental elucidation of the folding process has been hampered by its cooperativity and complexity, but there is now a substantial amount of experimental data about the mechanism of protein folding. Moreover, virtually all these data, obtained with a variety of techniques and a number of different proteins, appear to indicate a consistent scheme of protein folding mechanisms that differs substantially from most theoretical models.

The purpose of this communication is to point out the relevant experimental observations concerning folding of small single-domain proteins and to address some recent controversies.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Experimental Observations of Protein Folding Transitions

Defining the Problem. Proteins are observed to refold on the second to minute time scale, depending on the protein and the conditions, but with half-times as short as slightly less than a second. Such rates are rightly considered rapid for this complex process (1), yet they are very slow for a unimolecular process with no intrinsically high energy barriers. An unfolded protein molecule probably assumes a recognizably different conformation every 10^{-11} sec, so it is not feasible either experimentally or theoretically to determine every conformational transition undergone while a single molecule is folding. This fundamental difficulty is compounded experimentally, where populations of molecules (e.g., nmol to μmol; 10^{14}–10^{18} molecules) need to be observed. It is necessary to think about both populations of molecules and their distribution between different conformational states. In the unfolded state, every molecule of a population probably has a different conformation at each instant of time, whereas all molecules have basically the same conformation in the folded state.

Besides such vast numbers of conformations, protein folding is also complicated by the general instability of partially folded conformations. In view of these complications, what aspects of protein folding could possibly be determined experimentally? One would be the distribution of conformations of the unfolded state under refolding conditions. Do the different unfolded molecules tend to sample a common subset of conformations? If so, what are these conformations and at what rate are they interconverted? A second would be the number of rate-limiting steps by which different populations of molecules fold. If different molecules fold by the same rate-limiting step, at which stage do they converge on the common pathway? Third, what is the nature of the rate-limiting step—i.e., the transition state? Is it close to the native or fully unfolded states? Finally, what conformational transitions in the folded and unfolded states precede the rate-limiting step in, respectively, unfolding and refolding?

The basic answers to most of these questions are known from experimental studies of folding of simple single-domain proteins, where folding is a unimolecular process controlled solely by the covalent structure of the protein and the environmental conditions. Whilst exceptions to the following general scheme are known, they are not necessarily incompatible with it, for they could arise by extraneous factors, such as protein heterogeneity.

Refolding Kinetics. If different molecules, or subsets of a population, refold by different pathways, a corresponding number of different rates of folding should be observed kinetically. If there are very many pathways, a continuum of rates should be apparent. Some pathways might happen to

Abbreviation: BPTI, bovine pancreatic trypsin inhibitor.

have the same rate constant under some conditions, but this is unlikely to be the case under a wide variety of conditions.

Unfortunately, the kinetics of refolding are usually complicated by heterogeneity of the unfolded state, caused by intrinsically slow cis–trans isomerization of peptide bonds adjacent to proline residues (6–9). The folded state usually has a unique set of isomers of the various peptide bonds that is present in all the folded molecules. Upon unfolding, however, the constraints are released and the other isomer of each prolyl peptide bond is formed with significant frequency in the various unfolded molecules. For those with one or more nonnative isomers, complete folding is delayed to various extents by the cis–trans isomerization kinetic barrier (10).

In spite of this complication, the kinetics of refolding are observed to be relatively simple; each population of molecules with one set of cis–trans isomers appears to refold with a single rate constant.

The complication of cis–trans isomerization can be obviated by studying the refolding of proteins with no proline residues (11, 12) or with only a small number; for example, hen egg white lysozyme has only two trans prolyl peptide bonds, and only 10% of the unfolded molecules have incorrect cis isomers (13–16). Another approach is to study the refolding of transiently unfolded protein before substantial isomerization occurs (7). In virtually all of these cases, the kinetics have been observed under a variety of folding conditions to be simple, with a single kinetic rate constant applicable to all the molecules.

Unfolded Proteins Rapidly Equilibrate Between Different Conformations. The simple kinetics of refolding are possible with a conformationally heterogeneous population of unfolded molecules only if all the molecules are in rapid conformational equilibrium prior to refolding or if they converge rapidly onto a common pathway. Such rapid equilibration of unfolded molecules is also indicated by the observation that the rate of refolding depends only on the final conditions. Different unfolding conditions (i.e., denaturants, extremes of pH or temperature, etc.) are well known to produce unfolded states that differ in their average physical properties (17). Nevertheless, when such different populations are refolded under the same conditions, the rates of refolding are indistinguishable (18–22). Therefore, whatever the initial unfolded state, it rapidly attains the properties appropriate to the folding conditions.

Equilibration amongst the unfolded molecules appears to occur by them tending to adopt a limited subset of conformations under refolding conditions (22–30). When transferred to conditions where they can refold completely, unfolded proteins tend to adopt compact, but nonnative, conformations that are in rapid equilibrium with each other and with more unfolded conformations (25, 26); complete refolding occurs more slowly. These transient compact conformations appear (25–31) to be similar to the recently recognized "molten globule" state that is stable with some proteins under some conditions (32, 33).

The rapid equilibration of unfolded molecules is incompatible with the multiple pathway model (5), which required that protein assembly be essentially irreversible and unidirectional. Assembling a jigsaw is not a good analogy for folding proteins, because it is not a cooperative process in which partially assembled states are unstable and rapidly disassemble. The inappropriateness of this model is also shown by the disulfide folding pathways described below and by the existence of mutants that can block protein folding (34).

Kinetics of Unfolding. The unfolding of proteins is almost invariably observed to be an all-or-none process, with a single rate constant that applies to all molecules of a particular protein. Partially folded conformations are not detected in a significant fraction of the molecules, so unfolding occurs completely or not at all.

These observations are pertinent to folding, since the pathway of refolding must be the reverse of the unfolding pathway under the same conditions. The absence of partially unfolded conformations during unfolding, and their presence during refolding, suggest that the rate-limiting step is much nearer the native than the unfolded state (25, 26, 35, 36).

Transition State for Folding. The relatively simple kinetics of unfolding and refolding, except for slow prolyl cis–trans isomerization, makes it possible to characterize the rate-limiting step—i.e., the transition state, for both processes. The most detailed studies have been made by Segawa and Sugihara (15, 16) on hen egg white lysozyme, in which 90% of the unfolded molecules refold without the complication of cis–trans isomerization (13).

The rate of unfolding has been shown to vary uniformly with changes in temperature and denaturant concentrations (15). This is the classical proof that the reaction mechanism and the transition state are not changing as the conditions are altered. Consequently, it is possible to consider a single transition state, as for any simple chemical reaction. In contrast, the rate of refolding varies nonuniformly—e.g., first increasing, then decreasing, with changing conditions (15).

If the nature of the transition state is not changing, the nature of the unfolded state must be. This is consistent with the above observation that unfolded proteins tend to adopt compact conformations to an extent that depends on the folding conditions. Among these compact conformations must be at least some that are involved directly in complete refolding. The overall rate of refolding depends on both the concentrations of these conformations and on the rate at which they complete refolding; both depend on the folding conditions. Because all the nonrandom conformations are sampled rapidly, they will have indistinguishable kinetic behavior (37, 38).

The transition state for folding of lysozyme has been shown (15, 16) to be similar to the native state in its solvent exposure of nonpolar groups, to be nearly as compact, with solvent molecules excluded from the interior, and with Glu-35 and Trp-108 in the same region. However, the transition state does not specifically bind substrate analogues.

These observations, taken with the effects of cross-links on the folding kinetics of some other proteins (39), are consistent with the proposal that the transition state for both unfolding and refolding is, at least to a first approximation, a high-energy distorted form of the native conformation (35, 36). Of course, the transition state will not be native-like in all respects (40, 41), but the number of ways in which it is similar to the native state indicates that it is at least very different from the unfolded state. The high energy of the native-like transition state probably arises from the need to disrupt the entropic cooperativity of the interactions that stabilize the native state (35, 36, 42).

Nucleation rapid growth types of folding mechanisms are incompatible with the rate-limiting step in refolding, being very close along the reaction coordinate to the native state (2). Nucleation events could be involved in the rapid equilibration of unfolded molecules, but they are not rate limiting for folding. Instead of searching for nucleation sites in unfolded proteins, it might be more relevant to search for unfolding nucleation events in the native conformation.

Disulfide Folding Pathways

Protein folding pathways may be elucidated experimentally if folding requires disulfide bond formation between cysteine residues (38, 43, 44). The unique redox properties of the disulfide interaction permit the partially folded kinetic intermediates that define the folding pathway to be trapped,

identified, and characterized. Unfolding and refolding are controlled by varying specifically only the intrinsic strength of the redox disulfide interaction, so no general denaturants of uncertain action are required, and both unfolding and refolding are carried out under identical conditions. Moreover, the kinetics of making, breaking, and rearranging disulfide bonds can reflect the conformational transitions undergone, but only if the appropriate experimental methods are used and the rates reflect the protein intramolecular steps, not the disulfide chemistry involved (43, 44). Consequently, this approach has yielded the most detailed pathways known, for bovine pancreatic trypsin inhibitor (BPTI) (45), two homologues from black mamba venom (46), RNase A (47, 48) and RNase T1 (49). These disulfide folding pathways are consistent with the general properties of folding transitions described above.

BPTI. The folding transition of BPTI shown in Fig. 1 is cooperative (37) and reflects the conformational properties of the protein (50). All the molecules of unfolded reduced BPTI equilibrate at the one-disulfide stage, because the single-disulfide intermediates are in rapid equilibrium with each other. Only one of these intermediates, that with the 30–51 disulfide bond, is important for further productive folding. Thereafter, all the molecules tend to follow a limited number of pathways.

The rate-limiting intramolecular step in both unfolding and refolding separates the native-like conformation from all the more unfolded conformations, consistent with it involving a distorted form of the native conformation. Only very minor perturbations of the folded conformation precede unfolding, upon breaking the 14–38 disulfide bond, whereas many conformational transitions precede the rate-limiting step in refolding. Other than the rate-limiting step, all the transitions are rapid and reversible under redox conditions where the native conformation is only marginally stable.

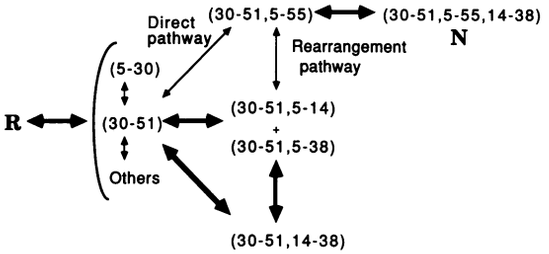

FIG. 1. Direct and rearrangement pathways for folding and disulfide formation in BPTI and homologous proteins. R, fully reduced protein; intermediates are designated by the residue numbers of the cysteine residues paired in disulfide bond. N, folded protein with the 30–51, 5–55, and 14–38 disulfides. The + between (30–51, 5–14) and (30–51, 5–38) indicates that they have comparable kinetic roles. R is unfolded and, consequently, makes initial disulfides essentially randomly; this is depicted with a single arrow and a bracket to encompass all the one-disulfide intermediates. The one-disulfide intermediates are in rapid equilibrium, and their various levels of accumulation depend on their relative free energies; only the two most predominant intermediates are depicted. Second disulfides are formed predominantly in (30–51), and three predominate: 14–38, 5–14, and 5–38. However, the resulting two-disulfide intermediates cannot complete refolding directly and rearrange intramolecularly to the native-like (30–51, 5–55). This intermediate can very rapidly form the 14–38 disulfide bond and complete refolding. Disulfide reduction and unfolding occur by the reverse of this pathway. A quasi-native dead-end intermediate (5–55, 14–38) that is formed readily from the minor intermediate (5–55) is omitted from the pathway for clarity. This disulfide rearrangement pathway predominates with BPTI because the 14–38, 5–14, and 5–38 disulfides are formed in (30–51) at a rate 10^5 times greater than the 5–55 disulfide of the direct pathway. Forming 5–55 directly in (30–51) has the same intramolecular rate as in forming it in the rearrangement step, as would be expected if the same conformational transitions are involved, irrespective of whether the disulfide being rearranged is 5–14 or 5–38 or the mixed-disulfide with the reagent (38). The direct pathway is energetically favorable with the BPTI homologues from black mamba venom (46) because the 5–55 disulfide is formed directly 10^3-fold more rapidly than in BPTI, while the rearrangement step has about the same rate. Consequently, the intramolecular steps involved in the direct pathway have lower free-energy barriers than those via the rearrangement pathway in these less stable BPTI homologues.

That the rate-limiting step involved rearrangement of two-disulfide intermediates with nonnative disulfides was unexpected and initially surprising (51). Nevertheless, the importance of the rearrangements was indicated by several independent experimental observations. One of the most dramatic was the kinetic consequences of blocking irreversibly the Cys-14 and Cys-38 thiols; the rates of both unfolding and refolding were dramatically decreased. Concern that this might be a consequence of the blocking groups on the thiols led Marks *et al.* (52) to compare the kinetic behavior of engineered variants of BPTI with alanine or serine residues at positions 14 and 38. They observed that "there are no discernible differences in the refolding behavior of BPTI alkylated at cysteines 14 and 38 and BPTI in which alanine or threonine have been substituted for these residues." Goldenberg (53) studied similar mutants in which serine residues were introduced and demonstrated that their refolding kinetics were consistent with the pathway of Fig. 1. Allegations by Wetlaufer *et al.* (54) that changes in the thiol and disulfide reagent concentrations during folding were not taken into account in the BPTI kinetic analysis are incorrect.

The disulfide rearrangement pathway of BPTI is now believed to reflect the extreme stability of its folded conformation (36, 46). The high energy required to distort it sufficiently for unfolding results, for unknown reasons, in the rearrangement pathway having the lowest free-energy transition state. That is not the case with less stable BPTI homologues (46) and possibly also with BPTI under conditions in which it is less stable (52); in these cases, the direct pathway of sequentially forming the three disulfide bonds (Fig. 1) has the lowest free-energy barrier. The transition state in the direct folding pathway is still a distorted form of the native-like conformation, and this folding pathway is still compatible with the general observations of protein folding transitions described above.

RNase A. Similar experimental studies of the disulfide folding pathway of RNase A have not elucidated a detailed pathway, for there are far too many intermediates (47, 48), plus complications of covalent modification (48, 55). Nevertheless, the overall scheme (Fig. 2*a*) and the kinetics and energetics of both unfolding and refolding were found to be similar to those observed with BPTI and to be consistent with the general properties of protein folding transitions. The rate-limiting intramolecular step in both refolding and unfolding involves, respectively, making and breaking a fourth correct disulfide bond of the native conformation.

A very different RNase A disulfide folding pathway (Fig. 2*b*), however, has been proposed by Scheraga *et al.* (56–63). They have concluded that there are very many rate-limiting steps. None of the putative rate-determining steps was identified directly, however, and several deficiencies in their experimental approach have been suggested (42, 43, 64). Scheraga *et al.* (56) have rejected these suggestions, however, so it is necessary to spell them out in greater detail.

Multiple pathways of refolding with different rate-limiting steps were claimed by Scheraga *et al.* (56–59) to be necessary because they were unable to fit all their kinetic data on the appearance of RNase enzyme activity under a wide variety of redox conditions with a single rate-limiting step. There are several reasons why they were unable to do so.

(*i*) They used an inappropriate type of disulfide reagent—namely, oxidized glutathione, with an intermolecular disul-

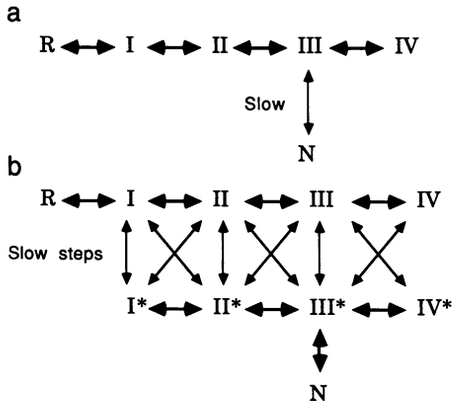

FIG. 2. Comparison of the different rate-limiting steps in refolding of reduced RNase A proposed by Creighton (48) (*a*) and Scheraga *et al.* (56–59) (*b*). In each case, R is the protein with no disulfides, and N is fully folded with the four correct disulfides. The intermediate species are indicated by the number of intramolecular protein disulfide bonds they contain (Roman numerals). Each is a mixture of very many different species, and each species is likely to have unique kinetic properties. Mixed disulfides with the reagent are ignored. The rate-limiting steps are those indicated by thin vertical arrows; not all members of each group of intermediates are likely to undergo the rate-limiting steps. All other steps indicated are more rapid in at least the productive intermediates. In *b*, the intermediates marked with asterisks are postulated to occur after the rate-limiting step in folding and to be converted rapidly to N.

fide bond. The rates of productive protein disulfide formation with this type of reagent usually do not reflect the conformational tendencies of the protein but simply the rate of the chemical exchange reaction between protein thiols and the disulfide reagent (38, 43–51). The intramolecular rate constants relevant to protein folding can be obtained from the rate of disulfide reduction by glutathione (45), but only if homogeneous species are being identified (see below).

(*ii*) Intermediates trapped with acid were fractionated only on the basis of the number of mixed disulfides with glutathione that they contained. Cysteine residues accumulate as mixed disulfides only if they cannot readily form a protein disulfide with any other free cysteine residues of the protein. Therefore, intermediates with mixed disulfides are those least likely to be on the productive pathway and most likely to be abortive dead-end intermediates. The more relevant concentrations of species with different numbers of protein disulfides were only estimated indirectly.

(*iii*) A single rate constant for each step was assigned to each class of species, defined simply by the number of intra- and intermolecular disulfides, even though each class contains many covalently different species. Scheraga *et al.* (56) maintain that their kinetic model is valid because each single rate constant is the weighted sum of the individual microscopic rate constants for each species. This would be true, however, only if all the species within each class are in rapid equilibrium. If their interconversion is slow, or nonexistent, the individual microscopic steps will be observed, not a single composite step with a single rate constant. Therefore, the kinetic analysis of Scheraga *et al.* (56–59) is valid only if all the various intermediates of RNase A that occur before the rate-limiting step in folding are interconverted rapidly on the time scale of minutes. Three types of kinetically significant heterogeneity are almost certain to occur. (*a*) Intermediates with the same number of inter- and intramolecular disulfides will differ in the cysteine residues involved. Altogether, a total of 7193 chemically distinct species are possible, but they were separated into only 25 classes (56). Intermediates that cannot possibly be in rapid equilibrium are those with all eight cysteine residues involved in disulfide bonds (i.e., those with four protein disulfides, with three protein disulfides and two mixed disulfides, etc.); interconverting them requires breaking at least one of the disulfide bonds. It is also unlikely that there is a rapid equilibrium between all the molecules with three protein disulfides, the intermediates that are believed to be involved in the true rate-limiting step (Fig. 2*a*). (*b*) Molecules with the same disulfides can differ in their conformations. Most conformations will be in rapid equilibrium, but possibly not if they differ in certain respects, such as different disulfide bond geometries (64), disulfide loop topologies, or cis–trans isomers of peptide bonds preceding proline residues (62). (*c*) Molecules indistinguishable in the preceding aspects can be heterogeneous covalently in ways that could affect their rates of folding. Sufficient evidence for covalent heterogeneity of RNase has been reported (47, 48, 55), at least some of which is produced during unfolding, for this to be a very plausible source of kinetic complexity. Therefore, the kinetic analysis of Scheraga *et al.* (56–59) is invalid in assuming that a single rate constant will apply collectively to all the members of each class.

(*iv*) Scheraga *et al.* trapped their species only by acidification, which is not sufficient to stop intramolecular disulfide rearrangements (43, 44). Consequently, a further explanation for the diversity of the rate of appearance of RNase activity is that the acid-trapped intermediates were enzymatically active or became active during the analysis. The activities of the intermediates were claimed by Scheraga *et al.* (59, 62) to be negligible, even though one fully active three-disulfide intermediate has been well-characterized (65), and others are likely to occur (48, 66). The allegation by Konishi and Scheraga (62) that traces of activity present in the iodoacetate-trapped intermediates (48) resulted from contamination with native protein are not correct. Even the fully reduced protein exhibits significant activity (67).

In summary, there are numerous reasons why RNase A activity would not appear in a single rate-limiting step under the inappropriate conditions used by Scheraga *et al.* (56–63). By incorrectly assuming that there should be a single rate constant for each defined step, they were forced to make other steps rate limiting to fit the kinetic data.

Finally, the incorrectness of the RNase A folding pathway proposed by Scheraga *et al.* (56–59) is demonstrated by the kinetics of unfolding. Their pathway would require the accumulation of one-, two-, and three-disulfide intermediates (I*, II*, III*) that precede the rate-limiting step in unfolding (Fig. 2*b*). Scheraga *et al.* (56–63) did not determine the kinetics of unfolding, but the absence of such intermediates in unfolding has been demonstrated directly (48). Unfolding native RNase A by reducing its four disulfides in an all-or-none process in which the rate-limiting step is breaking the first disulfide, and no one-, two-, or three-disulfide intermediates accumulate.

The experimental results obtained in the two independent studies of RNase A disulfide folding have recently been shown (S.J. Wearne and T.E.C., unpublished observations) not to differ as a result of using different conditions or disulfide reagents, as suggested by Scheraga *et al.* (56). The two sets of experimental results are not inconsistent, but the very different conclusions of Scheraga *et al.* (56–63) result from their use of inappropriate experimental methods, an incorrect kinetic analysis, and a failure to take into account the kinetics of unfolding.

1. Levinthal, C. (1968) *J. Chim. Phys.* **65**, 44–45.
2. Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 697–701.
3. Anfinsen, C. B. (1973) *Science* **181**, 223–230.
4. Karplus, M. & Weaver, D. C. (1976) *Nature (London)* **260**, 404–406.
5. Harrison, S. C. & Durbin, R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4028–4030.

6. Garel, J.-R. & Baldwin, R. L. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 3347–3351.
7. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975) *Biochemistry* **14,** 4953–4963.
8. Lin, L.-N. & Brandts, J. F. (1984) *Biochemistry* **23,** 5713–5723.
9. Kelley, R. F. & Richards, F. M. (1987) *Biochemistry* **26,** 6765–6774.
10. Kim, P. S. & Baldwin, R. L. (1982) *Annu. Rev. Biochem.* **51,** 459–489.
11. Brandts, J. F., Brennan, M. & Lin, L.-N. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4178–4181.
12. Lin, L.-N. & Brandts, J. F. (1978) *Biochemistry* **17,** 4102–4110.
13. Kato, S., Shimamoto, N. & Utiyama, H. (1982) *Biochemistry* **21,** 38–43.
14. Segawa, S.-I., Husimi, Y. & Wada, A. (1973) *Biopolymers* **12,** 2521–2537.
15. Segawa, S.-I. & Sugihara, M. (1984) *Biopolymers* **23,** 2473–2488.
16. Segawa, S.-I. & Sugihara, M. (1984) *Biopolymers* **23,** 2489–2498.
17. Tanford, C. (1968) *Adv. Protein Chem.* **23,** 121–282.
18. Creighton, T. E. (1974) *J. Mol. Biol.* **87,** 563–577.
19. Garel, J.-R., Nall, B. T. & Baldwin, R. L. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 1853–1857.
20. Denton, J. B., Konishi, Y. & Scheraga, H. A. (1982) *Biochemistry* **21,** 5155–5163.
21. Lynn, R. M., Konishi, Y. & Scheraga, H. A. (1984) *Biochemistry* **23,** 2470–2477.
22. Kato, S., Okamura, M., Shimamoto, N. & Utiyama, H. (1981) *Biochemistry* **20,** 1080–1085.
23. Schmid, F. X. & Baldwin, R. L. (1980) *J. Mol. Biol.* **135,** 199–215.
24. Schmid, F. & Blaschek, H. (1984) *Biochemistry* **23,** 2128–2133.
25. Creighton, T. E. (1980) *J. Mol. Biol.* **137,** 61–80.
26. Creighton, T. E. & Pain, R. H. (1980) *J. Mol. Biol.* **137,** 431–436.
27. Semsotnov, G. V., Rodionova, N. A., Kutyshenko, N. A., Ebert, B., Blanck, J. & Ptitsyn, O. B. (1987) *FEBS Lett.* **224,** 9–13.
28. Gilmanshin, R. I. & Ptitsyn, O. B. (1987) *FEBS Lett.* **223,** 327–329.
29. Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. & Nagamura, T. (1987) *FEBS Lett.* **221,** 115–118.
30. Touchette, N. A., Perry, K. M. & Matthews, C. R. (1986) *Biochemistry* **25,** 5445–5452.
31. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986) *Biochemistry* **25,** 6965–6972.
32. Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S. (1976) *J. Mol. Biol.* **106,** 359–373.
33. Damaschun, G., Gernat, Ch., Damaschun, H., Bychkova, V. E. & Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol.* **8,** 226–230.
34. King, J. (1986) *Bio/Technology* **4,** 297–303.
35. Creighton, T. E. (1981) in *Structural Aspects of Recognition and Assembly in Biological Macromolecules*, eds. Balaban, M., Sussman, J. L., Traub, W. & Yonath, A. (Balaban ISS, Rehovot), pp. 57–73.
36. Goldenberg, D. P. & Creighton, T. E. (1985) *Biopolymers* **24,** 167–182.
37. Creighton, T. E. (1977) *J. Mol. Biol.* **113,** 295–312.
38. Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol.* **33,** 231–297.
39. Goto, Y. & Hamaguchi, K. (1982) *J. Mol. Biol.* **156,** 911–926.
40. Lin, S. H., Konishi, Y., Nall, B. T. & Scheraga, H. A. (1985) *Biochemistry* **24,** 2680–2686.
41. Goto, Y. & Hamaguchi, K. (1987) *Biochemistry* **26,** 1879–1884.
42. Creighton, T. E. (1983) *Biopolymers* **22,** 49–58.
43. Creighton, T. E. (1984) *Methods Enzymol.* **107,** 305–329.
44. Creighton, T. E. (1986) *Methods Enzymol.* **131,** 83–106.
45. Creighton, T. E. & Goldenberg, D. P. (1984) *J. Mol. Biol.* **179,** 497–526.
46. Hollecker, M. & Creighton, T. E. (1983) *J. Mol. Biol.* **168,** 409–437.
47. Creighton, T. E. (1977) *J. Mol. Biol.* **113,** 329–341.
48. Creighton, T. E. (1979) *J. Mol. Biol.* **129,** 411–431.
49. Pace, C. N. & Creighton, T. E. (1986) *J. Mol. Biol.* **188,** 477–486.
50. Creighton, T. E. (1980) *J. Mol. Biol.* **144,** 521–550.
51. Creighton, T. E. (1977) *J. Mol. Biol.* **113,** 275–293.
52. Marks, C. B., Naderi, H., Kosen, P. A., Kuntz, I. D. & Anderson, S. (1987) *Science* **235,** 1370–1373.
53. Goldenberg, D. P. (1988) *Biochemistry* **27,** 2481–2489.
54. Wetlaufer, D. B., Branka, P. A. & Chen, G.-X. (1987) *Protein Eng.* **1,** 141–146.
55. Thannhauser, T. W. & Scheraga, H. A. (1985) *Biochemistry* **24,** 7681–7688.
56. Scheraga, H. A., Konishi, Y., Rothwarf, D. M. & Mui, P. W. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5740–5744.
57. Scheraga, H. A., Konishi, Y. & Ooi, T. (1984) *Adv. Biophys.* **18,** 21–41.
58. Konishi, Y., Ooi, T. & Scheraga, H. A. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5734–5738.
59. Konishi, Y., Ooi, T. & Scheraga, H. A. (1982) *Biochemistry* **21,** 4734–4740.
60. Konishi, Y., Ooi, T. & Scheraga, H. A. (1982) *Biochemistry* **21,** 4741–4748.
61. Konishi, Y., Ooi, T. & Scheraga, H. A. (1981) *Biochemistry* **20,** 3945–3955.
62. Konishi, Y. & Scheraga, H. A. (1980) *Biochemistry* **19,** 1308–1316.
63. Konishi, Y. & Scheraga, H. A. (1980) *Biochemistry* **19,** 1316–1322.
64. Creighton, T. E. (1985) *J. Phys. Chem.* **89,** 2452–2459.
65. Creighton, T. E. (1980) *FEBS Lett.* **118,** 283–288.
66. Garel, A. T. & Garel, J. R. (1982) *J. Biol. Chem.* **257,** 4031–4033.
67. Garel, J. R. (1978) *J. Mol. Biol.* **118,** 331–345.