Protein Science (1994), *3*:1945–1952. Cambridge University Press. Printed in the USA.
Copyright © 1994 The Protein Society

Fast folding of a prototypic polypeptide: The immunoglobulin binding domain of streptococcal protein G

JOHN KUSZEWSKI, G. MARIUS CLORE, AND ANGELA M. GRONENBORN

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

(RECEIVED August 15, 1994; ACCEPTED August 29, 1994)

Abstract

The folding of the small (56 residues) highly stable B1 immunoglobulin binding domain (GB1) of streptococcal protein G has been investigated by quenched-flow deuterium–hydrogen exchange. This system represents a paradigm for the study of protein folding because it exhibits no complicating features superimposed upon the intrinsic properties of the polypeptide chain. Collapse to a semicompact state exhibiting partial order, reflected in protection factors for ND–NH exchange up to 10-fold higher than that expected for a random coil, occurs within the dead time (≤1 ms) of the quenched flow apparatus. This is followed by the formation of the fully native state, as monitored by the fractional proton occupancy of 26 backbone amide groups spread throughout the protein, in a single rapid concerted step with a half-life of 5.2 ms at 5 °C.

Keywords: folding intermediates; protein folding; quenched-flow deuterium–hydrogen exchange

Although folding intermediates have attracted considerable attention over the last few years (Kuwajima, 1989; Kim & Baldwin, 1990; Matthews, 1993; Fersht & Dill, 1994; Ptitsyn, 1994), new experimental and theoretical studies have cast doubt on their importance as obligate states on the folding pathway of proteins. In particular, theoretical studies and computer simulations of the folding of simplified models for proteins has led to the proposal of a 3-stage random search mechanism of folding, involving the initial rapid collapse from a random coil state to a random semicompact globule, followed by a rate-determining search for one of the many possible transition states from which the chain rapidly folds to the native state (Camacho & Thirumalai, 1993; Chan & Dill, 1994; Sali et al., 1994). The results of this theoretical study further stimulate the debate (Baldwin, 1994; Creighton, 1994), initiated by a provocative paper by Sosnick et al. (1994), challenging the widely held assumption that transient, partly folded intermediates are responsible for the rapid folding of proteins. In 2 related studies (Elöve et al., 1994; Sosnick et al., 1994), it was shown that for cytochrome *c* at least, folding can occur extremely rapidly if trapped, incorrectly ligated heme intermediates are avoided. Indeed, the authors demonstrated that the experimental conditions could be adjusted such that up to 50–70% of the protein can attain the correct heme ligation directly and thus fold rapidly to the native state. These results reopen the debate of whether slow processes observed in folding reactions are of any significance kinetically for native protein folding or are simply the result of species trapped by optional barriers created by conditions specific to the folding solution. Such misorganization of the polypeptide chain may be connected to particular chemical requirements for the folding of the protein under investigation. These include side-chain protonation states, sulfhydryl chemistry, *cis–trans* proline isomerizations, and the presence of prosthetic groups or metal binding sites. To date, all experimental studies on folding have involved proteins that exhibit these complicating features superimposed on the intrinsic properties of the polypeptide chain itself (Kuwajima, 1989; Kim & Baldwin, 1990; Matthews, 1993; Fersht & Dill, 1994; Ptitsyn, 1994). Thus, the observation of slow steps in the folding of these proteins may, to a large degree, be caused by structural misorganization involving nonoptimal disulfide bond formation, *cis–trans* proline isomerization, or the incorrect ligation of a prosthetic group or metal ion. Consequently, it is of utmost importance to study the folding of a protein that is devoid of all these complications and that can therefore be regarded as a "generic polypeptide."

The B1 immunoglobulin domain (GB1) of streptococcal protein G fulfills this criterion. In addition, it possesses several other qualities that make it an ideal system to study the folding process: it is small (56 residues) and its structure has been deter-

Reprint requests to: G. Marius Clore or Angela M. Gronenborn, Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520.

mined at high resolution by both NMR (Gronenborn et al., 1991) and X-ray crystallography (Achari et al., 1992; Gallagher et al., 1994); it contains both helix and β-sheet structure; it is extremely stable to both high temperature and chemical denaturants (Gronenborn et al., 1991; Alexander et al., 1992a) and contains only a single tryptophan residue; 95% of the residues are involved in elements of secondary structure, and 26 hydrogen bonded backbone amides, spread throughout the protein (Fig. 1), have very slow hydrogen-deuterium (H-D) exchange rates ($t_{1/2} > 24$ h at pH 4.3; Gronenborn et al., 1991); and, finally, complete ^1H and ^{15}N resonance assignments are available (Gronenborn et al., 1991; Gronenborn & Clore, 1993). The experimental method that we employed to study the folding of GB1 involved the application of quenched flow D-H exchange (Roder et al., 1988; Udgaonkar & Baldwin, 1988; Roder, 1989; Englander & Mayne, 1992; Baldwin, 1993; Woodward, 1994). This technique is particularly powerful because it permits many specific sites within a protein, in this case the aforementioned 26 backbone amide groups, to be probed on the millisecond time scale.

Results and discussion

Equilibrium folding and unfolding of GB1

In order to establish conditions for the quenched-flow D-H experiments, we proceeded to establish the equilibrium unfolding curve of GB1 as a function of guanidine hydrochloride (GuHCl) at pH 4 and 5 °C. The results are shown in Figure 2. The unfolding transition is clearly 2-state, with the transition point occurring at about 3 M GuHCl. Analysis of the data, as described by Pace et al. (1989) yields a free energy of unfolding in water, $\Delta G(H_2O)$, of 4.8 kcal·mol^{-1}. With these data in hand, the kinetic folding experiments were carried out by diluting fully unfolded GB1 in 4.5 M GuHCl into a folding mixture containing 0.25 M GuHCl, at which concentration GB1 is fully folded.

In the quenched-flow D-H experiments, the protein is exposed to both low and high pH. We therefore examined both the 1-dimensional (1D) ^1H-NMR spectrum and the 2D ^1H-^{15}N correlation spectrum of GB1 as a function of pH. The protein remains fully folded over a range extending from pH 2 to pH 11.3 at 25 °C. Unfolding begins to occur above pH 11.3 and, at pH 11.6, GB1 is fully unfolded. Conditions for the quenched-flow D-H experiments were therefore chosen so that no significant contribution from the reverse reaction (i.e., unfolding) could occur. Thus, folding was allowed to occur at pH 4 and pulse labeling was carried out at pH 9.

Quenched-flow D-H exchange experiments

Previous stopped-flow optical measurements (Alexander et al., 1992b), as well as preliminary quenched-flow D-H exchange experiments (data not shown), indicated that GB1 folds within 5 ms at room temperature. To slow the folding reaction down, we therefore carried out all D-H exchange experiments at 5 °C. Even at this temperature, complete refolding to the native state occurs within 20 ms. For the simplest case of a 2-state folding transition, the processes that occur during the quenched-flow D-H exchange experiment can best be described as follows. Fully unfolded protein (U) in D_2O at pH 4, with all backbone amide groups completely deuterated (that is ND), is allowed to

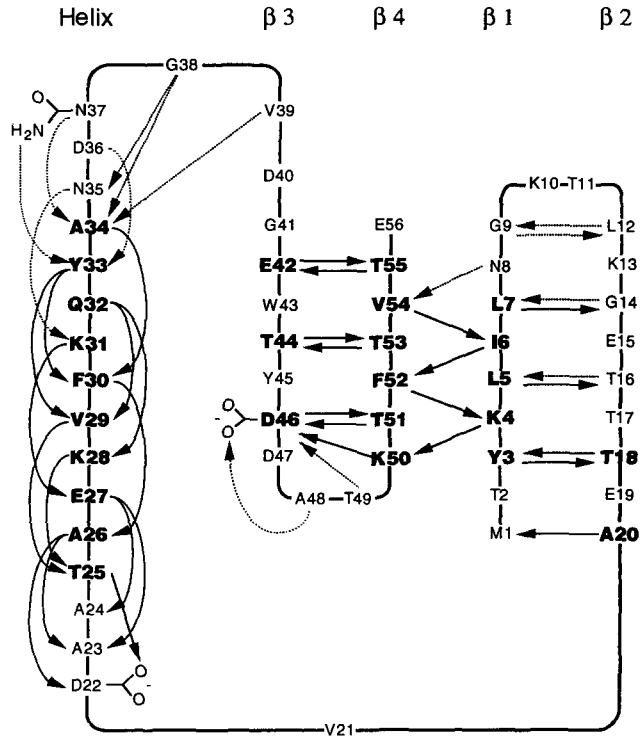

Fig. 1. Sequence, secondary structure, and hydrogen bonding of GB1 (Gronenborn et al., 1991). Hydrogen bonds and residues associated with the 26 backbone amide groups that exchange slowly in the folded state ($t_{1/2} > 24$ h at pH 4) are depicted as solid arrows and bold letters, respectively. Hydrogen bonds involving the remaining backbone amide groups exchange too rapidly in the folded state to be monitored in the quenched-flow D-H exchange experiment and are shown as dotted arrows.

fold for a length of time t_{fold} yielding native protein (N). During the time t_{fold}, the only reaction that occurs is that associated with the folding process itself:

$$\text{U (ND)} \xrightarrow{k_f(\text{low})} \text{N (ND)} \qquad (1)$$

where k_f(low) is the rate constant for folding at pH 4. Subsequently, the reaction mixture is subjected to a high pH exchange pulse in H_2O of duration t_{ex} (pH 9.1 and 25 ms), followed by a quench pulse at pH 3.5. During the high pH exchange period t_{ex}, there is direct competition between ND-NH exchange in the unfolded state and the folding process itself, such that the kinetics may be described by the scheme:

$$\begin{array}{ccc} \text{U (ND)} & \xrightarrow{k_f(\text{high})} & \text{N (ND)} \\ \Big\downarrow k_{ex} & & \\ \text{U (NH)} & \xrightarrow{k_f(\text{high})} & \text{N (NH)} \end{array} \qquad (2)$$

where k_f(high) is the rate constant for folding at pH 9. No significant ND-NH exchange occurs in the folded state during the short duration, t_{ex} (25 ms), of the high pH pulse (pH 9) so that

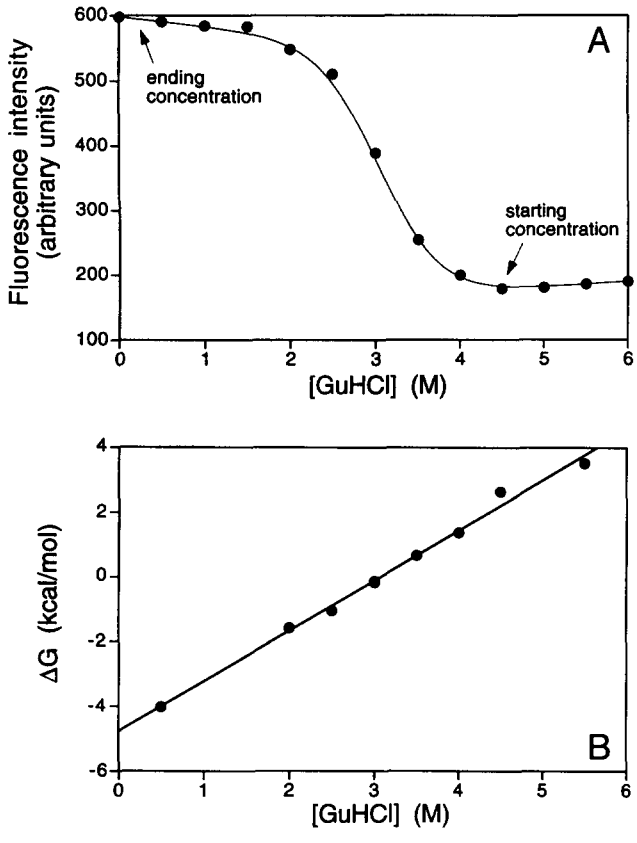

Fig. 2. Equilibrium unfolding of the GB1 domain at pH 4.0 in guanidine hydrochloride. **A:** Tryptophan fluorescence as a function of guanidine hydrochloride (fluorescence excitation was carried out at 295 nm and the emission was integrated from 305 to 450 nm). The arrows indicate the concentrations of guanidine hydrochloride used to unfold (starting concentration) and refold (ending concentration) the GB1 domain in the quenched-flow D-H exchange experiments. **B:** ΔG as a function of guanidine hydrochloride. Linear extrapolation yields a value of 4.8 kcal·mol^{-1} for $\Delta G(H_2O)$.

the N(ND) to N(NH) exchange reaction can be neglected. (Note that there is no change in intensity of the amide resonances in a sample of native GB1 at pH 9 following 1.5 s presaturation of the water resonance, indicating that the lifetime for amide exchange at this pH must be greater than 5 s.) Because folding of GB1 is so rapid, the exchange and folding processes occur on the same time scale. Once the protein is fully folded, the degree of protection of the backbone amide deuterons against solvent protons achieved at the end of the exchange period is monitored by recording a ^1H-^{15}N correlation spectrum on the now completely refolded protein. The fractional proton occupancy, F_i, of a particular backbone amide group i, as a function of t_{fold} and t_{ex} can then be described by the equation:

$$F_i(t_{fold}, t_{ex}) = \{k_{ex}(i)/[k_f(\text{high}) + k_{ex}(i)]\} \times \{1 - e^{-[k_f(\text{high})+k_{ex}(i)]t_{ex}}\} \cdot e^{-k_f(\text{low})t_{fold}}. \quad (3)$$

Because t_{ex} is long compared to $1/[k_f(\text{high}) + k_{ex}]$, Equation 3 can be simplified to

$$F_i(t_{fold}) = \{k_{ex}(i)/[k_f(\text{high}) + k_{ex}(i)]\} \cdot e^{-k_f(\text{low})t_{fold}}. \quad (4)$$

The experimental data illustrating the fractional proton occupancy of the 26 monitored amide groups as a function of the folding time t_{fold} is depicted in Figure 3.

We initially fitted all the experimental data in Figure 3 simultaneously to Equation 3, optimizing the values of k_f(high) and k_f(low) with the assumption that the exchange rate constants $k_{ex}(i)$ were the same as those expected for a random coil, $k_{xu}(i)$, calculated on the basis of model peptide data, if the unfolded state were a completely random coil (Bai et al., 1993; Connelly et al., 1993). That is to say, the protection factors P_i given by $k_{xu}(i)/k_{ex}(i)$, were assumed to have a value of 1.0. The resulting best-fit curves failed to fit the experimental data adequately and displayed systematic deviations in the distribution of residuals. Moreover, the value obtained in this manner for the high pH folding rate constant k_f(high) was a factor of 4.3 larger than that of the low pH one k_f(low), which appears unlikely because refolding of GB1 is slower at pH 11 than at neutral pH (Alexander et al., 1992b).

We then proceeded to fit all the experimental data in Figure 3 simultaneously to Equation 3, optimizing the value of the folding rate constant k_f(low) with the assumption that k_f(low) = k_f(high), and the values of the individual exchange rate constants $k_{ex}(i)$ for the 26 monitored amide groups. As is apparent from Figure 3, all the data can be fitted simultaneously with a single value of 133 ± 2 s^{-1} for k_f(low) with an overall standard deviation of the fit of 1.5%, random distribution of residuals, and good determination of all the optimized parameters. The values of $k_{ex}(i)$ expressed in terms of backbone amide protection factors, P_i, are plotted as a function of residue in Figure 4, and displayed spectrally on the structure of GB1 in Figure 5A.

The assumption that the rate constant for folding at pH 4, k_f(low), and pH 9, k_f(high), are the same does not affect the value obtained for k_f(low) because this is purely determined by the decay of the proton occupancy as a function of the folding time t_{fold} at pH 4. Moreover, whereas this assumption clearly affects the absolute values of the protection factors P_i, it does not affect the relative values of P_i. Thus, the proton occupancy at zero folding time t_{fold} is given by

$$F_i(0) = k_{ex}(i)/[k_f(\text{high}) + k_{ex}(i)] = k_{xu}(i)/[P_i k_f(\text{high}) + k_{xu}(i)], \quad (5)$$

so that the ratio of P_i and P_j for 2 amide protons i and j is independent of the folding rate constant k_f(high)

$$P_i/P_j = \{k_{xu}(i)[F_i(0)^{-1} - 1]\}/\{k_{xu}(j)[F_j(0)^{-1} - 1]\}. \quad (6)$$

In addition, the true absolute value of P_i is simply given by

$$P_i = k_f(\text{low})P_i(\text{app})/k_f(\text{high}), \quad (7)$$

where P_i(app) are the values of the apparent protection factors obtained on the assumption that k_f(low) = k_f(high). It should also be noted that P_i cannot be less than 1 because the observed exchange rate constants $k_{ex}(i)$ will not exceed the expected values $k_{xu}(i)$ for a random coil. Examination of Figure 4 indicates that Leu 7 has an apparent protection factor of 1. This implies that the rate constant for folding at pH 9 cannot be larger than

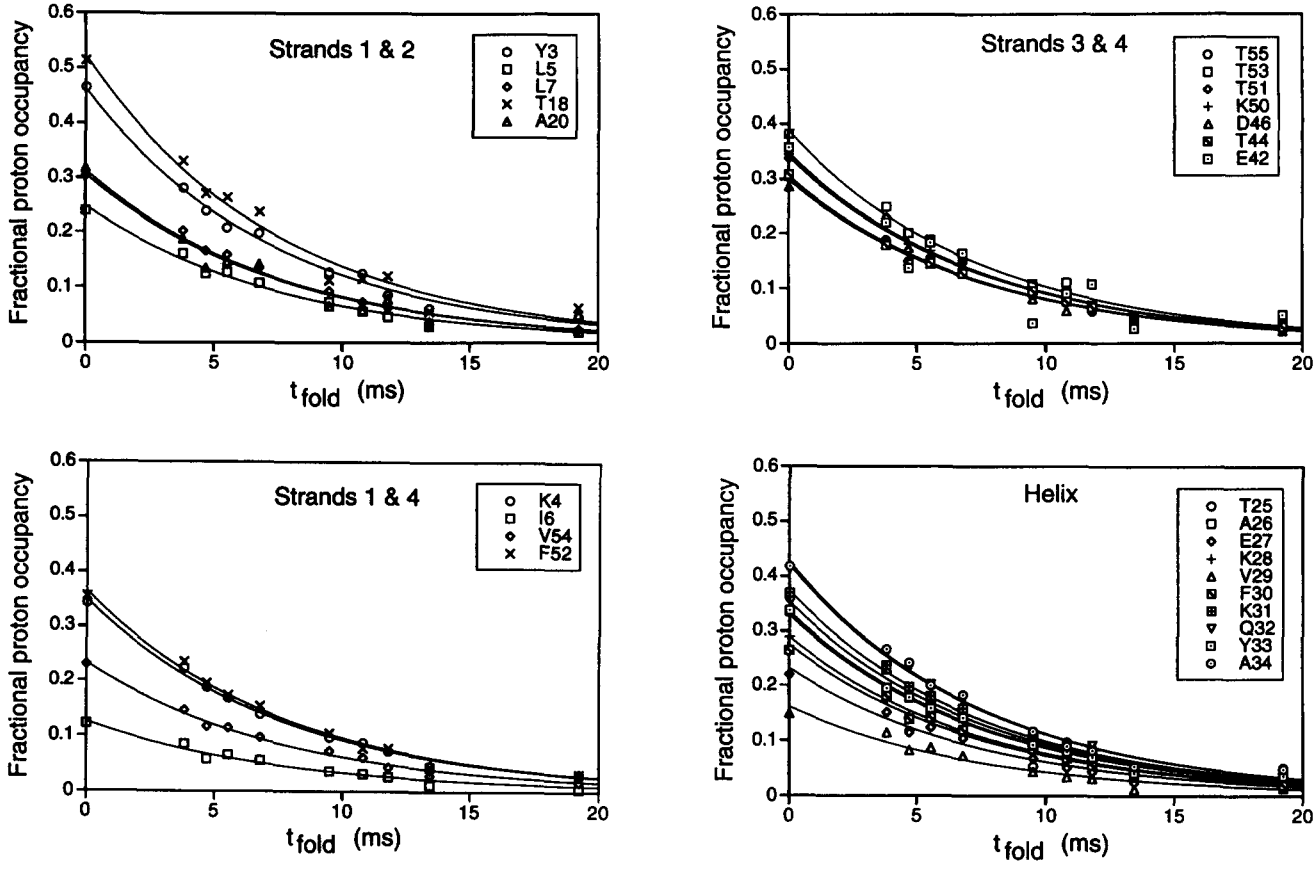

Fig. 3. Comparison of the observed time courses of NH proton occupancy as a function of the length of the folding time t_{fold} (depicted by the symbols) with the best-fit computed time courses (represented by the solid lines) calculated using Equation 3. All the experimental data that comprise the kinetic curves for the 26 backbone amide groups that are slowly exchanging in the folded state, were fitted simultaneously with a single value for the folding rate constant k_f (133 s^{-1}). The NH proton occupancies were measured from 2D ^1H-^{15}N heteronuclear single quantum coherence spectra recorded at 600 MHz.

that at pH 4, so that the values of P_i(app) plotted in Figure 4 represent lower limits on the true absolute values of P_i.

In principle one could measure the value of k_f(high) by simply refolding the protein directly into folding buffer at pH 9 for varying lengths of time t_{ex} prior to quenching. In this case, the fractional proton occupancy would be given by

$$F_i(t_{ex}) = \{k_{ex}(i)/[k_f(\text{high}) + k_{ex}(i)]\} \times \{1 - e^{-[k_f(\text{high})+k_{ex}(i)]t_{ex}}\}. \quad (8)$$

Unfortunately, however, the observed rate constant given by the sum of k_f(high) and $k_{ex}(i)$ is too large to permit us to measure the kinetics of this competition experiment between refolding and amide exchange.

Marginal protection factors, such as those observed here for the apparent initial state, have often been attributed to the existence of parallel folding pathways associated with multiphasic refolding kinetics and characterized by the presence of distinct plateaus (Radford et al., 1992; Miranker et al., 1993; Mullins et al., 1993; Dobson et al., 1994). In the present case, only a single exponential phase is observed for the refolding kinetics observed in the quenched-flow D–H exchange experiments. If the apparent protection factors obtained from the present experiments were due to a parallel folding pathway in which a fraction (1 – X) of the unfolded GB1 molecules refolded within the dead-time of the quenched flow apparatus (that is, ≤1 ms), while the remainder refolded with a half-life of 5.2 ms, the observed protection factors would be uniform for all residues and given by 1/X. The fact that the apparent protection factors vary over an order of magnitude therefore excludes this possibility.

The nature of the apparent initial state

The results of the quenched-flow D–H exchange experiments indicate that the formation of the folded state from either the fully unfolded state or a state formed within the dead-time of the quenched flow apparatus (that is, ≤1 ms) occurs in a single rapid concerted step with a half-life of 5.2 ms. The observed protection factors provide insight into the nature of the state from which the 5.2-ms folding process occurs. We will refer to this state as the apparent initial state and argue that it is a state distinct from the completely unfolded random coil state.

The overall apparent average protection factor for the 26 monitored amide groups is 4.3 ± 1.6, with values ranging from 0.98 to 8.4 (Fig. 4). As noted above, the absolute values of the protection factors will be larger by an amount equal to the ratio of the folding rate constants at pH 4.0 and 9.0. Thus, the

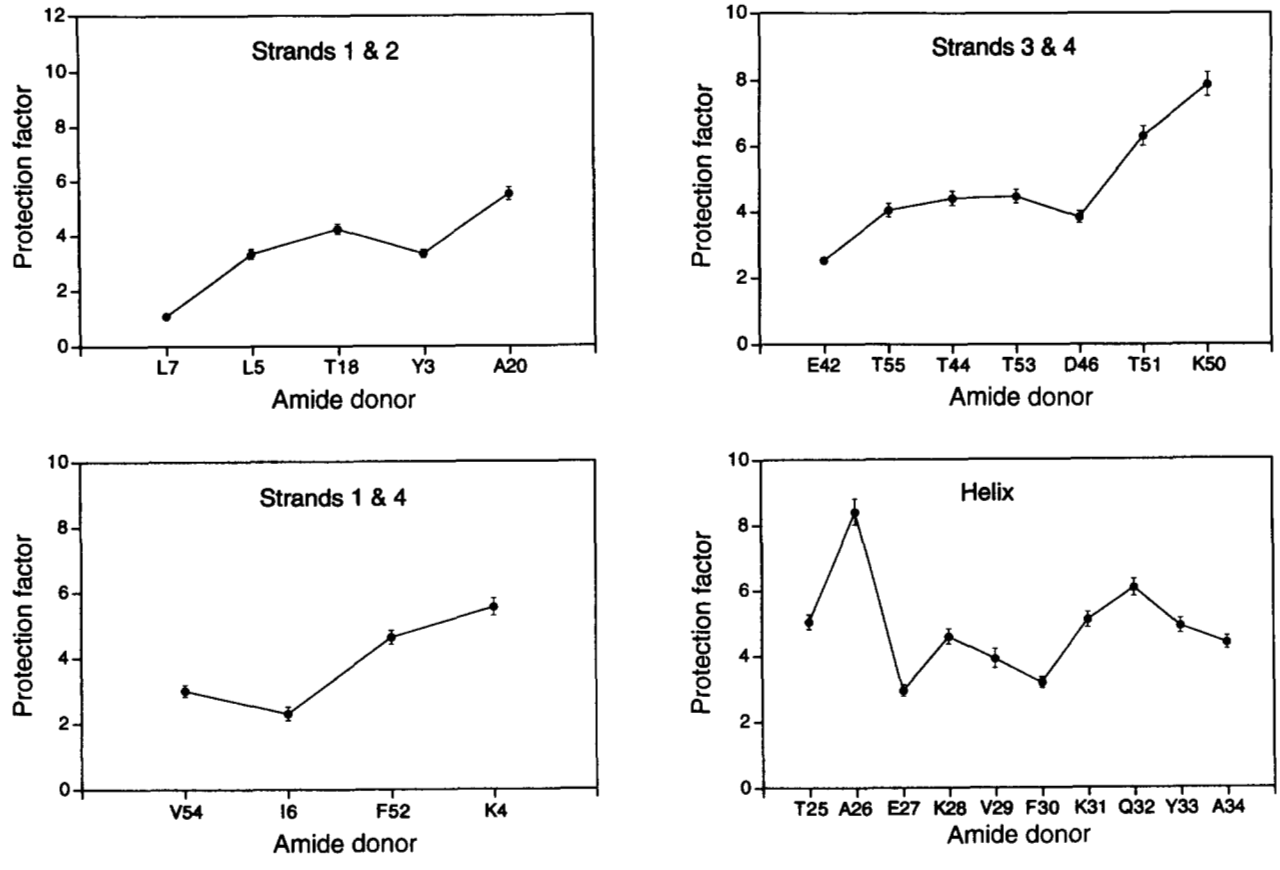

Fig. 4. Apparent amide protection factors for the apparent initial state in the folding of GB1. The protection factors are defined as the ratio of $k_{xu}(i)/k_{ex}(i)$ where $k_{xu}(i)$ is the random-coil exchange rate constant for residue i calculated as described by Bai et al. (1993) and Connelly et al. (1993), and $k_{ex}(i)$ is the corresponding exchange rate constant for the apparent initial state obtained from the global nonlinear least-squares best fit of the computed to the experimental curves using Equation 3 and shown in Figure 3. The error bars only represent the errors in the fitted values of $k_{ex}(i)$. The errors in the values of $k_{xu}(i)$, however, are at most of the order of 10%.

apparent initial state as far as the quenched-flow experiment is concerned contributes significantly to amide group protection. Several lines of evidence suggest that the protection factors relate to a state formed as a result of rapid hydrophobic collapse on the submillisecond time scale. First, recent experiments using flash photolysis to probe the folding of cytochrome c have indicated that collapse occurs on the microsecond time scale (Jones et al., 1993). Second, although nuclear Overhauser enhancement measurements have indicated the presence of partially ordered regions in the case of 2 proteins (the DNA binding domain of bacteriophage λ 434 repressor and FK506 binding protein) apparently fully denatured with urea (Neri et al., 1992; Logan et al., 1994), as well as in a recombinant model of denatured bovine pancreatic trypsin inhibitor (Lumb & Kim, 1994), these regions are strictly localized and involve only a few residues. In contrast, 13 of the 26 monitored amide groups in GB1 have apparent protection factors greater than 4, suggesting more extensive regions of structure partially shielded from solvent exchange. Third, the average backbone amide protection factor for hen egg white lysozyme unfolded in 8 M urea has a value of only 1.6, indicating that the urea-denatured state has little residual structure, even though the 4 disulfide bonds were kept intact (Buck et al., 1994). Moreover, the most protected amide group in denatured hen egg white lysozyme had a protection factor of only 5 and was located in a small cluster of protected residues in the immediate vicinity of one of the intact disulfide bridges. In comparing these values to the protection factors observed in the present work, it seems likely that the apparent initial state in the GB1 folding pathway is significantly more structured than would be expected for a fully denatured protein.

Finally, further evidence that the apparent initial state is indeed a partially collapsed state comes from stopped-flow measurements monitoring the fluorescence of the single tryptophan residue at position 43. In the native state, Trp 43 is partially exposed (Gronenborn et al., 1991; Achari et al., 1992; Gallagher et al., 1994). In the denatured state, the fluorescence intensity of Trp 43 is reduced relative to the native state (cf. Fig. 2). This reduction in fluorescence, presumably caused by increased solvent quenching, suggests that Trp 43 is more exposed to solvent in the denatured state than in the native state. We carried out stopped-flow measurements at both pH 4 and pH 9 (data not shown). In both cases, folding was too rapid to measure its rate and the fluorescence intensity of the native state was recovered within the dead-time of the stopped-flow apparatus (~5 ms). The observation that the native fluorescence intensity of Trp 43 is recovered more rapidly than the formation of stable hydrogen bonds monitored in the quenched-flow D-H exchange experiments implies that rapid partial or full collapse occurs.

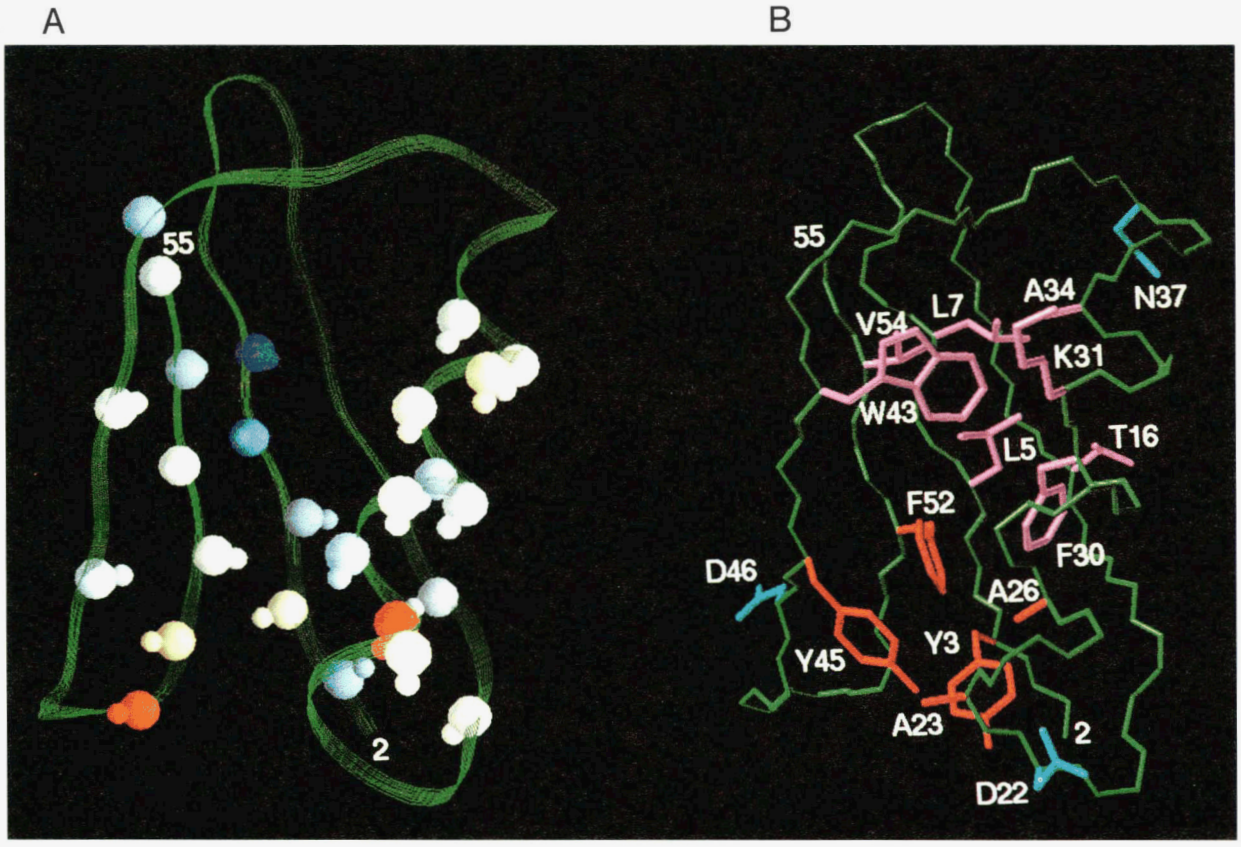

Fig. 5. A: Ribbon drawing of native GB1 with the amide groups color coded on the basis of their apparent protection factors in the apparent initial state in the folding of GB1. The amide groups are depicted as spheres and the color varies continuously from red (protection factor of 8.4) through white (protection factor of 4.7) to blue (protection factor of 0.9). **B:** Backbone of native GB1 in green with the 2 overlapping clusters of tightly packed buried hydrophobic residues in the native protein depicted in red and lilac, and the side chains involved in side-chain–backbone hydrogen bonds depicted in light blue. The pictures were produced with the programs GRASP (Nicholls, 1993) and VISP (deCastro & Edelstein, 1992), and the coordinates are taken from Gronenborn et al. (1991).

These results, however, do not necessarily imply that this collapsed state can be regarded as a definitive folding intermediate. Rather, they could be interpreted to indicate that the rapid perturbation of solution conditions caused by "instantly" diluting the chaotropic agent away may simply result in a smooth relaxation from the unfolded state to a partially collapsed state with no intervening potential barrier.

Possible structural characteristics of the collapsed state

There is a clear pattern of backbone amide protection in the apparent initial state when related to the structure of the native protein (Figs. 4, 5A). The average apparent protection factor (4.8 ± 1.8) for the residues in the antiparallel β-sheet formed by strands 3 and 4 is slightly larger than that in the parallel β-sheet formed by strands 1 and 4 (3.9 ± 1.5), which in turn is slightly larger than that in the antiparallel β-sheet formed by strands 1 and 2 (3.0 ± 1.3). Moreover, the average apparent protection factor for the residues in the helix (4.8 ± 1.6) is the same as that for the sheet between strands 3 and 4. The most protected amide groups are located in the reverse turn connecting strands 3 and 4 (Lys 50, 7.9; Thr 51, 6.3) and at the beginning of the helix (Thr 25, 5.0; Ala 26, 8.4). These apparent protection factors correspond to a stabilization of 0.9–1.2 kcal·mol^{-1} relative to a random coil state. The other amide groups in the 4-stranded sheet become progressively less protected as their distance from the reverse turn between strands 3 and 4 increases (Fig. 5A). By comparison with the structure of the native state, we suggest that 2 complementary factors are involved in the simultaneous stabilization of the reverse turn between strands 3 and 4 and the N-terminal end of the helix in the apparent initial state. In the native state these regions are in close proximity to a cluster of tightly packed buried hydrophobic residues comprising Tyr 3, Ala 23, Ala 26, Tyr 45, and Phe 52 (Fig. 5B). Thus, hydrophobic collapse centered around these residues may be an important contributing factor, resulting in the shielding of surrounding areas from solvent exchange. In addition, both the reverse turn between strands 3 and 4 and the N-terminus of the helix are stabilized in the folded state by side-chain–backbone hydrogen bonds, namely between the carboxylate side chain of Asp 46 and the NH of Ala 48, and the carboxylate of Asp 22 and the NH of Thr 25, respectively (Figs. 1, 5B). These interactions may also be formed in the apparent initial state.

Higher than average apparent protection factors are also observed around the middle of the helix (Lys 31, 5.1; Gln 32, 6.1; Tyr 33, 4.9), corresponding to a stabilization of 0.9–

1.0 kcal·mol^{-1} relative to a random coil state. Using the same arguments as above, this may be due to the formation of a hydrophobic cluster comprising Leu 5, Leu 7, Thr 16, Phe 30, the aliphatic portion of the side chain of Lys 31, Ala 34, Trp 43, Phe 52, and Val 54 (Fig. 5B), as well as the side-chain–backbone C-terminal capping interaction between the side-chain amide of Asn 37 and the backbone carbonyl of Tyr 33 (Figs. 1, 5B). In this regard, it is interesting to note that the last amides to exchange out from the native state are those of Ile 6, Val 29, Phe 30, and Val 54 (A.M. Gronenborn & G.M. Clore, unpubl. data), all of which are buried within this hydrophobic cluster.

Concluding remarks

Although previous studies have shown that GB1 folds rapidly (Alexander et al., 1992b), the present quenched-flow D–H exchange experiments clearly demonstrate that a large number of individual residues participate in this fast cooperative process. The data also suggest that the apparent initial state is different from the fully unfolded random coil state. Specifically, the apparent initial state exhibits a distinct pattern of amide protection factors and may represent a partially collapsed state. Thus, it is clear that, in the absence of complicating factors resulting in dead-end pathways, folding of a polypeptide chain to the native folded state can occur very rapidly, a result that is consistent with a number of recent theoretical studies on simplified models of protein folding (Camacho & Thirumalai, 1993; Chan & Dill, 1994; Sali et al., 1994).

Materials and methods

Sample preparation, and unfolding and refolding conditions

Uniformly (>95%) ^{15}N-labeled GB1 was expressed and purified as described previously (Gronenborn et al., 1991; Gronenborn & Clore, 1993). To assess appropriate unfolding and folding conditions for GB1, fluorescence of the single tryptophan residue was measured at pH 4.0 and 5 °C with GuHCl concentrations ranging from 0 to 6 M. Tryptophan fluorescence excitation was carried out at 295 nm and the emission was integrated from 305 to 450 nm.

Quenched-flow D–H experiments

The quenched-flow D–H experiments were carried out using a Biologic QFM-5 quench-flow apparatus, thermostated at 5 °C. Fully unfolded GB1 was prepared by taking up lyophilized GB1 at a concentration of 3 mg/mL in D_2O containing 4.5 M guanidine deuterochloride (GuDCl) and 50 mM sodium d_3-acetate, pH 4.0 (uncorrected for the deuterium isotope effect), and incubated at 37 °C overnight to ensure complete exchange of all its amide positions. Refolding was initiated by rapidly diluting the unfolded deuterated GB1 18-fold into protonated 50 mM sodium-acetate, pH 4.1. The resulting folding mixture contained 0.17 mg/mL GB1, 0.25 M GuDCl, pH 4.0. At this pH and temperature, the average exchange half-life of the completely exposed amide positions in GB1 is about 4 min. After allowing refolding to occur for varying lengths of time t_{fold}, the reaction mixture was subjected to a high pH exchange pulse of 25 ms duration. The values of t_{fold} ranged from 0 ms (obtained by subjecting unfolded GB1 directly to the high pH exchange pulse) to 370 ms. The high pH exchange pulse was achieved by the addition of 140 mM glycine, pH 10.1, in a ratio of 18:13, resulting in an exchange mixture containing 0.15 M GuDCl, 0.1 mg/mL GB1, 96% H_2O/4% D_2O, pH 9.1. ND–NH exchange was then quenched by adding 2 M acetic acid in a ratio of 31:3 to give a final pH of 3.5. The resulting samples were concentrated at 4 °C with Amicon centriprep and centricon ultrafiltration devices. Prior to NMR analysis, the buffer was exchanged for 50 mM sodium d_3-acetate, pH 4.0, in D_2O, using Amicon centricons, and the samples were incubated at 37 °C for 30 min to exchange out labile protons (Briggs & Roder, 1992). To determine the complete proton occupancy at each amide position a sample was prepared by taking up protonated, lyophilized GB1 in sodium d_3-acetate buffer, pH 4.0, and incubating it in the same way as the other samples. The proton occupancy at each amide position as a function of the folding time t_{fold} was then determined from the crosspeak intensities in a 2D ^{1}H–^{15}N heteronuclear single quantum coherence spectrum (Bax et al., 1990) recorded on a Bruker AMX-600 spectrometer. The resulting spectra were processed using the nmrPipe package (DeLaglio et al., 1994) and peak heights were measured using the program PIPP (Garrett et al., 1991). These were corrected for variations in protein concentration (~0.3–1 mM) in each sample (as measured by the intensity of the resonances in the aliphatic region of the ^{1}H-NMR spectrum). The observed intensities were expressed as fractions of the intensities observed in the "complete-occupancy" sample described above, corrected for the concentration of D_2O relative to H_2O (4:96) at the time of the exchange pulse.

Analysis of the quenched-flow D–H exchange kinetics

The complete set of quenched-flow D–H exchange data comprising 26 kinetic curves, each of which describes the proton occupancy as a function of 12 t_{fold} time points (including the "complete-occupancy" sample, which describes the true 0 time point) were fitted simultaneously by nonlinear least-squares optimization of Equation 3, using the initial value problem and optimization program FACSIMILE (Chance et al., 1979; Clore, 1983). The random coil exchange rate constants, $k_{xu}(i)$ for each residue were calculated as described by Bai et al. (1993) and Connelly et al. (1993) taking into account temperature, primary-structure effects, the pH of the exchange pulse, and the isotope effect of D–H versus H–D exchange. Although the effect of the presence of 0.15 M GuDCl in the exchange mixture on the amide exchange rates is difficult to quantify, previous studies have indicated that at most an increase of 5–10% in the random coil exchange rate constants may be observed (Loftus et al., 1986). Consequently, no attempt was made to correct for the presence of 0.15 GuDCl. The values of $k_{xu}(i)$ are as follows: Y3, 382 s^{-1}; K4, 399 s^{-1}; L5 135 s^{-1}; I6, 44.8 s^{-1}; L7, 60.4 s^{-1}; T18, 526 s^{-1}; A20, 276 s^{-1}; T25, 332 s^{-1}; A26, 618 s^{-1}; E27, 121 s^{-1}; K28, 252 s^{-1}; V29, 103 s^{-1}; F30, 163 s^{-1}; K31, 408 s^{-1}; Q32, 590 s^{-1}; Y33, 332 s^{-1}; A34, 438 s^{-1}; E42, 178 s^{-1}; T44, 258 s^{-1}; D46, 219 s^{-1}; K50, 564 s^{-1}; T51, 438 s^{-1}; F52, 356 s^{-1}; T53, 381 s^{-1}; V54, 123 s^{-1}; T55, 241 s^{-1}.

Stopped-flow fluorescence measurements

Stopped-flow tryptophan fluorescence measurements were carried out on a Bio-Logic SFM-3 stopped-flow apparatus, ther-

mostated at 5 °C. Unfolded GB1 (180 μM) in 4.5 M GuHCl, 50 mM sodium acetate, pH 4.01, was rapidly diluted 18-fold into either (1) 50 mM sodium acetate, pH 4.02, giving a refolding solution of 0.25 M GuHCl, 50 mM sodium acetate, pH 4.0, containing 10 μM GB1, or (2) into 50 mM sodium acetate, 140 mM glycine, pH 9.1, giving a refolding solution of 0.25 M GuHCl, 50 mM sodium acetate, 140 mM glycine, pH 9.0, containing 10 μM GB1. In order to determine the fluorescence intensity of unfolded GB1, the same unfolding buffer was diluted 18-fold into either 4.5 M GuHCl, 50 mM sodium acetate, pH 4.01, for the low pH control experiment or into 4.5 M GuHCl, 50 mM sodium acetate, 140 mM glycine, pH 9.1, for the high pH control experiment. Excitation was carried out at 280 nm, and fluorescence emission was measured through a long-pass filter (which blocks wavelengths below 305 nm) with a Bio-Logic PMS-200 phototube assembly. The output from the phototube was digitized every 0.5 ms. The dead-time of the stopped-flow apparatus is estimated to be about 5 ms.

Acknowledgments

We thank Drs. Attila Szabo and Dexter Kennedy for very stimulating and helpful discussions and Dr. Jay Knutson for use of his stopped-flow apparatus. This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C. and A.M.G.).

References

Achari A, Hale SP, Howard AJ, Clore GM, Gronenborn AM, Hardman KD, Whitlow M. 1992. The 1.67 Å X-ray structure of the B2 immunoglobulin domain of streptococcal protein G and comparison to the NMR structure of the B1 domain. *Biochemistry 31*:10449–10457.

Alexander P, Fahnstock SR, Lee T, Orban J, Bryan P. 1992a. Thermodynamic analysis of folding of streptococcal protein G IgG binding domains B1 and B2: Why small proteins tend to have high denaturation temperatures. *Biochemistry 31*:3597–3603.

Alexander P, Orban J, Bryan P. 1992b. Kinetic analysis of folding and unfolding of the 56 amino acid IgG-binding domain of streptococcal protein G. *Biochemistry 31*:7243–7248.

Bai Y, Milne JS, Mayne L, Englander SW. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct Funct Genet 17*:75–86.

Baldwin RL. 1993. Pulsed H/D-exchange studies of folding intermediates. *Curr Opin Struct Biol 3*:84–91.

Baldwin RL. 1994. Protein folding: Matching speed and stability. *Nature (Lond) 369*:183–184.

Bax A, Ikura M, Kay LE, Torchia DA, Tschudin R. 1990. Comparison of different modes of two-dimensional reverse correlation NMR for the study of proteins. *J Magn Reson 86*:304–318.

Briggs M, Roder H. 1992. Early hydrogen-bonding events in the folding reaction of ubiquitin. *Proc Natl Acad Sci USA 89*:2017–2021.

Buck M, Radford SE, Dobson CM. 1994. Amide hydrogen exchange in a highly denatured state: Hen egg-white lysozyme in urea. *J Mol Biol 237*:244–254.

Camacho CJ, Thirumalai D. 1993. Kinetics and thermodynamics of folding in model proteins. *Proc Natl Acad Sci USA 90*:6369–6372.

Chan HS, Dill KA. 1993. Transition states and folding dynamics of proteins and heteropolymers. *J Chem Phys 100*:9238–9257.

Chance EM, Curtis AR, Jones IP, Kirby CR. 1979. FACSIMILE: A computer program for flow and chemistry simulation, and general initial value problems. *AERE Report No R-8775*. Harwell, UK: UK Atomic Energy Research Authority.

Clore GM. 1983. Computer analysis of transient kinetic data. In: Geisow MJ, Barrett AN, eds. *Computing in biological science*. Amsterdam: Elsevier North-Holland. pp 313–348.

Connelly G, Bai Y, Jeng MF, Englander SW. 1993. Isotope effects in peptide group hydrogen exchange. *Proteins Struct Funct Genet 17*:87–92.

Creighton TE. 1994. The energetic ups and downs of protein folding. *Nature Struct Biol 1*:135–138.

deCastro E, Edelstein S. 1992. *VISP 1.0 user's guide*. University of Geneva.

DeLaglio F, Grzesiek S, Vuister G, Zhu J, Pfeifer J, Bax A. 1994. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. In: *Proceedings of the 35th Experimental Nuclear Magnetic Resonance Conference*. Abstr WP108. p 262.

Dobson CM, Evans PA, Radford SE. 1994. Understanding how proteins fold: The lysozyme story so far. *Trends Biochem Sci 19*:31–37.

Elöve G, Bhuyan AK, Roder H. 1994. Kinetic mechanism of cytochrome *c* folding: Involvement of the heme and its ligands. *Biochemistry 33*: 6925–6935.

Englander SW, Mayne L. 1992. Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu Rev Biophys Biomol Struct 21*:243–265.

Fersht AR, Dill KA. 1994. Folding and binding. *Curr Opin Struct Biol 4*: 67–68.

Gallagher T, Alexander P, Bryan P, Gilliland GL. 1994. Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. *Biochemistry 33*:4721–4729.

Garrett DS, Powers R, Gronenborn AM, Clore GM. 1991. A common sense approach to peak picking in two-, three- and four-dimensional spectra using automatic computer analysis of contour diagrams. *J Magn Reson 95*:214–220.

Gronenborn AM, Clore GM. 1993. Identification of the contact surface of a streptococcal protein G domain complexed with a human Fc fragment. *J Mol Biol 233*:331–335.

Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. 1991. A novel highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science 253*:657–661.

Jones CM, Henry ER, Hu Y, Chan CK, Luck SD, Bhuyan A, Roder H, Hofrichter J, Eaton WA. 1993. Fast events in protein folding initiated by nanosecond laser photolysis. *Proc Natl Acad Sci USA 90*:11860–11864.

Kim PS, Baldwin RL. 1990. Intermediates in the folding reactions of small proteins. *Annu Rev Biochem 59*:631–660.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure. *Proteins Struct Funct Genet 6*:87–103.

Loftus D, Gbenle GO, Kim PS, Baldwin RL. 1986. Effects of denaturants on amide proton exchange rates: A test for structure in protein fragments and folding intermediates. *Biochemistry 25*:1428–1436.

Logan TM, Theriault Y, Fesik SW. 1994. Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J Mol Biol 236*:637–648.

Lumb KJ, Kim PS. 1994. Formation of a hydrophobic cluster in denatured bovine pancreatic trypsin inhibitor. *J Mol Biol 236*:412–420.

Matthews CR. 1993. Pathways of protein folding. *Annu Rev Biochem 62*:653–683.

Miranker A, Robinson CV, Radford SE, Aplin RT, Dobson CM. 1993. Detection of transient protein folding populations by mass spectrometry. *Science 262*:896–899.

Mullins LS, Pace CN, Raushel FM. 1993. Investigation of ribonuclease T1 folding intermediates by hydrogen–deuterium amide exchange–two-dimensional NMR spectroscopy. *Biochemistry 32*:6152–6156.

Neri D, Billeter M, Wider G, Wüthrich K. 1992. NMR determination of residual structure in a urea-denatured protein. *Science 257*:1559–1563.

Nicholls AJ. 1993. *GRASP manual*. New York: Columbia University.

Pace CN, Shirley BA, Thomson JA. 1989. Measuring the conformational stability of a protein. In: Creighton TE, ed. *Protein structure – A practical approach*. Oxford, UK: IRL Press. pp 311–330.

Ptitsyn OB. 1994. Kinetic and equilibrium intermediates in protein folding. *Protein Eng 7*:593–598.

Radford SE, Dobson CM, Evans PA. 1992. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature (Lond) 358*:302–305.

Roder H. 1989. Structural characterization of protein folding intermediates by proton magnetic resonance and hydrogen exchange. *Methods Enzymol 176*:446–473.

Roder H, Elöve GA, Englander SW. 1988. Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR. *Nature 355*:700–704.

Sali A, Shakhanovich E, Karplus M. 1994. How does a protein fold. *Nature 369*:248–251.

Sosnick TR, Mayne L, Hiller R, Englander SW. 1994. The barriers in protein folding. *Nature Struct Biol 1*:149–156.

Udgaonkar JB, Baldwin RL. 1988. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature 355*:694–699.

Woodward CK. 1994. Hydrogen exchange rates and protein folding. *Curr Opin Struct Biol 4*:112–116.