Published in final edited form as:
Chem Rev. 2006 May ; 106(5): 1836–1861. doi:10.1021/cr040430y.

Early Events in Protein Folding Explored by Rapid Mixing Methods

Heinrich Roder[a,b,*], **Kosuke Maki**[a,c], and **Hong Cheng**[a]

a *Basic Science Division, Fox Chase Cancer Center, 333 Cottman Ave, Philadelphia, PA 19111*

b *Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104*

1. Introduction

Globular proteins often exhibit cooperative unfolding transitions in which only folded (native) and unfolded (denatured) molecules are populated at equilibrium.[1,2] This two-state behavior is thought to be a consequence of the fact that the native structure is stabilized by a large number of weak interactions, which are cooperative in nature, so that partially folded states are inherently unstable.[3] Even when partially structured states accumulate as transient intermediates in kinetic experiments, these are generally not sufficiently stable to be observable under strongly denaturing equilibrium unfolding conditions. On the other hand, non-native states with spectroscopic and hydrodynamic properties intermediate between those of the fully native and unfolded states (the so-called molten-globule state) often accumulate under mildly denaturing conditions, such as acidic or basic pH.[4–6] In some cases, deviations from simple two-state behavior can also be observed in denaturant-induced or thermal unfolding equilibria.[6,7] Although a stable native state that is energetically well separated from denatured states is expected to favor rapid folding into a unique structure,[8,9] there is growing evidence that kinetic intermediates are very common, even in small, single-domain proteins.[10–14] Thus, conformational states distinct from both the fully native and unfolded populations are readily accessible for many proteins, making them structurally less cooperative than thought previously.

The study of protein folding pathways is more than an academic exercise. Insight into the structural, thermodynamic and kinetic properties of protein folding intermediates is critical for understanding a wide range of diseases that can be linked to aggregation of partially denatured or misfolded forms of proteins.[15–20] Issues related to protein stability and folding also play a central role in understanding the biological consequences of mutations[21] and for designing proteins with new or modified functional properties.[22,23] Studies of protein folding mechanisms *in vitro* further provide the necessary framework for elucidating the folding of proteins in their cellular environments and other cellular processes, such as protein trafficking and degradation.[24,25]

Theoretical models and computer simulations describe the process of protein folding in terms of a diffusive motion of a particle on a high-dimensional free energy surface.[26–28] This ‘landscape’ description of protein folding predicts that a protein can choose among a large number of alternative pathways, which eventually converge toward a common free energy minimum corresponding to the native structure. In contrast, the time course of protein folding monitored by optical and other experimental probes generally shows relaxation kinetics with one or a few exponential phases, which are adequately described in terms of a simple kinetic scheme with a limited number of populated states (the chemical kinetics description). These

*Corresponding Author. Tel. 215-728-3123. Fax: 215-728-3574. *E-mail address:* H_Roder@fccc.edu.
[c]Present address: Department of Physics, University of Tokyo, Tokyo 113-0033 Japan

apparently conflicting models can be consolidated if the free energy surface is divided into several regions (basins) separated by substantial free energy barriers due to unfavorable enthalpic interaction or entropic factors (conformational bottlenecks). The protein can rapidly explore conformational space within each basin comprising a broad ensemble of unfolded or partially folded states, but has to traverse substantial kinetic barriers before entering another basin. This type of free energy surface can thus give rise to multi-exponential folding kinetics.

Progress in understanding dynamic and mechanistic aspects of protein folding has been closely linked with advances in kinetic methods. The development of commercially available stopped-flow instruments in the 1960's[29,30] enabled the first quantitative kinetic studies of protein folding reactions.[31–37] The combination of quenched-flow methods, initially developed for the study of enzyme reaction mechanisms,[38] with hydrogen exchange labeling and NMR has proven to be particularly fruitful for the structural characterization of transient folding intermediates.[39–41] Dielectric and ultrasonic relaxation measurements gave early insight into the dynamics of helix-coil transitions in homo-polypeptides,[42,43] and more recent studies on model peptides, using laser-induced temperature-jump methods,[44] established that the time window for formation of isolated helices and β-hairpins ranges from about 50 ns to several μs.[45–48] Together with other laser initiation techniques[49,50] and dynamic NMR methods,[51,52] laser T-jump studies have also been a rich source of information on the dynamics of folding on the microsecond time scale (reviewed in refs [53–56]). In particular, a number of small proteins and domains were shown to undergo two-state folding/unfolding transitions with relaxation times as short as a few microseconds.[57–65] These fast-folding proteins are of interest not only in an effort to find the ultimate "speed limit" for protein folding,[56] but also serve as bench marks for computer simulations of protein folding, which have advanced to the stage where conformational transitions on the microsecond time scale can be described with atomic resolution.[61,66–69]

The dynamics of secondary structure formation and two-state folding on the microsecond time scale has been reviewed extensively[53–55] and is not the main subject of this chapter. Instead, we will focus on recent advances in rapid mixing methods and their application to studies of early stages of protein folding, drawing mainly from work published since 1997. Solution mixing techniques have experienced a renaissance due to advances in mixer design and detection methods, which made it possible to extend the time resolution well into the microsecond time range.[70–73] Efficient turbulent mixers coupled with a variety of detection methods have yielded a wealth of information on early stages of protein folding.[61,71,74–91] Although rapid mixing techniques cannot compete with the perturbation methods mentioned above in terms of time resolution, they remain the method of choice for studies of protein folding reactions far from the equilibrium transition region where intermediate states are most likely to accumulate.

2. Rapid Mixing Techniques for Protein Folding Studies

As with any complex reaction, time-resolved data are essential for elucidating the mechanism of protein folding. Even in cases where the whole process of folding occurs in a single step, which is the case for many small proteins,[92] the kinetics of folding and unfolding provide valuable information on the rate-limiting step in folding, which, in analogy to chemical or enzymatic reaction mechanisms, is often described as an ensemble of transition states.[93] The effects of temperature and denaturant concentration give insight into activation energies and solvent-accessibility of the transition state ensemble,[1,94] and by measuring the kinetic effects of mutations, one can gain more detailed structural insight.[95–97] If the protein folding process occurs in stages, i.e., if partially structured intermediate states accumulate, kinetic studies can potentially offer much additional insight into the structural and thermodynamic properties of intermediate states and intervening barriers.[10,14,98–101]

2.1. Turbulent Mixing

Most rapid mixing schemes rely on turbulent mixing to achieve complete mixing of two (or more) solutions. Mixers of various designs are in use ranging from a simple T-arrangements to more elaborate geometries, such as the Berger ball mixer.[102] The goal is to achieve highly turbulent flow conditions in a small volume. The turbulent eddies thus generated can intersperse the two components down to the μm distance scale. However, the ultimate step in any mixing process relies on diffusion in order to achieve a homogeneous mixture at the molecular level. Given that the diffusion time t varies as the square of the distance r over which molecules have to diffuse, it takes a molecule with a diffusion constant $D=10^{-5}$ cm^2/s about 1 ms to diffuse over a distance of 1 μm ($t=r^2/D$). Thus, the mechanical mixing step has to intersperse the two components on a length scale of less than 1 μm in order to achieve sub-ms mixing times. The onset of turbulence is governed by the Reynolds number, *Re*, defined as

$$Re=\rho v d/\eta \quad (1)$$

where ρ is the density (g/cm^3), v is the flow velocity (cm/s), d describes the characteristic dimensions of the channel (cm), and η is the viscosity of the fluid (e.g., 0.01 poise for water at 20 °C). To maintain turbulent flow conditions in a cylindrical tube, *Re* has to exceed values of about 2,000.

Turbulence is important not only for achieving efficient mixing, but also for maintaining favorable flow conditions during observation. In stopped-flow and quenched-flow experiments, turbulent flow insures efficient purging of the flow lines. In continuous-flow measurements, turbulent flow conditions in the observation channel lead to an approximate “plug flow” profile, which greatly simplifies data analysis compared to the parabolic profile obtained under laminar flow conditions. The time resolution of a rapid mixing experiment is governed not only by the mixing time, which in practice is difficult to quantify, but also the delay between mixing and observation. The effective delay between initiation of the reaction and the first reliably measurable data point is defined as the dead time, Δt_d. In both stopped- and continuous-flow experiments, any unobservable volume (dead volume), ΔV, between the point where mixing is complete and the point of observation contributes an increment $\Delta t = \Delta V/(dV/dt)$ to the dead time (dV/dt is the flow rate). Additional contributions to the effective dead time include the time delay to stop the flow and any artifacts that can obscure early parts of the kinetic trace.

Increasing the flow rate promotes more efficient mixing by generating smaller turbulent eddies, and yields shorter time delays Δt, and thus should lead to shorter dead times. However this trend does not continue indefinitely. Aside from practical problems due to back pressure and, in the case of stopped-flow measurements, various stopping artifacts, the time resolution of a rapid mixing experiment is ultimately limited by cavitation phenomena.[103] Under extreme conditions, the pressure gradients across turbulent eddies can become so large that the solvent begins to evaporate, forming small vapor bubbles that can take a long time to dissolve. The result is an intensely scattering plume that makes meaningful detection of the kinetic signal virtually impossible. Rapid mixing experiments involving large changes in denaturant concentration, which is often the case in studies of protein folding or unfolding, pose additional challenges related to the increase in solvent viscosity, which slows down the mixing process, convective flow due to density gradients across the mixer, and the heat of mixing, which can give rise to thermal artifacts.

To achieve efficient turbulent mixing conditions requires high flow rates and relatively large channel dimensions (the practical limit below which backpressure becomes prohibitive seem to lie around 100 μm), which can consume substantial amounts of material. An interesting alternative to turbulent mixing with improved sample economy uses hydrodynamic focusing to mix solutions under laminar flow conditions.[104,105] By compressing a fluid phase into a

thin layer, the technique can potentially achieve mixing times in the low microsecond range. However, the slow onset of mixing has so far limited practical applications of hydrodynamic focusing to dead times near 1 ms.[73,106,107]

2.2. Stopped-flow Methods

The simple, but versatile, stopped-flow technique, coupled with optical detection (typically fluorescence, absorbance or circular dichroism), has long been the primary source of kinetic insight into protein folding reactions.[10,36,98,99,108,109] In a typical stopped-flow experiment, a few hundred μl of solution are delivered to the mixer *via* two syringes driven by a pneumatic actuator or stepper-motors. Flow rates in the range of 5 to 15 ml/s with channel diameters of the order of 1 mm insure turbulent flow conditions (*Re*>5,000). After delivering a volume sufficient to purge and fill the observation cell with freshly mixed solution, the flow is stopped abruptly when a third syringe hits a stopping block or a valve is closed. Commercial instruments can routinely reach dead times of a few ms. Recent improvements in mixer and flow-cell design by several manufacturers of stopped-flow instruments resulted in dead times well under 1 ms. The upper end of the time scale that can be reliably measured in a stopped-flow experiment is determined by the stability of the mixture in the flow cell, which is limited by convective flow or diffusion of reagents in and out of the observation volume. For slow reactions with time constants longer than a few minutes, manual mixing experiments are generally more reliable. Stopped-flow mixing is usually coupled with real-time optical observation using absorbance (UV through IR), fluorescence emission or circular dichroism (CD) spectroscopy. In addition, the stopped-flow technique has been implemented in conjunction with many other biophysical techniques, such as fluorescence lifetime measurements,[110,111] NMR,[112,113] and small-angle x-ray scattering.[114] An extension of the stopped-flow technique makes use of two or more consecutive mixing steps in order to prepare the system in a particular initial state (double-jump stopped-flow), which can provide important information on kinetic mechanisms in protein folding.[85,115]

The interpretation of stopped-flow data requires a careful calibration of the instrumental dead time by measuring a pseudo-first-order reaction tuned to the time scale of interest (i.e., a single-exponential process with a rate-constant approaching the expected dead time) and an optical signal matching the application. Common test reactions for absorbance measurements include the reduction of 2,6-dichlorophenolindophenol (DCIP) or ferricyanide by ascorbic acid.[116] A convenient test reaction for tryptophan fluorescence measurements is the irreversible quenching of N-acetyltryptophanamide (NATA) by N-bromosuccinimide (NBS). For fluorescence studies in or near the visible range, one can follow the pH-dependent association of the Mg^{2+} ion with 8-hydroxyquinoline, which results in a fluorescent chelate,[117] or the binding of the hydrophobic dye 1-anilino-8-naphlalene-sulfonic acid (ANS) to bovine serum albumin (BSA), which is associated with a large increase in fluorescence yield.[118]

2. 3. Continuous-flow Techniques

In a continuous-flow experiment, the reaction is again triggered by turbulent mixing, but, in contrast to stopped-flow, the progress of the reaction is sampled under steady-state flow conditions as a function of the distance down-stream from the mixer.[38,119] This avoids artifacts related to arresting the flow and makes it possible to use relatively insensitive detection methods. Thus, continuous-flow measurements can achieve shorter dead times compared to stopped-flow, but this comes at the expense of sample economy. Earlier versions of this experiment involved point-by-point sampling of the reaction profile while maintaining constant flow at high rates (several ml/s for a conventional mixer). The prohibitive amounts of sample consumed limited the impact of continuous-flow techniques until advances in mixer design made it possible to achieve highly efficient mixing at lower flow rates,[70,72,75,120] and an improved detection scheme allowed simultaneous recording of a complete reaction

profile in a few seconds.[72] These developments lowered both the dead time and sample consumption by at least an order of magnitude, and made routine measurements on precious samples with dead times as short as 50 μs possible.

In 1985, Regenfuss et al.[70] described a capillary jet mixer consisting of two coaxial glass capillaries with a platinum sphere placed at their junction. The reaction progress was monitored in a free-flowing jet, using conventional photography to measure fluorescence *vs.* distance from the mixer. Measurements of the binding kinetics of ANS to bovine serum albumin indicated that dead times less than 100 μs can be achieved with this mixer design. More recently, several laboratories reported continuous-flow resonance Raman and fluorescence studies of enzyme and protein folding reactions on the sub-millisecond time scale, using machined mixers with dead-times in the 100 μs range.[71,75,120] More widespread use of these methods has been hampered by a number of technical and experimental difficulties. Continuous-flow experiments involving a free-flowing jet[70,75,120,121] are fraught with difficulties due to instability and scattering artifacts. The use of a conventional camera with high-speed monochrome film for fluorescence detection[70] is inadequate due to the low sensitivity of the film in the UV region, limited dynamic range and the non-linearity of the film response. Finally, prohibitive sample consumption makes continuous-flow experiments that record a kinetic trace one point at a time feasibly only for highly abundant proteins.[71,74,79,84]

We were able to overcome many of these limitations by combining a highly efficient quartz capillary mixer, inspired by the design of Regenfuss et al.,[70] with a flow cell and an improved detection system involving a digital camera system with a UV-sensitized CCD detector.[72] A diagram of the experimental arrangement is shown in Figure 1. For solution delivery, we use a motor-driven syringe pump (Update, Madison, WI), which injects the reagents to be studied at high rate into each of the two coaxial capillaries (typically using volume ratios of 1:5 or 1:10). The outer capillary consists of a thick-walled quartz tube, which is pulled to a fine tip (~200 μm i.d. at the end), using a glassblowing lathe or a simple gravity method. A smaller inner capillary (360 μm o.d., 150–180 μm i.d.) with a platinum sphere (~250 μm diameter) suspended at the tip is positioned inside the tapered end of the outer capillary. Solid glass rods fused to the inner wall of the outer capillary (tapering down to diameter of ~20 μm) prevent the sphere from plugging the outlet. The reagents are forced through the narrow gap between the sphere and the outer wall where mixing occurs under highly turbulent flow conditions. The mixture emerging is injected into the flow channel (250 μm × 250 μm) of a fused-silica observation cell joined to the outer capillary by means of a hemispherical ground-glass joint. Typical flow rates are 0.6 to 1.5 ml/s, resulting in linear flow velocities of 10 to 25 m/s through the 0.25×0.25 mm^2 channel of the observation cell.

The reaction progress in a continuous-flow mixing experiment is measured by recording the fluorescence profile *vs.* distance downstream from the mixer. A conventional light source consisting of an arc lamp, collimating optics and monochromator, is used for fluorescence excitation. Relatively uniform illumination of the flow channel over a length of 10–15 mm is achieved by means of a cylindrical lens. A complete fluorescence *vs.* distance profile is obtained by imaging the fluorescent light emitted at a 90°angle onto the CCD detector of a digital camera system (Micromax, Roper Scientific, Princeton NJ) containing a UV-coated Kodak CCD chip with an array of 1317 × 1035 pixels. The camera is equipped with a fused silica magnifying lens and a high-pass glass filter or a band-pass interference filter to suppress scattered incident light. In a typical continuous-flow fluorescence experiment, a raw reaction profile, $I_r(d)$, is obtained by averaging the emitted light intensity across the flow channel *vs.* the distance *d* downstream from the mixer. To account for the non-uniform distribution of incident light intensity, $I_c(d)$, a matching intensity profile is recorded for a “fluorescent control” (e.g., dilution of the unfolded protein solution with denaturing buffer, or dilution of NATA with buffer

without quencher). The scattering background, $I_b(d)$, is measured by passing water (or buffer) through both capillaries. Distance is converted into time on the basis of the known flow rate and cross-sectional area of the flow channel (0.0625 mm^2 in our setup) and the length of the channel being imaged. Finally, a corrected fluorescence trace, $fl_{rel}(t)$, is obtained by taking a ratio of the background-corrected reaction and control traces,

$$fl_{rel}=(I_r - I_b)/(I_c - I_b). \quad (1)$$

To estimate the dead time of the continuous-flow experiment, we routinely measure the pseudo-first-order NATA-NBS quenching reaction at a final NATA concentration of 40 μM and several NBS concentrations in the range 2–32 mM. Figure 2*a* shows a plot of the relative fluorescence, fl_{rel}, calculated according to eq 1, along with exponential fits (solid lines). Since $fl\text{-}_{rel} = 1$ corresponds to the unquenched NATA signal expected at t=0, the intercept of the fits with $fl_{rel} = 1$ indicates the time point t=0 where the mixing reaction begins. The delay between this point and the first data point that falls onto the exponential fit corresponds to the dead time of the experiment, Δt_d, which in this example is 50 ± 5 μs. In Figure 2*b*, the rate constants obtained by exponential fitting are plotted as a function of NBS concentration. The slope of a linear fit (solid line) yields a second-order rate constant for the NBS-induced chemical quenching of NATA of 7.9×10^5 $M^{-1}s^{-1}$, which agrees well with the rate constant obtained from stopped-flow measurements at lower NBS concentration.[122] Together with the linearity of the second-order rate plot (Figure 2*b*), this documents the accuracy of the continuous-flow measurement. Instrumental dead times can be substantially longer if one or both of the solutions contains a denaturant, such as guanidine hydrochloride (GuHCl) or urea at high concentration due to a combination of viscosity-dependent decreases in mixing efficiency and thermal mixing artifacts. Thus, it is important to calibrate mixer performance for each set of solution conditions and flow rates used, for example by including the appropriate denaturant and buffers in a NATA-NBS quenching experiment. A straightforward method for assessing mixing efficiency is to compare the reaction profile measured during continuous-flow mixing of NATA with water (under the appropriate buffer/denaturant conditions) with that of a pre-mixed solution at the same final NATA concentration. For an efficient mixer at sufficiently high flow rates, we find that the ratio of these flow profiles approaches unity within a fraction of a mm downstream of the mixing region (Figure 1). Alternatively, mixing efficiency could be assessed using a very fast reaction completed within the dead time, such as the quenching of tryptophan (or NATA) by sodium iodide. To detect any flow-related artifacts, such as caviation, we simply compare the flow profiles of a dilute NATA solution (introduced into both inner and outer capillary) recorded during continuous flow and immediately after stopping the drive.

Although capillary mixers remain highly competitive in terms of mixing efficiency and dead time, some of the drawbacks of this design are the challenging manual manufacturing process, limited reproducibility and difficulty of cleaning after a blockage. Several groups have relied on various micromachining techniques to construct mixers consisting of thin (~100 μm) channels, often using a simple T-arrangement.[71,75,79,84,91,123,124] A recent example is the microfluidic mixer design by Bilsel at al.[91] illustrated in Figure 3. Using laser micromachining they introduced a series of fine (50–125 μm diameter) channels into a thin film (127 μm thick) made of a hard, chemically inert polymer (PEEK) sandwiched between two fused silica windows. The two inlet channels meet at a 90°angle with the outlet channel bisects the two, resulting in a large fluid momentum change that promotes turbulent mixing. Bilsel et al. incorporated this mixer in a laser-based optical setup capable of both steady-state and time-resolved fluorescence detection.[91] Using the NATA-NBS quenching reaction (c.f. Fig. 2), they recorded a dead time of 50 μs, which is comparable to that obtained with our capillary mixer and shorter than those reported using other microfluidic mixers.

3. Detection Methods

A major strength of rapid mixing methods is that they can be readily combined with a wide range of detection methods. Data acquisition in the stopped-flow mode requires rapid sampling of the signal at intervals shorter than the dead time while continuous-flow experiments permit steady-state spectroscopic measurements for longer time intervals limited only by the capacity of the solution delivery system. Table 1 lists common detection methods used in rapid mixing studies of the kinetics of folding, which are illustrated in this section with selected examples.

3.1. Tryptophan Fluorescence

The fluorescence emission properties of tryptophan and tyrosine side chains provide information on the local environment of these intrinsic chromophores. For example, a fully solvent-exposed tryptophan in the denatured state of a protein typically shows a broad emission spectrum with a maximum near 350 nm and quantum yield of ~0.14, similar to that of free tryptophan or its derivative, NATA. Burial of the tryptophan side chain in an apolar environment within the native state or a compact folding intermediate can result in a substantial blue-shift of the emission maximum (by as much as 30 nm) and enhanced fluorescence yield. These changes are a consequence of the decrease in local dielectric constant and shielding from quenchers, such as water and polar side chains. In other cases, close contact with certain (polar) side-chains gives rise to a decrease in fluorescence yield upon folding. Many polar amino acid side chains (as well as main chain amino groups) are known to quench tryptophan fluorescence, probably via excited-state electron or proton transfer.[125,126] Thus, the straightforward measurement of fluorescence intensity *vs.* folding or unfolding time can provide useful information on solvent accessibility and proximity to quenchers of an individual fluorescence probe. Complications due to the presence of multiple fluorophores can be avoided by using mutagenesis to replace any additional tryptophans.[127,128] Because Trp is a relatively rare amino acid, proteins with only one tryptophan are not uncommon; in the case of Trp-free proteins, a unique fluorophore can be introduced by using site-directed mutagenesis.[129,130]

The use of tryptophan fluorescence to explore early stages of protein folding is illustrated in Figure 4, which shows results on staphylococcal nuclease (SNase) recently obtained in our laboratory.[87] A variant with a unique tryptophan fluorophore in the N-terminal β-barrel domain (Trp76 SNase) was obtained by replacing the single typtophan in wild-type SNase, Trp140, with His in combination with Trp substitution of Phe76. The fluorescence of Trp76 is strongly enhanced and blue-shifted under native conditions relative to the denatured state in the presence of urea (panel *a*), indicating that upon folding the indole ring of Trp76 moves from a solvent-exposed location to an apolar environment within the native structure. An intermediate state with a fluorescence emission spectrum similar to, but clearly distinct from the native state was detected in equilibrium unfolding experiments (dashed line in Figure 4*a*). In contrast to WT* SNase (P47G, P117G and H124L background), which shows no changes in tryptophan fluorescence prior to the rate-limiting folding step (~100 ms), the F76W/W140H variant shows additional changes (enhancement) during an early folding phase with a time constant of about 80 μs (Figure 4*b*). The observation that the main folding phase has essentially the same rate for both variants (10.4 s^{-1} for WT* and 11.3 s^{-1} for the Trp76 variant) indicates that the overall folding mechanism is not perturbed by the F76W/W140H mutations (the differential behavior at long times suggest that the mutations modulate a slow-folding species containing non-native peptide-bond isomers). Thus, the observed amplitude differences can be largely attributed to the fact that Trp76 reports on the rapid formation of a hydrophobic cluster in the N-terminal β-sheet region while the wild-type Trp140 is silent during this early stage of folding.

3.2. ANS Fluorescence

Valuable complementary information on the formation of hydrophobic clusters at early stages of folding can be obtained by using ANS as extrinsic fluorescence probe.[131–133] Figure 4*c* illustrates this with recent results on the Trp76 variant of SNase introduced above.[87] Continuous-flow measurements showed a major enhancement in ANS fluorescence during early stages of SNase folding to fully native conditions (U→N), and also during formation of the compact acid-denatured state of SNase (U→A). For comparison, we also measured the kinetics of ANS binding to the pre-formed A-state. While the rate of ANS binding to the A-state shows the linear dependence on ANS concentration characteristic of a second-order binding process, the rates observed during refolding reactions ending either in the native state or the A-state level off at ~120 μM ANS. The limiting ANS-independent rate at higher concentrations is thus due to an intramolecular conformational event that precedes ANS binding. The rate of this process closely matches that of the earliest phase detected by intrinsic fluorescence of Trp76 (Figure 4*b*), confirming that both processes reflect a common early folding step. This agreement confirms that ANS can serve as a faithful probe for detecting early folding intermediates. This is especially important in view of the fact that ANS has been shown to perturb the kinetics of late folding steps,[133] and is known to shift folding equilibrium transitions.[134]

Maki et al.[87] also observed a major enhancement in ANS fluorescence during folding of WT* SNase with a time constant of 80 μs, which closely matches that of the Trp76 mutant. This further supports the notion that this rapid process reflects an intrinsic conformational event rather than being a consequence of the amino acid changes (F76W/W140H). Taken together, the tryptophan and ANS fluorescence data are consistent with the rapid accumulation of an ensemble of states containing a loosely packed hydrophobic core involving primarily the β-barrel domain. In contrast, the specific interactions in the α-helical domain involving Trp140 are formed only during the final stages of folding.

3.3. Fluorescence Energy Transfer

Fluorescence resonant energy transfer (FRET) can potentially give more specific information on the changes in average distance between fluorescence donors and acceptors. For example, horse cytochrome *c* (cyt *c*) contains an intrinsic fluorescence donor-acceptor pair, Trp59 and the covalently attached heme group, which quenches tryptophan fluorescence *via* excited-state energy transfer.[135] We have made extensive use of this property to characterize the folding mechanism of cyt *c*,[40,136,137] including the initial collapse of the chain on the microsecond time scale.[76,77]

We recently combined ultrafast mixing experiments with FRET in order to monitor large-scale structure changes during early stages of folding of acyl-CoA binding protein (ACBP), a small (86-residue) four-helix bundle protein.[83] ACBP contains two tryptophan residues on adjacent turns of helix 3, which served as fluorescence donors. An AEDANS fluorophore covalently attached to a C-terminal cysteine residue introduced via mutagenesis was used as an acceptor (Figure 5*a*). Earlier equilibrium and kinetic studies, using intrinsic tryptophan fluorescence, showed a cooperative unfolding transition and single-exponential (un)folding kinetics consistent with an apparent two-state transition. Even when using continuous-flow mixing to measure intrinsic tryptophan fluorescence changes on the sub-ms time scale (Figure 5*b*), we found only minor deviations from two-state folding behavior. However, when we monitored the fluorescence of the C-terminal AEDANS group while exciting the tryptophans, we observed a large increase in fluorescence during a fast kinetic phase with a time constant of 80 μs, followed by a decaying phase with a time constant ranging about 10 ms to 500 ms, depending on denaturant concentration (Figure 5*c*). The large enhancement in FRET efficiency is attributed to a major decrease in the average distance between helix 3 and C-terminus of

ACBP. The fact that the early changes are exponential in character suggests that the initial compaction of the polypeptide is limited by an energy barrier rather than chain diffusion. The subsequent decrease in AEDANS fluorescence during the final stages of folding is attributed to a sharp decrease in the intrinsic fluorescence yield of the two tryptophans due to intramolecular quenching. The specific side chain interactions responsible for quenching are established only in the close-packed native structure and are not present during the initial folding event. These observations indicate that the early (80 μs) folding phase marks the formation of a collapsed, but loosely packed and highly dynamic ensemble of states with overall dimensions (in terms of fluorescence donor-acceptor distance) similar to that of the native state.

3.4. Continuous-flow Absorbance

Although fluorescence is inherently more sensitive, our capillary mixing instrument can also be adapted for continuous-flow absorbance measurements on the microsecond time scale. The fully transparent flow cell used for fluorescence measurements is replaced with a custom-made partially opaque absorbance flow cell of the same dimensions (0.25 mm path length). Relatively uniform illumination with minimal changes to the optical arrangement (see Figure 1) was achieved by using a scattering cell containing an opaque solution (non-dairy creamer works well). As in fluorescence measurements, a complete reaction profile can be recorded in a single 2–3 s continuous-flow run by imaging the flow channel onto the CCD detector. Using the reduction of DCIP by ascorbic acid as a test reaction,[116] we measured dead times as short as 40 μs at the highest flow rate tested (1.1 ml/s).

As a first application of absorbance-detected continuous-flow mixing, we measured the changes in heme absorbance in the Soret region (~360–430 nm) associated with the folding of oxidized horse cyt *c*.[138] The reaction was initiated by a rapid jump from pH 2, where the protein is fully unfolded, to pH 4.7, where folding occurs rapidly with minimal complications due to non-native histidine-heme ligands. A series of kinetic traces covering the time window from 40 μs to ~1.5 ms were measured at different wavelengths spanning the Soret region (Figure 6). A parallel series of stopped-flow experiments (data not shown) was performed under matching conditions to extend the data to longer times (2 ms to 10 s). Global fitting of the family of kinetic traces to sums of exponential terms yielded three major kinetic phases with time constants of 65 μs, 500 μs and 2 ms, respectively, consistent with accumulation of two intermediate species, I and M (Scheme 1), with absorbance properties distinct from both the initial (U) and final (N) states. In previous continuous-flow fluorescence measurements on horse cyt *c*,[76,77] we also observed three kinetic phases with very similar time constants, indicating that a basic four-state mechanism is sufficient to describe the folding process of cyt *c* in the absence of complications due to non-native heme ligation and other slow events, such as *cis-trans* isomerization of peptide bonds.

3.5. Probes of Secondary Structure Formation

CD spectroscopy in the far-UV (peptide) region, which provides a measure of average secondary structure content of a protein,[139,140] has been an important tool for protein folding research (reviewed in refs [10,98,99,101,141,142]). Stopped-flow CD measurements on numerous proteins have shown that a substantial fraction of the overall change in the far-UV CD spectrum associated with refolding can appear within the instrumental dead time while changes in the near-UV region of the CD spectrum often do not occur until the late stage of folding.[143–150] Such observations have been generally interpreted in terms of a model involving the transient accumulation of early intermediates that appear native-like in terms of the CD-detectable secondary structure, but still unfolded in terms of the environment of aromatic side chains observed in the near-UV region of the CD spectrum. However, until recently the time resolution of stopped-flow CD measurements has been limited to the 10 ms range by the low inherent sensitivity of the technique and various flow artifacts, such as strain-

induced birefringence. Akiyama et al.[79] were able to record CD at folding times as short as 400 μs by coupling an efficient turbulent mixer (T-design) with a commercial CD spectrometer. Their continuous-flow measurements of CD spectral changes in the far-UV region revealed the kinetics formation of (helical) secondary structure during early stages of folding of cyt *c*[79] and apomyoglobin.[89]

Infrared spectroscopy is a promising alternative for monitoring acquisition of secondary structure during protein folding. The amide I band, which is dominated by backbone carbonyl stretching vibrations, is a valuable source of information on secondary structure,[151] especially after recent work on peptides led to a firmer assignment of the spectral contributions of specific secondary structure types and degrees of hydration.[152] However, the implementation of IR spectroscopy in conjunction with rapid mixing poses a series of technical challenges, including the problem of rapidly sampling a wide spectral range and the design of a flow cell with sufficiently short optical path-lengths to allow transmission of infrared light through the solvent. The need for protein concentrations in the mM range is another impediment. By interfacing a Fourier transform infrared (FTIR) spectrometer in rapid-scan mode with a stopped-flow mixer, it is possible to record complete IR spectra on transient states populated on the 10 ms time scale.[153–155] Marinkovic et al.[124] were able record transient IR spectra with a time resolution in the ms range by using a synchrotron as an IR radiation source in conjunction with a micro-fabricated mixer developed previously for resonance Raman studies[71]. In another recent paper Kimura et al. [156] used FTIR-detected continuous-flow measurements to monitor helix formation in poly-L-glutamic acid in the sub-ms time range.

3.6. Small-angle X-ray Scattering

Small-angle X-ray scattering (SAXS) experiments can provide unique information on the molecular dimensions and shape of proteins, including dynamic ensembles of partially or fully denatured states.[157,158] In order to monitor size and shape changes during early stages of folding, several laboratories have combined synchrotron-based SAXS measurements with stopped-flow[114,159] or continuous-flow instruments,[84,89] as well as a novel hydrodynamic mixing method.[73,106] Akiyama et al.[84] constructed a mixer/flow-cell assembly with a dead time as short as 160 μs for continuous-flow SAXS measurements on a synchrotron. This challenging experiment enabled them to follow the changes in size (radius of gyration, R_g) and shape (pair distribution derived from scattering profiles) associated with refolding of cyt *c*,[84] myoglobin,[89] ribonuclease A (RNase A)[160] and single-chain monellin.[161]

3.7. Other Detection Methods

Continuous-flow measurements have been coupled with a number of other biophysical techniques, including resonance Raman spectroscopy[71] and EPR.[80,162] In their pioneering work, Takahashi et al.[71] used resonance Raman spectroscopy to monitor changes in heme coordination during folding of cyt *c* on the sub-milliscond time scale. Their findings confirmed and extended prior results on the involvement of heme ligation in folding of cyt *c*, based on stopped-flow absorbance and fluorescence measurements.[163–165] Grigoryants et al.[80] developed a unique capillary mixer for continuous-flow measurements in an EPR spectrometer. The short relaxation times (compared to NMR) of free radicals and paramagnetic metal centers makes it possible to record EPR signals on short-lived species. Scholes and colleagues have used this technique to measure the kinetics of folding of spin-labeled yeast cyt *c* on the sub-millisecond time scale.[162]

4. Quenched-flow H/D Exchange Studies of Early Folding Events

Quenched-flow hydrogen exchange labeling experiments coupled with NMR analysis[39–41,166–168] continue to be the main source of residue-specific structural information on folding

intermediates.[82,169–174] Mass-spectrometric analysis has emerged as an alternative analytical technique for H/D exchange labeling studies of protein folding and other conformational changes.[175,176] H/D exchange labeling methods and their application to protein folding have been reviewed extensively.[101,166,177–181] We will, therefore, limit our discussion to recent NMR-based H/D exchange studies aimed at the structural characterization of early folding intermediates.

4.1. A quenched-flow Method for H/D Exchange Labeling on the μs Time Scale

Most protocols for H/D exchange labeling rely on commercial quenched-flow equipment to carry out two or three sequential mixing steps, which limits the time resolution to a few milliseconds or longer. However, amide protons are often found to become protected from H/D exchange within the dead time of quenched-flow mixing (2–10 ms), indicating that hydrogen bonded structure can form on the sub-millisecond time scale.[167,168,170,182–185]

Bökenkamp *et al.* [123] described a quenched-flow device with substantially improved time resolution (~100 μs minimum delay between sequential mixing events). They constructed a series of mixers by etching 200 μm diameter channels into silicon chips, using a microfabrication method. Each chip contains a pair of T-mixers connected by a delay line of variable length. Solutions are delivered at flow rates sufficient for efficient turbulent mixing, using a BioLogic (Claix, France) quenched-flow apparatus.

We recently devised an alternative approach for extending quenched-flow measurements into the microsecond time range.[138] The device uses a quartz capillary mixer similar to that used for optical measurements (Figure 1), but without observation cell, in order to generate a homogeneous mixture of two solutions (e.g., unfolded protein and refolding/exchange buffer). The mixture emerges from the capillary as a fine (200 μm diameter) jet with a linear velocity of up to 40 m/s at the highest flow rate used (1.25 ml/s). A second mixing event can be achieved simply by injecting the jet into a test tube containing a third solution (quench buffer); the high flow velocity ensures very efficient mixing. To determine the dead time (i.e., the shortest delay between the two mixing events), we carried out a series of H/D exchange experiments on a pentapeptide (YGGFL). Rapid exchange of the backbone amide protons with solvent deuterons was achieved by mixing an H_2O solution of the peptide with D_2O buffered at basic pH (9.7). The exchange reaction was quenched by injecting the mixture into ice-cooled acetate buffer at pH 3. Under these quench conditions, the rate of exchange for some of the peptide NH groups (Gly3, Phe4 and Leu5) is sufficiently slow (10, 45 and 70 min, respectively) to determine their residual NH intensity by recording one-dimensional ^1H NMR spectra. To vary aging time the capillary was raised from direct contact with the quench solutions to a distance of about 40 mm corresponding to an upper time limit of ~1 ms. For Gly3 and Phe4, exponential fits of the decay in residual NH intensity with the incremented time delay yields exchange rates of 5,600 and 4,400 s^{-1}, respectively, in agreement with published intrinsic exchange rates.[186] Extrapolation of the fits up to the NH intensity expected at t=0 (measured in a separate control) indicates that the first measurement corresponds to an effective exchange time of 60 ± 10 μs, thus defining the dead time of the measurement.

4.2. H/D Exchange Labeling of Early Folding Intermediates

A straightforward, yet informative, technique for measuring amide protection patterns in early folding intermediates is the burst-phase labeling method, which is closely related to the competition method used in the first study of protein folding by NMR-detected H/D exchange.[39] In contrast to the pulse labeling protocol used for following the kinetics of protection at longer times (>5 ms), the competition method requires only two sequential mixing steps, which can be applied in rapid succession, using the capillary quenched-flow device.[138] For a three-state folding mechanism with a populated intermediate (K_{UI}>1), the competition experiment

can be described by Scheme 2, where k_{cU} and k_{cI} are the exchange rates of a given NH group in the unfolded and intermediates states, respectively. $k_c(U)$ corresponds to the intrinsic (chemical) exchange rate, k_c, estimated on the basis of model peptides.[186] If $k_{cI} \ll k_{cU}$, i.e., the NH group is protected in the intermediate, the proton occupancy, Y_H, for a given amide proton will initially rise at a rate k_c+k_{UI} and approach at long times a steady-state level determined by the ratio of folding and exchange rates.[39] Under conditions where exchange and folding rates (k_{UI}) are comparable, both k_c and k_{UI} can, in principle, be determined from a series of experiments involving variation of the competition time at constant pH. However, a more straightforward experiment is to vary pH at a fixed competition time.[39] By minimizing the time period (competition time) during which the pH-dependent competition between folding and H/D exchange can occur, it is possible to measure protection factors in early intermediates, even if amide protons are only marginally protected.[167,168] In this burst-phase labeling scheme, proton occupancies, Y_H, are measured as a function of labeling pH, using a competition time of a few ms. Any shift in the Y_H vs. pH profile to higher pH, compared to that expected for an unprotected proton (measured in a control experiment in the presence of denaturant), provides an estimate for the effective protection from exchange due to accumulation of early intermediates within the competition time. If the I-state is a relatively stable (late) intermediate, the results are expressed in terms of protection factors, $P = k_{ex}(U)/k_{ex}(I)$, where $k_{ex}(U)$ and $k_{ex}(I)$ are the effective exchange rates of a given amide protons in the unfolded population and intermediate state, respectively (see also ref [187]). Although the model-peptide rates, k_c, can be used to estimate $k_{ex}(U)$, it is more reliable to measure this rate under destabilizing conditions in order to account for any residual structure in the denatured state.[168] This modification greatly improves the accuracy of the results, making it possible to measure protection factors as small as 2 (since $P = K_{IU} + 1$, this corresponds to 50% population of a hydrogen-bonded species).

To study structure formation on the sub-ms time scale, one has to extend the competition experiment to pH values above 10 where intrinsic exchange times are of the order of 1 ms or less. Under such extreme conditions, we can no longer assume that the exchange reaction is limited by the intrinsic exchange rate, k_c (a limiting case commonly referred to as EX_2 exchange).[188] If k_c exceeds the rate of folding (k_{UI}), the rate of unfolding (k_{IU}) becomes rate-limiting for exchange (EX_1 exchange). Eq 2 is a more general steady-state rate expression that covers both EX_2 as well as EX_1 exchange conditions (assuming $k_{UI} \gg k_{IU}, k_c$).[173,188–190] In eq 2,

$$Y_H(pH)=1-\exp\left(\frac{k_{IU}\times k_c(pH)}{k_{IU}+k_{UI}+k_c(pH)}\times t_p\right) \quad (2)$$

k_{UI} and k_{IU} are the rates of formation and unfolding of the intermediate, respectively, $k_c = k_c(U)$ is the intrinsic exchange rate from the unfolded state, and t_p is the duration of the exchange/folding competition time. This analysis assumes that k_c is the only pH-dependent rate, and that any direct exchange from the I-state ($k_c(I)$ in Scheme 2) is negligible compared to $k_c(U)$. The rates of formation and unfolding of intermediates are likely to become pH-dependent under very basic conditions, which may give rise to deviations from the behavior predicted by eq 2 at high pH. To describe situations with marginally stable conformational states ($k_{UI} \sim k_{IU}$), it is necessary to use the full double-exponential solution for a three-state kinetic mechanism.[173,188]

4.3. Probing the Stability of Folding Intermediates by Pulse Labeling

Residue-specific information on the structure and stability of intermediates populated at longer folding times (>ms) can be obtained by allowing the protein to fold for a fixed refolding time before applying a short labeling pulse at variable pH.[170,173,182,190–192] This labeling protocol calls for three sequential timed mixing steps at accurately timed intervals ranging from

a few ms to about 10 s, which can be achieved using commercially available quenched-flow instrumentation (shorter times are, in principle, accessible using a microfluidic device[123]). Figure 7 illustrates this approach with recent results by Krishna et al.[173], who used a pulse pH variation strategy to characterize the structural and dynamic properties of a late intermediate populated during folding of cytochrome *c*. Previous pulsed hydrogen exchange studies have shown that this intermediate accumulates within about 10 ms of refolding at pH 6 or above, contains stable structure primarily in a pair of interacting N- and C-terminal α-helices and is stabilized by the presence of a non-native histidine-heme ligand.[40,163,164] The symbols in Figure 7*a* show the pH-dependent proton occupancy for selected residues monitored by 2D NMR analysis of samples that were exposed to a 50 ms labeling pulse at variable pH following a refolding delay of 100 ms.[173] The dashed lines indicate the pH profiles expected if a particular NH group were unprotected in the intermediate. For many residues, especially those in the N- and C-terminal regions of the protein, the observed labeling profiles are displaced toward higher pH, indicating that they are involved in stable hydrogen bonds that protect them from solvent exchange. Elöve and Roder[190] have previously reported similar behavior for a subset of residues that can be resolved by 1D NMR and concluded that the complex labeling curves were indicative of a heterogeneous population of molecules folding along parallel pathways. Krishna et al.[173] were able to fit the observed pH profiles quantitatively by using a generalized version of eq 2. Assuming that exchange occurs from the unfolded state at the intrinsic exchange rate k_c,[186] the model has only two free parameters, k_{UI} and k_{IU}, which reflect the dynamics and stability of individual hydrogen bonds. A plot of the equilibrium constant for formation of the I-state, $K_{UI} = k_{UI}/k_{IU}$, *vs.* residue number (Figure 7*b*) indicates that the structured regions of the intermediate ($K_{UI} \geq 1$) are limited to the N- and C-terminal helices and a handful of intervening residues. The gray bars indicate residues (Cys14, Ala15 and His18) involved in residual hydrogen-bonded structure even under fully denaturing conditions.[163,168,193] The analysis of Krishna et al.[173] can account for the complex variation in the slope and saturation behavior of the data in terms of a transition from EX_2 (exchange-limited) to EX_1 (opening-limited) exchange kinetics. Thus, the ensemble of intermediates populated over the 10–100 ms time window appears to be more homogeneous than we thought initially.[190]

5. Early Intermediates in the Folding of Small Proteins

5.1. Stopped-flow Evidence for Rapid Folding Events

The first indications of protein folding events on the microsecond time scale came from stopped-flow experiments, which often show unresolved changes in optical signals occurring within the dead time of measurements.[145,148,170,183,194–196] To illustrate this so-called burst-phase effect, Figure 8 shows the kinetics of refolding of horse cytochrome *c* (cyt *c*)[136] measured by stopped-flow fluorescence (panel *a*) along with equilibrium fluorescence data *vs*. denaturant concentration (panel *b*). The protein was unfolded by addition of 4.5 M GuHCl, which lies in the baseline region above the cooperative unfolding transition, and the refolding reaction was triggered by 6-fold dilution with buffer (0.1 M sodium acetate, pH 5), resulting in a final GuHCl concentration of 0.75 M, well within the folded baseline region. The data points in panel *a* were recorded by sampling the fluorescence emission above 325 nm (using a glass cutoff filter) at logarithmically spaced time intervals. The first time point corresponds to the instrumental dead time of 2.5 ms, which was calibrated using the NATA-NBS test reaction (see Figure 2). The observed kinetics on the ms time scale describes a double-exponential decay (solid line) comprising a major phase with time constants of ~8 ms and a minor one with a time constant of about 500 ms. Extrapolation of the observed kinetics back to t=0 yields the initial signal, S(0), which is compared in panel *b* (arrow) with the equilibrium unfolding transition plotted on the same fluorescence scale (relative to unfolded protein at 4.5 M GuHCl). The initial signal observed in this and a series of additional stopped-flow

experiments at different final GuHCl concentrations is consistently below the relative fluorescence of the unfolded state, $S_{pred}(U)$, predicted by linear extrapolation from the unfolded baseline region to lower GuHCl concentrations (dashed line in panel *b*). The difference between the predicted and observed initial amplitude, $S_{burst} = S_{pred}(U) – S(0)$, reflects conformational events occurring within the dead time of the stopped-flow experiment (the so-called burst-phase amplitude). If the formation of intermediate(s) during the dead time is kinetically well resolved from slower folding steps, a pre-equilibrium is established between unfolded and intermediate states, whose stability can be estimated from the denaturant dependence of S_{burst} (provided that pre- and post-transition baseline values can be defined). Similar observations for many different proteins using various spectroscopic parameters gave clear evidence for the existence of rapid conformational events that cannot be resolved with conventional mixing techniques, and provided a strong incentive for the development of faster methods for triggering and observing structural changes during the first millisecond of refolding.[71,72]

A second line of evidence supporting accumulation of early folding intermediates has been based on the dependence of folding rates on denaturant concentration (usually guanidine HCl or urea), as illustrated schematically in Figure 9. In the absence of populated intermediates, a logarithmic plot of the measured rate constant of folding, k_f, *vs.* denaturant concentration (panel *a*) often describes a V-shape with a decreasing linear dependence at denaturant concentrations below the midpoint of the unfolding transition (C_m), a minimum at C_m, and a linearly increase for the rate of unfolding at denaturant concentrations above C_m (hence the name “chevron plot”).[95,197,198] As illustrated by Figure 9*b*, accumulation of intermediate states prior to the rate-limiting folding step can give rise to a non-linear folding branch of the chevron plot, usually a downward curvature with decreasing denaturant concentration. The discovery of this so-called rollover effect is often attributed to Matouschek et al.,[96] who found that the chevron plot for barnase showed a pronounced curvature at low denaturant concentrations. However, Matthews and colleagues described similar complexities in the folding kinetics of the α-subunit of tryptophan synthase as early as 1981.[199] A quantitative connection between folding intermediates and the rollover effect was established be several groups in the mid-1990’s who used quantitative kinetic models to explain the curvature in the $\log(k_f)$ *vs.* [denaturant] dependence in terms of accumulation of an early (obligatory) folding intermediate.[195,200,201] Khorasanizadeh et al.[195] further showed that both the rollover and burst-phase effects can be modeled quantitatively by a three-state folding mechanism (Scheme 3) involving rapid formation of an intermediate prior to the rate-limiting formation of the native state. In this scenario, a partially folded state is sufficiently stable to become populated at low denaturant concentrations and limits the overall rate of folding, which gives rise to both deviation from linearity in the chevron plot (Figure 9*b*) and a drop in the observable kinetic amplitude of the main (rate-limiting) folding phase (Figure 9*d*). In many cases, the observed folding times at low denaturant concentrations approach values in the millisecond range, which rules out the possibility that the rate-limiting process represents *cis*/*trans* isomerization of a proline peptide bond.[37,202,203] Moreover, the amplitude behavior shown in Figure 9*d* where the fast kinetic phase is dominant at low denaturant concentrations (a_o approaching 1) and decreases in amplitude at the expense of the slow phase (a_1) is characteristic of a sequential folding process with an on-pathway intermediate. In contrast, slow isomerization steps, such as proline isomerization, that precedes the formation of the native state generally gain amplitude under destabilizing conditions approaching the midpoint of the unfolding transition.[203]

In other cases, a complex dependence of folding or unfolding rates can occur even without appreciable population of intermediate states.[13] For example, the rate of unfolding for several c-type cytochromes was found to level off at high denaturant concentrations, giving rise to a chevron plot with a highly non-linear unfolding branch.[85,136,163,172,204,205] Sauder et al.[204] were able to fit this behavior quantitatively on the basis of a kinetic model involving a

poorly populated, but obligatory intermediate in unfolding, which leads to a change in the rate-limiting transition state under strongly destabilizing conditions. In the case of the cytochromes, this effect can be attributed to cleavage of the ligand bond between the sulfur of a methionine and the heme iron,[163,204] but a similar phenomenon has been reported for non-heme proteins. For example, Walkenhorst et al.[206] used a mechanism involving a high-energy unfolding intermediate to explain the non-linearity observed in the unfolding kinetics of SNase. In a recent comprehensive survey of about 25 proteins, Sanchez and Kiefhaber[13] showed that non-linear chevron plots are a common feature, even for proteins without detectable intermediates, and can be interpreted quantitatively in terms of sequential folding mechanisms with high-energy obligatory intermediates. Native-like intermediates can lead to curvature in the kinetics of unfolding while less ordered high-energy intermediates account for complexities in refolding. Although alternative models, such as a solvent-induced shift of a broad folding barrier,[207] cannot be ruled out in every case, a three- or four-state mechanism is the simplest scenario and provides a consistent description of the effects of mutations and solvent conditions on the kinetics of folding and unfolding.

5.2. B1 Domain of Protein G

The observation of a burst-phase, such as that in Figure 8, suggested deviations from two-state behavior for many proteins. This list includes not only large proteins with over 100 residues, but also several smaller ones.[10–12,101,129,208] An illustrative example is the 57-residues B1 domain of protein G (GB1), which is among the smallest globular protein domains that do not depend on disulfide bonds or metals for stabilization.[209] Like many other small proteins,[92] GB1 was initially thought to fold according to a two-state mechanism.[210] However, continuous-flow fluorescence measurements of the GB1 folding kinetics[78] showed clear deviations from the first-order (single-exponential) kinetics expected for a simple two-state process (Figure 10*a*). The time course of refolding from the guanidine-denatured state revealed a prominent exponential phase with a time constant of 600–700 μs followed by a second, rate-limiting, process with a time constant of 2 ms or longer, depending on denaturant concentration (Figure 10*b*). The fast phase dominates the kinetics at low denaturant concentrations and accounts for the total fluorescence change associated with the burial of Trp43 upon folding, including the previously unresolved burst-phase signal.[208] In Figure 11, the rate constants obtained by fitting two exponential terms to the observed traces and the corresponding amplitudes are plotted *vs.* denaturant concentration. The biphasic kinetics of folding observed over a range of GuHCl concentrations can be modeled quantitatively on the basis of a three-state folding mechanism (Scheme 3), where I represents an ensemble of intermediate states with native-like fluorescence properties (i.e., Trp43 is buried). The dependence of elementary rate constants, k_{ij}, on denaturant concentration c (dashed lines in Figure 11) is governed by the following relationship

$$\ln k_{ij} = \ln k_{ij}^{0} + (m_{ij}^{\ddagger}/RT) \times c \quad (3)$$

where k_{ij}^{0} represents the elementary rate constant in the absence of denaturant, and $m_{ij}^{\ddagger}$ describes its dependence on denaturant concentration (kinetic m-value). The system of linear differential equations describing Scheme 3 were solved by determining the eigenvalues and eigenvectors of the corresponding rate matrix, using standard numeric methods.[211,212] Although a three-state kinetic mechanism can be solved analytically (e.g., ref [13]), the rate-matrix approach has the advantage that it can be readily expanded to more complex first-order kinetic mechanisms. After optimizing the four elementary rate constants and corresponding m-values, the two observable rates (eigenvalues) and associated amplitudes predicted by the model (solid lines in Figures 11*a* and *b*) simultaneously fit both the observed rate profile (log (rate) *vs*. GuHCl concentration) and kinetic amplitudes at each denaturant concentration, as well as the midpoint and slope of the equilibrium unfolding transition (diamonds in panel *b*). The three-state mechanism explains the kinetic behavior at low (<1 M) and intermediate (1–3

M) GuHCl concentrations in terms of two distinct kinetic limits reminiscent of the EX_1 and EX_2 limits of hydrogen exchange.[188] At low denaturant concentration, the intermediate is stable and transiently populated (the free energy of I in Figure 11*c* is lower than that of U), and the rate of the slower folding phase approaches the elementary rate constant of the final folding step, k_{IN}. At the same time, the fast (sub-ms) folding phase gains amplitude at the expense of the slow (rate-limiting) phase. At intermediate denaturant concentrations, the intermediate is destabilized (Figure 11*c*) and no longer accumulates. However, if I is an obligatory intermediate, the observed folding rate approaches the limiting value $k_f = K_{UI} \times k_{IN}$, where $K_{UI} = k_{UI} / k_{IU}$ is the equilibrium constant of the U ⇔ I transition. Thus, the sharp decrease in the net folding rate as the denaturant concentration approaches the midpoint of the unfolding transition is explained in terms of the unfavorable U ⇔ I pre-equilibrium involving a high-energy intermediate.

Three-state analysis of a complete set of kinetic data, such as that of GB1 reported by Park et al.,[78] provides a comprehensive description of the folding mechanism in terms of the size of the free-energy barriers separating the intermediate from the unfolded and native states (labeled TS_1 and TS_2 in Figure 11*c*) and the stability of the intermediate ($\Delta G = -RT \ln (k_{UI}/k_{IU})$. In addition, the denaturant-dependence of the four elementary rate constants provides valuable insight into the changes in solvent-accessible surface area associated with each transition. In Figure 11*c*, each state and intervening transition state is labeled with the corresponding α-value (sometimes called β_T in reference to Tanford[1]) obtained by the cumulative kinetic m-values with respect to the total equilibrium m-value. According to this analysis, the initial barrier, TS_1, represents a well solvated ensemble of states (α =0.29) while both I and TS_2 are nearly as solvent-shielded as the native state (α =0.85). The relative changes in fluorescence associated with the two folding phases of GB1 provide additional insight into the structural properties of the intermediate. The large increase in fluorescence during the fast phase (Figure 11), which accounts for nearly all of the fluorescence change at equilibrium, indicating that Trp43 becomes largely buried already during the initial phase of folding. Given the central location of Trp43 at the interface between the C-terminal β-hairpin and the α-helix, this indicates that the intermediate contains a well-developed hydrophobic core.

Krantz et al. recently questioned the validity of our analysis and argued that the folding kinetics of GB1 should be modeled as a two-state process.[213] This implies that the complete time course of folding can be fitted by a single exponential. In contrast, we find that a satisfactory fit of our combined continuous- and stopped-flow data requires two exponential phases, while a single exponential fit leads to non-random residuals of the order of 10%, which is unacceptable, given the quality of the data (Figure 10*a*). It should be noted that the two phases differ sufficiently in rate (>3.5-fold) to make their separation unambiguous. This conclusion is strengthened by the absence of additional, slower phases in GB1, which contains no proline residues. Below 1 M GuHCl, where the slower phase of our double-exponential fit levels off (Figure 11*a*), the apparent rate obtained by single-exponential fitting continues to increase, approximating the linear chevron behavior of a two-state system (cf. Figure 9*a*). This phenomenon is explained by the fact that the approximate rate obtained by fitting a single exponential represents a weighted average of the rates obtained by double-exponential fitting (see section 5.3). Because the fast phase dominates at low and the slow phase at higher denaturant concentrations, the result is a relatively linear rate profile. We further note that the population of the intermediate becomes negligible at GuHCl concentrations approaching the transition region, resulting in an apparent two-state unfolding equilibrium. Thus, contrary to Krantz et al.,[213] the predicted equilibrium behavior cannot be used to discriminate between the two models. Finally, the folding kinetics of GB1 was found to be independent of protein concentration,[208] indicating that intermolecular interactions are not involved in stabilizing the intermediate.

As a further validation of the three-state model, we used the elementary rate constants obtained by modeling the data in Figure 11*a* along with the relative fluorescence values for the various states (panel *b*) to predict the time course of folding. The fact that we can simultaneously reproduce all of the curves measured at different denaturant concentration (solid lines in Figure 10*b*) is a compelling demonstration that Scheme 3 is fully consistent with all observations (see refs [87,206] for earlier applications of this global kinetic modeling approach). Alternative three-state mechanisms with nonproductive or nonobligatory intermediates lead to somewhat poorer fits of the data at low denaturant concentration, but cannot be ruled out definitively on the basis of available data (see section 6.4). However, a simple two-state mechanism with a linear chevron plot (cf. ref [213]) leads to poor predictions of the observed time course of folding at all except the lowest and highest GuHCl concentrations for which we have both continuous- and stopped-flow data (dashed lines in Figure 10*b*).

5.3. Ubiquitin

The 76-residue α/β protein ubiquitin is another well-studied small protein for which three-state folding behavior has been reported under some conditions. Khorasanizadeh et al.[129,195] found deviations from two-state behavior in the folding kinetics of a tryptophan-containing ubiquitin variant (F45W mutant), including a downward curvature in the rate profile (log(rate) *vs*. [denaturant] plot) and a concomitant drop in the relative amplitude of the main folding phase at low denaturant concentration. They were able to account for both phenomena (rollover and burst phase) in terms of a three-state mechanism with an obligatory on-path intermediate (Scheme 3). As detailed above for GB1, this simple scheme explains the leveling-off of the rate constant and diminishing amplitude of the principal (rate-limiting) folding phase at GuHCl concentrations below 1 M. However, alternative explanations cannot be ruled out without direct observation of the inferred fast folding phase. Moreover, the uncertainty in the burst-phase amplitude is substantial if the main observable folding phase approaches the dead time of the kinetic experiment, which is the case for ubiquitin under stabilizing conditions. Krantz and Sosnick[214] remeasured the folding kinetics of F45W ubiquitin, using a stopped-flow instrument with a dead time of ~1 ms. They found a linear rate profile for the main folding phase and no indications of a burst phase, and concluded that our earlier evidence for a folding intermediate was based on fitting artifacts.

In an effort to settle this debate, we recently revisited the folding kinetics of F45W ubiquitin, [14] combining continuous-flow fluorescence measurements with stopped-flow experiments with an improved dead time, which allowed us to continuously monitor the time course of folding from about 100 μs to 100 s (see Figure 12*a* for two representative kinetic traces). The initial increase in fluorescence relative to the GuHCl-unfolded state with a time constant τ_0 = 150 μs becomes more pronounced in the presence of sodium sulfate (data not shown) and is consistent with a decrease in the solvent accessibility of the fluorophore, Trp45, at an early stage of folding. However, further studies are required to determine whether this process reflects formation of a folding intermediate or a nonspecific collapse event. The subsequent fluorescence decay is attributed to intramolecular quenching of Trp45 fluorescence upon folding.[215] A thorough kinetic analysis, using multi-exponential fitting functions, indicates that a minimum of four distinct phases are required to obtain a satisfactory fit of the data over the 0.1 to 1 s time window (Figure 12*a*); a minor additional fluorescence decay at longer times has been attributed to proline isomerization.[129] In contrast, if we use only three exponentials (including one increasing and two decaying phases) to fit the data over the same time window, we obtain nonrandom residuals with amplitudes much larger than the scatter of the data points (Figure 12*b*). Under the most stabilizing conditions studied (0.5 M GuHCl), a satisfactory fit of the main fluorescence decay from 300 μs to 10 ms requires two distinct phases, λ_1 and λ_2, with time constants τ_1 = 1.8 and τ_2 = 7.1 ms, respectively. Another phase (τ_3 ~ 70 ms) is necessary to account for a minor decay observed between 20 and 200 ms. The fact that both

the rate and amplitude of phase λ_3 are essentially independent of denaturant concentration points toward a heterogeneous process, perhaps involving a minor population of molecules containing non-native (*cis*) proline isomers. Although this process is clearly too fast to be assigned to proline isomerization, the presence of a *cis* peptide bond in a critical location may slow down a conformational folding step. A plot of the rate constants for the three decaying phases *vs*. GuHCl concentration (Figure 13) shows a pronounced rollover for λ_2 (squares) with a denaturant-independent regime below 0.75 M followed by a linear decrease above 1 M GuHCl, whereas λ_1 shows a shallow upward curvature. As in the case of GB1 (Figure 11), this behavior is fully consistent with a three-state mechanism (Scheme 3). Figure 13 also shows the apparent rates for the main decaying phase obtained by fitting only three exponentials (symbol x). Since they represent an average of λ_1 and λ_2 weighted according to the relative amplitudes, they continue to increase with decreasing GuHCl concentration below 1 M. Thus, the controversy whether an intermediate accumulates during folding of ubiquitin again boils down to a curve fitting issue. In our initial study, we detected only the slower process, λ_2, which exhibits the rollover and missing-amplitude effects indicative of a burst-phase intermediate.[195] By extending their stopped-flow measurements to shorter times (~1 ms), Krantz and Sonick[214] obtained a faster rate without apparent rollover corresponding to a weighted average of the two underlying processes. In a recent stopped-flow study, Went et al.[216] found that F45W ubiquitin exhibits three-state folding kinetics under some conditions (in the presence of GuHCl or urea plus salt) and two-state behavior under other conditions (in the presence of urea at low ionic strength). They further report that the folding rate for F45W ubiquitin shows moderate variation with protein concentration, suggesting that transient association may stabilize the intermediate. All of these observations, including our earlier findings on the effects of core mutations,[195] can be explained by the presence of a marginally stable intermediate, which affects the kinetics of folding only under sufficiently stabilizing conditions.

The examples presented in this and the preceding sections highlight some of the curve fitting problems that often plague the kinetic analysis of complex processes. Although any unimolecular system, irrespective of the number of states involved, can be described by a system of linear differential equations whose solution is a sum of exponential terms[211], it is often difficult or impossible to determine the appropriate number of terms to be used in fitting an individual kinetic trace. While the problem can in principle be approached using Laplace transforms or maximum entropy methods, these require extremely high data quality to yield meaningful results. For example, in the case of ubiquitin (Figure 12), the residuals obtained with three exponentials are clearly not satisfactory, and addition of a fourth term seems well justified, but the residuals are still not perfect and could be improved further by introducing a fifth phase. However, at some point we are beginning to fit systematic errors, such as baseline drifts, digitization noise, stop artifacts, etc., which are inevitable especially in stopped-flow experiments. A related problem is that the number of apparent phases depends on the exact conditions used. For example, increasing concentrations of denaturant typically simplify the kinetics as folding intermediates are destabilized. At some (arbitrary) point one has to drop a phase, which leads to a discontinuity in amplitudes.

Many of the pitfalls of empirical data fitting can be avoided by “fitting” a kinetic mechanism globally to a whole family of kinetic traces recorded as a function of an extrinsic variable, such as denaturant concentration, stabilizing salts, temperature or pH. As illustrated above for the case of GB1 (section 5.2, Figure 10), this circumvents the problem of having to chose the order of the fitting function, and fully exploits both rate and amplitude information. However, for more complex systems (more than three states), it is nontrivial and very time consuming to explore alternative scheme in an effort to determine the minimal mechanism (i.e., the least complex one consistent with the data). We have in the past used similar strategies to elucidate the folding mechanism of SNase.[87,206] With as many as seven states populated along two parallel pathways, SNase is a far more challenging system than GB1, which tests the limits of

our manual fitting procedures (numeric instabilities for certain combinations of rate constants have thus far prevented us from implementing the rate-matrix method in combination with nonlinear least-squares fitting algorithms).

5.4. Cytochrome c

Cyt *c* has played a central role in the development of new kinetic approaches for exploring early events in protein folding. The presence of a covalently attached heme group in this 104-residue protein (Figure 14*a*) serves as a useful optical marker, and its redox and ligand binding properties provide unique experimental opportunities for rapid initiation and observation of folding.[49,50,71] Its sole tryptophan residue, Trp59, is located within 10 Å of the heme iron (Figure 14*a*), resulting in efficient fluorescence quenching through a Förster-type energy transfer mechanism. Strongly denaturing conditions (e.g., guanidine HCl concentrations >4 M, or acidic pH at low ionic strength) result in a large increase in Trp59 fluorescence (up to ~60% of that of free tryptophan in water) indicative of an expanded chain conformation with an average tryptophan-heme distance greater than 35 Å.[75,135] While numerous studies have shown that folding of oxidized cyt *c* is accompanied by changes in coordination of the heme iron,[40,71,136,148,163,164,217–219] these complications can be largely avoided by working at mildly acidic pH (4.5–5) where the protein is still stable, but histidine residues are protonated and no longer can bind to the heme iron,[74,163] or by mutation of His33, which is primarily responsible for non-native heme ligation.[165]

Our capillary mixing apparatus[72] enabled us to resolve the entire fluorescence-detected folding kinetics of cyt *c*, including the elusive initial collapse of the chain[76]. Figure 14*b* shows the decay in Trp59 fluorescence observed during refolding of acid-unfolded cyt *c* (pH 2, 10 mM HCl) induced by a pH jump to native conditions (pH 4.5, 22 °C). The continuous-flow data covering the time range from about 50 μs to 1 ms are accurately described by a biexponential decay with a major rapid phase (time constant 59 ± 6 μs) and a minor process in the 500 μs range. This fit extrapolates to an initial fluorescence of 1.0 ± 0.05 (relative to the acid-unfolded protein) at t = 0, indicating the absence of additional, more rapid fluorescence changes that remain unresolved in the 50 μs dead time. By combining continuous-flow data with kinetic measurements on a conventional stopped-flow instrument, we were able to follow the time course of folding over six orders of magnitude in time.[76] An Arrhenius plot of the rate of the initial phase (Figure 14*b*, inset) yields an apparent activation enthalpy of 30 kJ/mol, which is significantly larger than that expected for a diffusion-limited process.[220] In subsequent laser T-jump studies, Hagen, Eaton and colleagues[221,222] detected a relaxation process with similar rates and activation energy, confirming the presence of a free energy barrier between unfolded and collapsed conformations of cyt *c*.

In other continuous-flow experiments, we measured the kinetics of folding of cyt *c* starting from either the acid-unfolded (pH 2, 10 mM HCl) or the GuHCl-unfolded state (4.5 M GuHCl, pH 4.5 or pH 7), and ending under various final conditions (pH 4.5 or pH 7 and GuHCl concentrations from 0.4 to 2.2 M).[76] In each case, we observed a prominent initial decay in fluorescence with a time constant ranging from 25 to 65 μs. In particular, the rate of the initial phase, measured under the same final conditions (pH 4.5, 0.4 M GuHCl), was found to be independent of the initial state (acid- or GuHCl-unfolded). These observations clearly indicate that a common rate-limiting step is encountered during the initial stages of cyt *c* folding. The large amplitude of the initial phase (as much as 70% of the total change in fluorescence under strongly native conditions) is consistent with the formation of an ensemble of compact states, which was confirmed by more recent SAXS experiments.[84,106]

Using their continuous-flow CD setup with a dead time of 400 μs, Akiyama et al.[79] were able to resolve the changes in the far-UV region of the CD spectrum associated with the second (τ~500 μs) and third (τ~2 ms) phase of folding of oxidized cyt *c*. The initial folding phase (τ

~60 μs) observed in our absorbance- and fluorescence-detected experiments (Figures 6 and 14), [76,116] appears as a burst phase in their continuous-flow CD data. The CD signal prior to the 400 μs phase indicates that about 20% of the native α-helical structure is present at this stage. This helix content is comparable to that of the acid-unfolded state at pH 2 (which retains some residual helical structure), but lower than that of the guanidine-unfolded form. Thus, a small, but measurable, amount of helical secondary structure is present already during the initial collapse of the cyt *c* polypeptide chain. In a subsequent study, Akiyama et al.[84] used continuous-flow SAXS measurements on a synchrotron to follow the changes in radius of gyration (R_g) associated with the various stages of the cyt *c* folding reaction. Starting with the expanded acid-denatured state characterized by a radius of gyration, R_g, of 24 Å, they observed accumulation of a partially collapsed state (R_g = 20.5 Å) within the 160 μs dead time of their experiment. In a subsequent phase on the ms time regime, the protein passes through a second, more compact intermediate (R_g = 18 Å) before reaching the native state (R_g = 13.9 Å). These findings confirm that cyt *c* undergoes a partial chain collapse during the initial folding phase, which corresponds to the process on the 10–100 μs time scale detected via the heme-induced quenching of Trp59 fluorescence in our previous microsecond mixing measurements of Trp-heme fluorescence energy transfer (Figure 14)[76] and heme absorbance changes (Figure 6), [138] as well as recent T-jump experiments by Hagen and colleagues.[221,222]

5.5. Apomyoglobin

Apomyoglobin (apoMb) is another important model protein that has been the subject of numerous studies focusing on early stages of folding.[89,171,174,183,223–230] On extracting the heme, the apo-form of myoglobin, retains its tightly packed globular structure[226,231, 232] along with seven α-helical segments (labeled A–E, G and H in Figure 15; the F-helix is disordered in the absence of the heme).

Uzawa et al.[89] recently capitalized on their unique capability of measuring secondary structure and molecular dimensions with submillisecond time resolution to explore the folding of horse apoMb. By combining continuous-flow CD measurements[79] with conventional stopped-flow data, the authors were able to measure the formation of helical secondary structure associated with the folding of acid-denatured apoMb (triggered by a rapid jump in pH from 2.2 to 6) over the time range from 300 μs to 1 s. The changes in the CD signal at 222 nm are consistent with a stepwise increase in helix content (expressed in terms of the fraction of helical residues, f_H) in three kinetic phases, including a major process occurring within the instrumental dead time (burst phase) and two observable phases with time constants of 5 and 50 ms, respectively. Continuous-flow SAXS measurements[84] under matching folding conditions showed a major collapse of the chain within the 300 μs dead time of the experiment, followed by a further decrease in chain dimensions during the final (50 ms) folding phase. Uzawa et al.[89] interpret their findings in terms of a four-state folding mechanism (Figure 15),[233] based on their observation of three distinct folding phases, as well as previous kinetic data by Jamin et al. [229]. Figure 15 also shows cartoons of the various states populated during apoMb folding consistent with the recent information on size and secondary structure, as well as additional data on the location of stable helical segments obtained in previous NMR-based hydrogen exchange labeling studies.[171,183] The acid-unfolded state, U, contains only a small amount of α-helical structure detectable by CD and is significantly expanded compared to the native state. However, detailed NMR studies by Dyson, Wright and coworkers[234] show clear deviations from the behavior expected for a random coil, including evidence for fluctuating residual secondary structure and restricted backbone motions. The high helix content of the I_1 intermediate (~60% of the native level) indicates that it contains helical segments in regions outside the A, G and H helix core observed in the initial pulsed hydrogen exchange study. [183] Likely candidates are the B and E helices, which, according to a more recent hydrogen exchange labeling study,[171] are partially solvent-protected in an ensemble of early

intermediates corresponding to I_1. Although these experiments cannot directly discriminate native and non-native interactions, there is recent evidence, based on the effect of mutations in the B-helix on amide protection patterns, indicating that the early intermediate contains primarily native-like helix-helix interactions.[235] The shape information (pair distribution function) extracted from the SAXS measurements at early folding times further supports a bipartite structure with a compact core comprising the bulk of the chain and some disordered segments.[89] Interestingly, the five folded helices correspond to peaks in the hydrophobicity plot of the myoglobin sequence while the still unfolded C and D helices are among the most polar regions of the protein. The I_2 intermediate formed during a later stage of folding is characterized by a further increase in helix content to about 80% of the native level, which can be explained by an increase in the length of existing helices (amide protection data detect no newly formed helices at this time). However, the overall size and shape of the protein remains unchanged, consistent with a partially hydrated (swollen) conformation containing some disordered loops.

Together with previous T-jump evidence for very rapid conformational events,[224,227] the recent microsecond mixing results[89] demonstrate that major structural event, both in terms of molecular dimensions and helix formation, occur within the first few hundred microseconds of initiating the apoMb folding reaction. Although earlier folding times remain to be explored, secondary structure formation and chain compaction appear to be coupled during the early stages of folding. This conclusion is consistent with continuous-flow fluorescence results by Jamin et al.[229] indicating that a partially structured equilibrium state with properties similar to the I_1 intermediate exhibits two-state folding/unfolding kinetics with a folding time of ~50 μs in the absence of denaturant (obtained by extrapolation of data in the presence of urea). Cyt *c*, the only other protein whose submillisecond folding kinetics has been followed by CD and SAXS, also shows a large decrease in overall dimensions on 100 μs time range,[73,76,84] but the helix content of the early intermediate is lower than that of apoMb.[79] Thus, the exact balance between chain compaction and secondary structure is likely to vary considerably from one protein to another.

6. Significance of Early Folding Events

6.1. Barrier-limited Folding vs. Chain Diffusion

In the preceding sections we discussed the folding kinetics of six proteins studied by microsecond mixing, GB1, cyt *c*, ubiquitin, ACBP, SNase and apoMb, all of which exhibit two or more distinct folding phases with major conformational changes occurring on the microsecond time scale.[76,78,81,83,87] Other examples include cytochrome c_{551} from *Pseudomonas aeruginosa*,[85] β-lactoglobulin,[82] a Trp-containing variant of RNase A,[88] and the bacterial immunity protein Im7,[81] which will be discussed further below. For several proteins (GB1, ubiquitin, Im7, β-lactoglobulin, SNase and RNase), we observed an increase in intrinisc fluorescence during the fast phase, which can be attributed to the burial of a tryptophan within an apolar environment. For three proteins (cyt *c*, cytochrome c_{551} and ACBP), we measured a large increase in energy transfer efficiency during the fast phase, confirming that these early conformational events include a large-scale collapse of the polypeptide chain. Since all of these proteins require a few ms or longer for traversing the rate-limiting barrier for folding, there is no doubt that important conformational events, including a large-scale collapse of the chain, occur prior to the ultimate barrier in the acquisition of the native structure. In every case the initial phase follows an exponential time course, indicating that a discrete free energy barrier separates the ensemble of compact states from the more expanded ensemble of unfolded states. In contrast, if the collapse of the polypeptide chain were a continuous process governed by diffusive dynamics,[8] one might expect non-exponential ('distributed') kinetics with a broad distribution of time constants.[26,236] In fact, Gruebele and colleagues[59,237,238] observed such distributed folding kinetics in temperature-jump

experiments under conditions where folding can proceed downhill from a saddle point in the free energy surface. However, other theories predict that polymer collapse can be characterized by a single relaxation time[239], and the observation of exponential kinetics may not always be a sufficient condition for the presence of an activation barrier.[240] Nevertheless, the case for a barrier-limited process is strengthened considerably if the process also exhibits a significant activation enthalpy, as observed for cyt *c*.[76] Finally, comparison of the pH-induced folding traces of cyt c monitored by Trp59 fluorescence (Figure 14) and heme absorbance (Figure 6) indicates that the time constant of the initial phase is independent of the probe used (59 ± 6 μs by fluorescence compared to 65 ± 4 μs by absorbance). This observation strongly supports the presence of a barrier, since a diffusive collapse process is expected to lead to probe-dependent kinetics, as observed in a recent T-jump study.[238]

The existence of a prominent free energy barrier between expanded and compact states justifies the description of the initial folding events in terms of a model involving an early folding intermediate. The first large-scale conformational change during folding can, thus, be described as a reversible two-state transition. The finding that the rate of collapse in cyt *c* is insensitive to initial conditions[76] suggests that the barrier may represent a general entropic bottleneck encountered during the compaction of the polypeptide chain, which is consistent with the moderate activation enthalpy associated with the initial transition (Figure 14). Bryngelson *et al.*[26] classified this type of folding behavior as a 'type I' scenario where a free energy barrier arises because of the unfavorable reduction in conformational entropy, which can be compensated by favorable enthalpic and solvent interactions only during the later stages of collapse.

6.2. Chain Compaction: Random Collapse vs. Specific Folding

Although denatured proteins often retain far more structure than is expected for a random coil polymer,[241] they are substantially more expanded than the native state.[242] Therefore, the folding process must be accompanied by a net decrease in chain dimensions. However, there are persistent controversies surrounding the question whether chain contraction occurs prior to or concurrent with the rate-limiting folding step, and whether compact states are the result of random hydrophobic collapse or a more specific structural events.[89,100,233,243–245] Our earlier conclusion that the optical changes on the submillisecond time scale (burst-phase events) seen for cyt *c* and many other proteins[10,101] reflect the formation of productive folding intermediates has been challenged by Englander and coworkers[244,246] on the basis of fluorescence and CD measurements on two heme-containing peptide fragments of cyt *c* (residues 1–65 and 1–80, respectively). Although the fragments are unable to assume a folded structure, they nevertheless show rapid (<ms) denaturant-dependent fluorescence and far-UV CD changes resembling the burst-phase behavior of the intact polypeptide chain (cf. Figure 8). Assuming that the fragments remain fully unfolded under all conditions, Sosnick *et al.*[244,246] concluded that both the optical properties of the fragments and the burst-phase changes seen during refolding of intact cyt *c* reflect a rapid solvent-dependent readjustment of the unfolded polypeptide chain rather than formation of partially folded states. However, this conclusion is inconsistent with our observation that refolding of both GuHCl- and acid-denatured cyt *c* is accompanied by an exponential fluorescence decay,[76] which indicates that the ensemble of states formed on the submillisecond time scale are separated by a free energy barrier from the unfolded states found immediately after adjustment of the solvent conditions. That these states are not just part of a broad distribution of more or less expanded denatured conformations is further supported by the observation of non-random amide protection patterns.[168] Apparently, both the full length and truncated forms of cyt *c* can assume a compact ensemble of states upon lowering of the denaturant concentration. This leads to the prediction that the fragments will also exhibit an exponential fluorescence change on a time scale similar to that of the intact protein[76] (40–60 μs at 22 °C and GuHCl concentrations of 0–0.4 M). This

prediction was confirmed in a recent study by Qiu et al.,[222] who observed an exponential relaxation process with a time constant of 20–30 μs at room temperature and a substantial activation energy (60–70 kJ/mol) for the 1–65 and 1–80 fragments of cyt *c*, using a laser T-jump apparatus.

In a related paper on RNase A, Englander and colleagues[247] reported that the GuHCl-dependence of the CD signal for the fully disulfide-reduced form closely matches the initial kinetic amplitude observed in stopped-flow refolding experiments on the protein with all four disulfide bonds intact. As in the case of the cyt *c* fragments, Qi *et al.*[247] assumed that reduced RNase remains in a random denatured state even under nondenaturing solvent conditions, which needs to be explored further. As an alternative interpretation, we again suggest that both in the presence and absence of disulfide bonds, RNase A rapidly forms a nonrandom ensemble of states, which is consistent with the shallow, but distinctly sigmoidal GuHCl dependence of the far-UV CD signal.[247] Since the native RNase A structure relies on disulfide bonds for its stability, folding of the reduced form cannot proceed beyond this early intermediate while the oxidized protein continues to fold to the native state. This scenario is consistent with recent a recent microsecond mixing study in which Kumura et al.[160] used CD and SAXS to early conformational events in oxidized and reduced RNas A.

The folding pathway of apomyoglobin as sketched in Figure 15 is a striking example of a complex folding mechanism where a series of conformational ensembles with increasing structural organization precede the final, rate-limiting, step in the acquisition of the native structure. The intermediate populated on the submillisecond time regime already contains a partially structured core giving rise to non-spherical contributions to the x-ray scattering profile, a distinctly helical far-UV CD component[89] and solvent-shielded amide protons in a cluster of interacting α-helices.[235] These properties are far from random, arguing against the notion that rapid conformational changes may reflect a random collapse of the chain driven by non-specific hydrophobic interactions.[213] For all but the simplest proteins, the rate-limiting transition state for folding cannot be reached in a single, concerted step, resulting in transient accumulation of partially structured conformations. Rapid secondary structure formation and concomitant compaction of the chain, guided by a few critical tertiary contacts, may limit the conformational space to be searched during folding, providing the bias toward the native structure necessary for efficient folding on a relatively smooth energy surface.[26]

6.3. Kinetic Role of Early Folding Intermediates

Several small proteins exhibit two-state behavior under certain conditions (e.g., elevated denaturant concentration, destabilizing mutations), but show evidence for populated intermediates (burst phase and/or nonlinear rate profiles) under other, more stabilizing, conditions.[50,75,83,129,195,248–250] Thus, apparent two-state behavior can be considered a limiting cases of a multistate folding mechanism with unstable (high-energy) intermediates[13,195]. On the other hand, there are many well documented examples of small proteins that show two-state folding behavior even under stabilizing conditions (reviewed in refs [92,251]). In several cases, the rate of folding at low denaturant concentrations was found to extend into the submillisecond time range and could be resolved only by methods such as NMR lineshape analysis,[51,252] electron transfer triggering,[253] pressure-jump[254] or T-jump techniques.[56]

While these findings clearly indicate that many small proteins can fold rapidly without going through intermediates, there is little support for the notion that early intermediates slow down the search for the native conformation,[164,255,256] except in cases where they contain features inconsistent with the native fold, such as nonnative proline isomers, metal ligands[163,164,257] or intermolecular interactions.[258] In fact, solution conditions that favor accumulation of intermediates, such as low denaturant concentrations or stabilizing salts, generally accelerate

the overall rate of folding.[10] Conversely, mutations that destabilize or abolish early intermediates often result in much slower rates of folding,[195,259,260] which is a strong indication that they are productive states. Although continuous-flow results are fully consistent with a sequential mechanism involving obligatory intermediates,[76,78,83,138] it has been difficult to rigorously demonstrate that early intermediates are obligatory states on a direct path to the native state (Scheme 3). Alternative mechanisms involving formation of nonproductive states (Scheme 4) or mechanisms with parallel pathways and non-obligatory intermediates (Scheme 5) can be ruled out only if (i) the two transitions are kinetically coupled (i.e., they have similar rates), (ii) both phases are directly observable and kinetically resolved, and (iii) the experimental probe used to monitor folding can discriminate native from intermediate and unfolded populations. Under these circumstances, the transient accumulation of an obligatory intermediate leads to a detectable lag in the appearance of the native population whereas an off-path intermediate gives rise to a rapid increase in the population of N during the initial phase. Clear evidence for such a lag phase was obtained in a stopped-flow fluorescence study on a proline-free variant of staphylococcal nuclease.[206] On the other hand, a thorough kinetic analysis, using a double-jump protocol, indicated that a late intermediate during folding of lysozyme is non-obligatory.[115] Other efforts to detect a lag phase during folding of interlukin-1β[261,262] and apoMb[230] were inconclusive because of the disparate time scales of early and rate-limiting folding events. In order to determine the kinetic role of early folding intermediates, it is necessary to directly measure the population of native molecules on the submillisecond time scale, either by double-jump experiments or *via* a specific spectroscopic probe for the native state.

A particularly favorable protein for investigating the kinetic importance of early folding intermediates is the bacterial immunity protein Im7, an 86 residue protein with a simple four-helix bundle structure (Figure 16*a*). Radford and colleagues have previously shown that Im7 populates a hyperfluorescent intermediate during the 2.5 ms dead time of their stopped-flow instrument.[249] However, the data could be fitted equally well to models involving either on- or off-pathway intermediates. Capaldi et al.[81] extended these measurements into the microsecond time range, using our continuous-flow mixing instrument to detect the changes in fluorescence of a single tryptophan (Trp75) associated with folding under various conditions (Figure 16*b*). The initial increase in fluorescence above the level of the denatured state in 6 M urea over the 100 μs to 1 ms time range is consistent with the rapid formation of a compact state resulting in a solvent-shielded environment for Trp75. The native structure is formed within a few milliseconds or longer (depending on the final concentration of urea), which is accompanied by partial quenching of Trp75 fluorescence due to close contact with a histidine. The tight kinetic coupling of the two folding events, together with the distinct fluorescence properties of the three conformational states involved, allowed us to discriminate among various possible kinetic mechanisms. The time course of folding predicted by solving the kinetic equations corresponding to the on-path mechanism (Scheme 3) accurately describe the observed behavior (Figure 16*b*, solid line). By contrast, the off-pathway mechanism (Scheme 4) failed to reproduce the data, especially at low urea concentrations where the intermediate is well populated (Figure 16*b*, dashed line). The amplitude of the submillisecond folding phase predicted by the off-pathway mechanism was far too small to account for the initial increase in the signal associated with the formation of a hyperfluorescent intermediate. Likewise, the quality of the fits does not improve when a second, parallel pathway is introduced (Scheme 5) and deteriorates when more than 25% of the flux bypassed the intermediate. These findings were further confirmed by a more detailed analysis of the dependence of the two observable rate constants on urea concentration (Figure 16*c*).[81] While Scheme 3 reproduces all of the kinetic data, there are major discrepancies in the rate of the fast phase predicted by Schemes 4 and 5, which can thus be ruled out.

By eliminating alternative mechanisms, we have, thus, been able to show that accumulation of a compact intermediate on the submillisecond time scale is a productive and obligatory event in folding of Im7.[81] This supports the notion that rapid formation of compact states can facilitate the search for the native conformation. Interestingly, a recent mutational study revealed that the Im7 folding intermediate contains non-native tertiary interactions among the three major α-helices, A, B and D, while the short helix C forms only during the final stages of folding.[263] Apparently, intermediates can be kinetically productive even if they contain some structural features not found in the final native state.

7. Conclusions and Implications

We have focused this review on recent developments in rapid mixing techniques coupled with improved biophysical techniques for monitoring conformational changes associated with early stages of protein folding. The findings clearly indicate that many proteins fold in stages with significant conformational events occurring within milliseconds of refolding, long before the rate-limiting step in formation of the native state. These early structural events can be attributed to accumulation of partially folded intermediates, which may act as stepping stones in finding the unique native conformation. Table 2 lists the time constants of the earliest folding phase measured by continuous-flow fluorescence for some of the proteins discussed in this chapter along with their size and structural type. While the number of proteins studied is still too small to warrant general conclusions, it is interesting to note that the initial folding times vary by an order of magnitude, but show no apparent correlation with protein size. This argues against the notion that the early stages of folding are dominated by a nonspecific hydrophobic polymer collapse, in which case the rate of the initial phase would depend primarily on the size of the hydrophobic core and would be insensitive to structural details.

There appears to be no simple relationship between the time scale of early folding events and secondary structure content. For example, the β-sheet containing proteins, GB1 and SNase, are near opposite ends of the time range, and Im7 has a ~5-fold slower initial rate compared to the structurally similar ACBP. Thus, the individual structural and topological features have a profound effect already during early stages of folding. The rate of the early folding events also shows no apparent correlation with protein stability (e.g., ACBP and Im7 have similar stabilities, but are near opposite ends of the rate scale). The finding that the earliest detectable folding event in GB1 domain is an order of magnitude slower than that of cyt *c* may well be related to the fact that GB1 contains β-sheet structure while cyt *c* is primarily α-helical (Figure 14*a*). With its parallel pairing of β-strands from opposite ends of the chain, GB1 has a high relative contact order (a metric of fold complexity that was found to correlate with the logarithm of the folding time of proteins with two-state mechanisms[264]). Such a trend is not expected for proteins with multistate folding mechanisms, since the rate-limiting step in structurally and kinetically more complex proteins is likely to involve structural events unrelated to the overall fold, such as side chain packing and specific tertiary interactions. However, if compact early intermediates have a native-like chain topology, one might expect a correlation between the time constant of the initial phase and contact order. To test this prediction, we include in Table 2 the relative contact order calculated according to Plaxco et al.[264] Some of the trends seen are consistent with this idea. For instance, cyt c and ACBP have low contact order and collapse much faster than GB1, which has the highest contact order among the proteins studied. The surprisingly fast initial phase of SNase (75 μs) in comparison to the much smaller GB1 (600 μs) may be related to the fact that the folding step in SNase detected by Trp76 and ANS fluorescence reflects a relatively local structural event within its antiparallel β-barrel domain, whereas the fluorescence of Trp43 in protein G may report on a more global conformational change involving the parallel pairing of N- and C-terminal β-strands. Finally, we note that several small proteins (or isolated domains of larger proteins) can complete the process of folding on a comparable time scale as the early stages of folding for some of the multi-state

proteins discussed here[58,61,64,265] (see ref [56] for additional examples). The underlying conformational events may be similar, but the fast-folding two-state proteins can directly reach the native state in a concerted collapse/folding event while most proteins encounter additional steps after crossing the initial barrier.

Acknowledgements

We are indebted to past and present group members for their invaluable contributions. We thank O. Bilsel, C. R. Matthews, M. Krishna and S. W. Englander for allowing us to include their figures. We thank the reviewers for their insightful comments and suggestions. The work was supported by grants GM056250 and CA06927 from the National Institutes of Health, a grant from the National Science Foundation (MCB-079148) and an Appropriation from the Commonwealth of Pennsylvania.

Biographies

Hong Cheng was born and raised in Shanghai, China. He received a B.Sc. degree (1983) from the Fudan University, Shanghai, China. After earning his Ph.D. in Physical Chemistry from the City University of New York in 1991 under the direction of Prof. William V. Sweeney, he moved to the University of Wisconsin at Madison as a postdoctoral research associate to study iron-sulfur proteins by using NMR under the direction of Prof. John L. Markley. In 1995 he moved to the Fox Chase Cancer Center where he currently is a staff scientist and manager of the NMR facility. The focus of his research is on protein structure, function and folding.

Kosuke Maki was born in Tokyo, Japan in 1969. He received his B.Sc. in physics from University of Tokyo in 1994 and his Ph.D. in physics from University of Tokyo in 1999 under the supervision of Prof. Kunihiro Kuwajima. He spent two years as a postdoctoral fellow in Prof. Masahiro Iwakura's group at the Advanced Institute of Science and Technology (AIST), Japan. Then he worked with Prof. Heinrich Roder at the Fox Chase Cancer Center as a postdoctoral associate from 2001 to 2003, where he explored the mechanisms of protein folding, especially in the early stages of the folding by means of the ultrarapid mixing technique. In 2003, he rejoined Prof. Kunihiro Kuwajima's group as a research associate where he continues to work on protein folding problems.

Heinrich Roder was born and raised in Switzerland. He studied Physics at the Federal Polytechnic (ETH) in Zürich and earned his Ph.D. in Biophysics under the direction of Prof. Kurt Wüthrich (1981). In 1984, following postdoctoral work with Prof. Hans Frauenfelder at the University of Illinois at Urbana, he joined the faculty of the Department of Biochemistry and Biophysics at the University of Pennsylvania in Philadelphia (Assoc. Prof. 1990; Adj. Prof. 1994). In 1991 he moved across town to the Fox Chase Cancer Center where he is currently a Senior Member and Chair of the Program in Biomolecular Structure and Function. Much of his research career has been devoted to the protein folding problem, focusing mainly on kinetic mechanisms and the role of intermediates. He made critical contributions to the development of NMR-based hydrogen exchange labeling methods and rapid mixing techniques.

Abbreviations

ACBP
: acyl-CoA binding protein

AEDANS
: 5-((((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid

ANS
: 1-anilino-8-naphlalene-sulfonic acid

ApoMb
: apomyoglobin

BSA
: bovine serum albumin

CD
: circular dichroism

CCD
: charge coupled device

cyt *c*
: cytochrome *c*

DCIP
: 2,6-dichlorophenolindophenol

EPR
: electron paramagnetic resonance

FRET
: fluorescence resonant energy transfer

FTIR
: Fourier transform infrared

GB1
B1 domain of protein G

GuHCl
guanidine hydrochloride

H/D
hydrogen/deuterium

NATA
N-acetyltryptophanamide

NBS
N-bromosuccinimide

NMR
nuclear magnetic resonance

PEEK
polyetheretherketone

RNase
ribonuclease A

SAXS
small-angle X-ray scattering

SNase
staphylococcal nuclease

T-jump
temperature jump (laser-induced)

References

1. Tanford C. Adv Protein Chem 1968;23:121. [PubMed: 4882248]
2. Privalov PL. Adv Protein Sci 1979;33:167.
3. Creighton TE. Biochem J 1990;270:1. [PubMed: 2204340]
4. Ptitsyn OB. Adv Protein Chem 1995;47:83. [PubMed: 8561052]
5. Kuwajima K. Proteins: Struct, Funct, Genet 1989;6:87. [PubMed: 2695928]
6. Fink AL. Annu Rev Biophys Biomol Struct 1995;24:495. [PubMed: 7663125]
7. Privalov PL. J Mol Biol 1996;258:707. [PubMed: 8637003]
8. Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS. Protein Sci 1995;4:561. [PubMed: 7613459]
9. Sali A, Shaknovich E, Karplus M. Nature 1994;369:248. [PubMed: 7710478]
10. Roder H, Colón W. Curr Opin Struct Biol 1997;7:15. [PubMed: 9032062]
11. Clarke AR, Waltho JP. Curr Opin Biotechnol 1997;8:400. [PubMed: 9273848]
12. Bilsel O, Matthews CR. Adv Protein Chem 2000;53:153. [PubMed: 10751945]
13. Sanchez IE, Kiefhaber T. J Mol Biol 2003;325:367. [PubMed: 12488101]
14. Roder, H.; Maki, K.; Latypov, RF.; Cheng, H.; Shastry, MCR. Protein folding handbook. In: Buchner, J.; Kiefhaber, T., editors. Part I. Wiley-VCH; Weinheim: 2005.
15. Prusiner SB. Proc Natl Acad Sci U S A 1998;95:13363. [PubMed: 9811807]
16. Kelly JW. Nat Struct Biol 2000;7:824. [PubMed: 11017183]
17. Rochet JC, Lansbury PT Jr. Curr Opin Struct Biol 2000;10:60. [PubMed: 10679462]
18. Selkoe DJ. Nature 2003;426:900. [PubMed: 14685251]

19. Dobson CM. Nature 2003;426:884. [PubMed: 14685248]
20. Kuwata K, Matumoto T, Cheng H, Nagayama K, James TL, Roder H. Proc Natl Acad Sci U S A 2003;100:14790. [PubMed: 14657385]
21. Joerger AC, Ang HC, Veprintsev DB, Blair CM, Fersht AR. J Biol Chem 2005;280:16030. [PubMed: 15703170]
22. Betz SF, Raleigh DP, DeGrado WF. Curr Opin Struct Biol 1993;3:601.
23. Hill RB, Raleigh DP, Lombardi A, DeGrado WF. Acc Chem Res 2000;33:745. [PubMed: 11087311]
24. Spiess C, Meyer AS, Reissmann S, Frydman J. Trends Cell Biol 2004;14:598. [PubMed: 15519848]
25. Young JC, Agashe VR, Siegers K, Hartl FU. Nat Rev Mol Cell Biol 2004;5:781. [PubMed: 15459659]
26. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Proteins: Struct, Funct, Genet 1995;21:167. [PubMed: 7784423]
27. Dobson CM, Karplus M. Curr Opin Struct Biol 1999;9:92. [PubMed: 10047588]
28. Onuchic JN, Wolynes PG. Curr Opin Struct Biol 2004;14:70. [PubMed: 15102452]
29. Gibson QH, Milnes L. Biochem J 1964;91:161. [PubMed: 5833381]
30. Chance, B. Rapid Mixing and Sampling Techniques in Biochemistry. Academic Press; New York: 1964.
31. Schechter AN, Chen RF, Anfinsen CB. Science 1970;167:886. [PubMed: 5410853]
32. Ikai A, Tanford C. Nature 1971;230:100. [PubMed: 4927005]
33. McPhie P. Biochemistry 1972;11:879. [PubMed: 5059893]
34. Ikai A, Fish WW, Tanford C. J Mol Biol 1973;73:165. [PubMed: 4347736]
35. Ikai A, Tanford C. J Mol Biol 1973;73:145. [PubMed: 4689947]
36. Tanford C, Aune KC, Ikai A. J Mol Biol 1973;73:185. [PubMed: 4347737]
37. Hagerman PJ, Baldwin RL. Biochemistry 1976;15:1462. [PubMed: 4087]
38. Gutfreund H. Methods Enzymol 1969;16:229.
39. Roder H, Wüthrich K. Proteins: Struct, Funct, Genet 1986;1:34. [PubMed: 2835760]
40. Roder H, Elöve GA, Englander SW. Nature 1988;335:700. [PubMed: 2845279]
41. Udgaonkar JB, Baldwin RL. Nature 1988;335:694. [PubMed: 2845278]
42. Schwarz G, Seelig J. Biopolymers 1968;6:1263. [PubMed: 5669466]
43. Gruenwald B, Nicola CU, Lustig A, Schwarz G, Klump H. Biophys Chem 1979;9:137. [PubMed: 427245]
44. Phillips CM, Mizutani Y, Hochstrasser RM. Proc Natl Acad Sci USA 1995;92:7292. [PubMed: 7638183]
45. Williams S, Causgrove P, Gilmanshin R, Fang KS, Callender RH, Woodruff WH, Dyer RB. Biochemistry 1996;35:691. [PubMed: 8547249]
46. Thompson PA, Eaton WA, Hofrichter J. Biochemistry 1997;36:9200. [PubMed: 9230053]
47. Muñoz V, Thompson PA, Hofrichter J, Eaton WA. Nature 1997;390:196. [PubMed: 9367160]
48. Huang CY, Klemke JW, Getahun Z, DeGrado WF, Gai F. J Am Chem Soc 2001;123:9235. [PubMed: 11562202]
49. Jones CM, Henry ER, Hu Y, Chan CK, Luck SD, Bhuyan A, Roder H, Hofrichter J, Eaton WA. Proc Natl Acad Sci USA 1993;90:11860. [PubMed: 8265638]
50. Pascher T, Chesick JP, Winkler JR, Gray HB. Science 1996;271:1558. [PubMed: 8599112]
51. Huang GS, Oas TG. Proc Natl Acad Sci USA 1995;92:6878. [PubMed: 7624336]
52. Spector S, Raleigh DP. J Mol Biol 1999;293:763. [PubMed: 10543965]
53. Callender RH, Dyer RB, Gilmanshin R, Woodruff WH. Annu Rev Phys Chem 1998;49:173. [PubMed: 9933907]
54. Gruebele M. Annu Rev Phys Chem 1999;50:485. [PubMed: 15012420]
55. Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, Hofrichter J. Annu Rev Biophys Biomol Struct 2000;29:327. [PubMed: 10940252]
56. Kubelka J, Hofrichter J, Eaton WA. Curr Opin Struct Biol 2004;14:76. [PubMed: 15102453]

57. Mayor U, Johnson CM, Daggett V, Fersht AR. Proc Natl Acad Sci U S A 2000;97:13518. [PubMed: 11087839]
58. Qiu L, Pabit SA, Roitberg AE, Hagen SJ. J Am Chem Soc 2002;124:12952. [PubMed: 12405814]
59. Yang WY, Gruebele M. Nature 2003;423:193. [PubMed: 12736690]
60. Kubelka J, Eaton WA, Hofrichter J. J Mol Biol 2003;329:625. [PubMed: 12787664]
61. Mayor U, Guydosh NR, Johnson CM, Grossmann JG, Sato S, Jas GS, Freund SM, Alonso DO, Daggett V, Fersht AR. Nature 2003;421:863. [PubMed: 12594518]
62. Wang M, Tang Y, Sato S, Vugmeyster L, McKnight CJ, Raleigh DP. J Am Chem Soc 2003;125:6032. [PubMed: 12785814]
63. Nguyen H, Jager M, Moretto A, Gruebele M, Kelly JW. Proc Natl Acad Sci U S A 2003;100:3948. [PubMed: 12651955]
64. Zhu Y, Alonso DO, Maki K, Huang CY, Lahr SJ, Daggett V, Roder H, DeGrado WF, Gai F. Proc Natl Acad Sci U S A 2003;100:15486. [PubMed: 14671331]
65. Wang T, Zhu Y, Gai F. J Phys Chem 2004;B108:3694.
66. Duan Y, Kollman PA. Science 1998;282:740. [PubMed: 9784131]
67. Snow CD, Nguyen H, Pande VS, Gruebele M. Nature 2002;420:102. [PubMed: 12422224]
68. Snow CD, Sorin EJ, Rhee YM, Pande VS. Annu Rev Biophys Biomol Struct 2005;34:43. [PubMed: 15869383]
69. Daggett V, Fersht AR. Trends Biochem Sci 2003;28:18. [PubMed: 12517448]
70. Regenfuss P, Clegg RM, Fulwyler MJ, Barrantes FJ, Jovin TM. Rev Sci Instrum 1985;56:283.
71. Takahashi S, Yeh SR, Das TK, Chan CK, Gottfried DS, Rousseau DL. Nat Struct Biol 1997;4:44. [PubMed: 8989323]
72. Shastry MCR, Luck SD, Roder H. Biophys J 1998;74:2714. [PubMed: 9591695]
73. Pollack L, Tate MW, Darnton NC, Knight JB, Gruner SM, Eaton WA, Austin RH. Proc Natl Acad Sci U S A 1999;96:10115. [PubMed: 10468571]
74. Yeh SR, Takahashi S, Fan B, Rousseau DL. Nat Struct Biol 1997;4:51. [PubMed: 8989324]
75. Chan CK, Hu Y, Takahashi S, Rousseau DL, Eaton WA, Hofrichter J. Proc Natl Acad Sci USA 1997;94:1779. [PubMed: 9050855]
76. Shastry MCR, Roder H. Nat Struct Biol 1998;5:385. [PubMed: 9587001]
77. Shastry MCR, Sauder JM, Roder H. Acc Chem Res 1998;31:717.
78. Park SH, Shastry MCR, Roder H. Nat Struct Biol 1999;6:943. [PubMed: 10504729]
79. Akiyama S, Takahashi S, Ishimori K, Morishima I. Nat Struct Biol 2000;7:514. [PubMed: 10881201]
80. Grigoryants VM, Veselov AV, Scholes CP. Biophys J 2000;78:2702. [PubMed: 10777766]
81. Capaldi AP, Shastry RMC, Kleanthous C, Roder H, Radford SE. Nat Struct Biol 2001;8:68. [PubMed: 11135674]
82. Kuwata K, Shastry R, Cheng H, Hoshino M, Batt CA, Goto Y, Roder H. Nat Struct Biol 2001;8:151. [PubMed: 11175905]
83. Teilum K, Maki K, Kragelund BB, Poulsen FM, Roder H. Proc Natl Acad Sci U S A 2002;99:9807. [PubMed: 12096190]
84. Akiyama S, Takahashi S, Kimura T, Ishimori K, Morishima I, Nishikawa Y, Fujisawa T. Proc Natl Acad Sci U S A 2002;99:1329. [PubMed: 11773620]
85. Gianni S, Travaglini-Allocatelli C, Cutruzzola F, Brunori M, Shastry MC, Roder H. J Mol Biol 2003;330:1145. [PubMed: 12860134]
86. Khan F, Chuang JI, Gianni S, Fersht AR. J Mol Biol 2003;333:169. [PubMed: 14516751]
87. Maki K, Cheng H, Dolgikh DA, Shastry MC, Roder H. J Mol Biol 2004;338:383. [PubMed: 15066439]
88. Welker E, Maki K, Shastry MC, Juminaga D, Bhat R, Scheraga HA, Roder H. Proc Natl Acad Sci U S A 2004;101:17681. [PubMed: 15574490]
89. Uzawa T, Akiyama S, Kimura T, Takahashi S, Ishimori K, Morishima I, Fujisawa T. Proc Natl Acad Sci USA 2004;101:1171. [PubMed: 14711991]

90. Garcia P, Bruix M, Rico M, Ciofi-Baffoni S, Banci L, Ramachandra Shastry MC, Roder H, de Lumley Woodyear T, Johnson CM, Fersht AR, Barker PD. J Mol Biol 2005;346:331. [PubMed: 15663948]
91. Bilsel O, Kayatekin C, Wallace LA, Matthews CR. Rev Sci Instr 2005;76:014302.
92. Jackson SE. Fold Des 1998;3:R81. [PubMed: 9710577]
93. Fersht AR, Matouschek A, Serrano L. J Mol Biol 1992;224:771. [PubMed: 1569556]
94. Tanford C. Adv Protein Chem 1970;24:1. [PubMed: 4912353]
95. Matthews CR. Methods Enzymol 1987;154:498. [PubMed: 3431461]
96. Matouschek A, Kellis JT Jr, Serrano L, Fersht AR. Nature 1989;340:122. [PubMed: 2739734]
97. Jackson SE, elMasry N, Fersht AR. Biochemistry 1993;32:11270. [PubMed: 8218192]
98. Matthews CR. Annu Rev Biochem 1993;62:653. [PubMed: 8352599]
99. Evans PA, Radford SE. Curr Opin Struct Biol 1994;4:100.
100. Baldwin RL, Rose GD. Trends in Biochemical Science 1999;24:77.
101. Roder, H.; Elöve, GA.; Shastry, RMC. Mechanisms of protein folding. Pain, RH., editor. Oxford University Press; New York: 2000.
102. Berger RLBB, Chapman HF. Rev Sci Instrum 1968;39:493. [PubMed: 5641804]
103. Zeff BW, Lanterman DD, McAllister R, Roy R, Kostelich EJ, Lathrop DP. Nature 2003;421:146. [PubMed: 12520296]
104. Knight JB, Vishwanath A, Brody JP, Austin RH. Phys Rev Lett 1998;80:3863.
105. Pabit SA, Hagen SJ. Biophys J 2002;83:2872. [PubMed: 12414719]
106. Pollack L, Tate MW, Finnefrock AC, Kalidas C, Trotter S, Darnton NC, Lurio L, Austin RH, Batt CA, Gruner SM, Mochrie SG. Phys Rev Lett 2001;86:4962. [PubMed: 11384392]
107. Russell R, Millett IS, Tate MW, Kwok LW, Nakatani B, Gruner SM, Mochrie SG, Pande V, Doniach S, Herschlag D, Pollack L. Proc Natl Acad Sci U S A 2002;99:4266. [PubMed: 11929997]
108. Kim PS, Baldwin RL. Annu Rev Biochem 1982;51:459. [PubMed: 6287919]
109. Roder H, Shastry MCR. Curr Opin Struct Biol 1999;9:620. [PubMed: 10508774]
110. Beechem JM, James L, Brand L. SPIE Proceedings 1990;1204:686.
111. Jones BE, Beechem JM, Matthews CR. Biochemistry 1995;34:1867. [PubMed: 7849046]
112. Balbach J, Forge V, van Nuland NAJ, Winder SL, Hore PJ, Dobson CM. Nat Struct Biol 1995;2:865. [PubMed: 7552710]
113. Frieden C. Biochemistry 2003;42:12439. [PubMed: 14580188]
114. Segel DJ, Bachmann A, Hofrichter J, Hodgson KO, Doniach S, Kiefhaber T. J Mol Biol 1999;288:489. [PubMed: 10329156]
115. Kiefhaber T. Proc Natl Acad Sci USA 1995;92:9029. [PubMed: 7568066]
116. Tonomura B, Nakatani H, Ohnishi M, Yamaguchi-Ito J, Hiromi K. Anal Biochem 1978;84:370. [PubMed: 626384]
117. Brissette P, Ballou DP, Massey V. Anal Biochem 1989;181:234. [PubMed: 2510550]
118. Gibson, QH.; Antonini, E. Hemes and Hemeproteins. Chance, B.; Estabrook, RW.; Yonetani, T., editors. Academic Press; New York: 1966.
119. Hartridge H, Roughton FJW. Proc Royal Soc (Lond): A Math Phys Sci 1923;104:376.
120. Takahashi S, Ching Y-c, Wang J, Rousseau DL. J Biol Chem 1995;270:8405. [PubMed: 7721733]
121. Paeng K, Paeng I, Kincaid J. Anal Sci 1994;10:157.
122. Peterman BF. Anal Biochem 1979;93:442. [PubMed: 464271]
123. Bökenkamp D, Desai A, Yang X, Tai YC, Marziuff EM, Mayo SL. Anal Chem 1998;70:232.
124. Marinkovic NS, Adzic AR, Sullivan M, Kovacs K, Miller LM, Rousseau DL, Yeh S, Chance MR. Rev Sci Instrum 2000;71:4057.
125. Chen Y, Barkley MD. Biochemistry 1998;37:9976. [PubMed: 9665702]
126. Van Gilst M, Tang C, Roth A, Hudson B. Fluorescence 1994;4:203.
127. Clark PL, Weston BF, Gierasch LM. Fold Des 1998;3:401. [PubMed: 9806942]
128. Shao X, Matthews CR. Biochemistry 1998;37:7850. [PubMed: 9601046]
129. Khorasanizadeh S, Peters ID, Butt TR, Roder H. Biochemistry 1993;32:7054. [PubMed: 8392867]

130. Eftink MR, Shastry MCR. Methods Enzymol 1997;278:258. [PubMed: 9170317]
131. Stryer L. Science 1968;162:526. [PubMed: 5706935]
132. Semisotnov GV, Rodionova NA, Kutyshenko VP, Ebert B, Blanck J, Ptitsyn OB. FEBS Lett 1987;224:9. [PubMed: 2824244]
133. Engelhard M, Evans PA. Protein Sci 1995;4:1553. [PubMed: 8520481]
134. Shi L, Palleros DR, Fink AL. Biochemistry 1994;33:7536. [PubMed: 8011619]
135. Tsong TY. J Biol Chem 1974;249:1988. [PubMed: 4361834]
136. Colón W, Elöve GA, Wakem LP, Sherman F, Roder H. Biochemistry 1996;35:5538. [PubMed: 8611545]
137. Colón W, Roder H. Nat Struct Biol 1996;3:1019. [PubMed: 8946855]
138. Roder H, Maki K, Cheng H, Shastry MC. Methods 2004;34:15. [PubMed: 15283912]
139. Woody RW. Methods Enzymol 1995;246:34. [PubMed: 7538625]
140. Greenfield NJ. Methods Enzymol 2004;383:282. [PubMed: 15063655]
141. Kuwajima, K. Circular Dichroism and the Conformational Analysis of Biomolecules. Fasman, GD., editor. Plenum Press; New York: 1996.
142. Woody RW. Methods Enzymol 2004;380:242. [PubMed: 15051341]
143. Kuwajima K, Hiraoka Y, Ikeguchi M, Sugai S. Biochemistry 1985;24:874. [PubMed: 3994996]
144. Labhardt, AM. Unfolding and Refolding of Proteins. 131. Academic Press, Inc.; 1986.
145. Kuwajima K, Yamaya H, Miwa S, Sugai S, Nagamura T. FEBS Lett 1987;221:115. [PubMed: 3040467]
146. Gilmanshin RI, Ptitsyn OB. FEBS Lett 1987;223:327. [PubMed: 3666154]
147. Kuwajima K, Garvey EP, Finn BE, Matthews CR, Sugai S. Biochemistry 1991;30:7693. [PubMed: 1868049]
148. Elöve GA, Chaffotte AF, Roder H, Goldberg ME. Biochemistry 1992;31:6876. [PubMed: 1322172]
149. Hamada D, Segawa SI, Goto Y. Nat Struct Biol 1996;3:868. [PubMed: 8836104]
150. Kuwajima K, Yamaya H, Sugai S. J Mol Biol 1996;264:806. [PubMed: 8980687]
151. Surewicz WK, Mantsch HH, Chapman D. Biochemistry 1993;32:389. [PubMed: 8422346]
152. Walsh ST, Cheng RP, Wright WW, Alonso DO, Daggett V, Vanderkooi JM, DeGrado WF. Protein Sci 2003;12:520. [PubMed: 12592022]
153. White AJ, Drabble K, Wharton CW. Biochem J 1995;306(Pt 3):843. [PubMed: 7702582]
154. Troullier A, Reinstadler D, Dupont Y, Naumann D, Forge V. Nat Struct Biol 2000;7:78. [PubMed: 10625432]
155. Fabian H, Naumann D. Methods 2004;34:28. [PubMed: 15283913]
156. Kimura T, Takahashi S, Akiyama S, Uzawa T, Ishimori K, Morishima I. J Am Chem Soc 2002;124:11596. [PubMed: 12296715]
157. Lattman EE. Curr Opin Struct Biol 1994;4:87.
158. Doniach S, Bascle J, Garel T, Orland H. J Mol Biol 1995;254:960. [PubMed: 7500363]
159. Chen L, Wildegger G, Kiefhaber T, Hodgson KO, Doniach S. J Mol Biol 1998;276:225. [PubMed: 9514723]
160. Kimura T, Akiyama S, Uzawa T, Ishimori K, Morishima I, Fujisawa T, Takahashi S. J Mol Biol 2005;350:349. [PubMed: 15935376]
161. Kimura T, Uzawa T, Ishimori K, Morishima I, Takahashi S, Konno T, Akiyama S, Fujisawa T. Proc Natl Acad Sci U S A 2005;102:2748. [PubMed: 15710881]
162. DeWeerd K, Grigoryants V, Sun Y, Fetrow JS, Scholes CP. Biochemistry 2001;40:15846. [PubMed: 11747462]
163. Elöve GA, Bhuyan AK, Roder H. Biochemistry 1994;33:6925. [PubMed: 8204626]
164. Sosnick TR, Mayne L, Hiller R, Englander SW. Nat Struct Biol 1994;1:149. [PubMed: 7656032]
165. Colón W, Wakem LP, Sherman F, Roder H. Biochemistry 1997;36:12535. [PubMed: 9376358]
166. Roder H. Methods Enzymol 1989;176:446. [PubMed: 2811697]
167. Gladwin ST, Evans PA. Fold Des 1996;1:407. [PubMed: 9080187]

168. Sauder JM, Roder H. Fold Des 1998;3:293. [PubMed: 9710575]
169. Kern G, Handel T, Marqusee S. Protein Sci 1998;7:2164. [PubMed: 9792104]
170. Walkenhorst WF, Edwards JA, Markley JL, Roder H. Protein Sci 2002;11:82. [PubMed: 11742125]
171. Nishimura C, Dyson HJ, Wright PE. J Mol Biol 2002;322:483. [PubMed: 12225742]
172. Hoang L, Bedard S, Krishna MM, Lin Y, Englander SW. Proc Natl Acad Sci U S A 2002;99:12173. [PubMed: 12196629]
173. Krishna MM, Lin Y, Mayne L, Englander SW. J Mol Biol 2003;334:501. [PubMed: 14623190]
174. Nishimura C, Dyson HJ, Wright PE. Proc Natl Acad Sci U S A 2005;102:4765. [PubMed: 15769860]
175. Miranker A, Robinson CV, Radford SE, Aplin RT, Dobson CM. Science 1993;262:896. [PubMed: 8235611]
176. Zhang Z, Smith DL. Protein Sci 1993;2:522. [PubMed: 8390883]
177. Englander SW, Mayne L. Ann Rev Biophys Biomol Struct 1992;21:243. [PubMed: 1525469]
178. Scholtz, JM.; Robertson, AD. Methods in Molecular Biology. Shirley, BA., editor. 40. Humana Press, Inc.; Totowa, NJ: 1995.
179. Miranker A, Robinson CV, Radford SE, Dobson CM. FASEB J 1996;10:93. [PubMed: 8566553]
180. Englander SW, Sosnick TR, Englander JJ, Mayne L. Curr Opin Struct Biol 1996;6:18. [PubMed: 8696968]
181. Krishna MM, Hoang L, Lin Y, Englander SW. Methods 2004;34:51. [PubMed: 15283915]
182. Lu J, Dahlquist FW. Biochemistry 1992;31:4749. [PubMed: 1591236]
183. Jennings PA, Wright PE. Science 1993;262:892. [PubMed: 8235610]
184. Jacobs MD, Fox RO. Proc Natl Acad Sci USA 1994;91:449. [PubMed: 8290547]
185. Kuszewski J, Clore GM, Gronenborn AM. Protein Sci 1994;3:1945. [PubMed: 7703841]
186. Bai Y, Milne JS, Englander SW. Proteins: Struct, Funct, Genet 1993;17:75. [PubMed: 8234246]
187. Houry WA, Sauder JM, Roder H, Scheraga HA. Proc Natl Acad Sci USA 1998;95:4299. [PubMed: 9539731]
188. Hvidt A, Nielsen SO. Adv Protein Sci 1966;21:287.
189. Roder H, Wagner G, Wüthrich K. Biochemistry 1985;24:7396. [PubMed: 2417625]
190. Elöve GA, Roder H. ACS Symp Ser 1991;470:50.
191. Udgaonkar JB, Baldwin RL. Proc Natl Acad Sci USA 1990;87:8197. [PubMed: 2236032]
192. Radford SE, Dobson CM, Evans PA. Nature 1992;358:302. [PubMed: 1641003]
193. Bai Y, Sosnick TR, Mayne L, Englander SW. Science 1995;269:192. [PubMed: 7618079]
194. Jones BE, Matthews CR. Protein Sci 1995;4:167. [PubMed: 7757007]
195. Khorasanizadeh S, Peters ID, Roder H. Nat Struct Biol 1996;3:193. [PubMed: 8564547]
196. Raschke TM, Marqusee S. Nat Struct Biol 1997;4:298. [PubMed: 9095198]
197. Jackson SE, Fersht AR. Biochemistry 1991;30:10428. [PubMed: 1931967]
198. Schindler T, Herrler M, Marahiel MA, Schmid FX. Nat Struct Biol 1995;2:663. [PubMed: 7552728]
199. Crisanti MM, Matthews CR. Biochemistry 1981;20:2700. [PubMed: 7016182]
200. Parker MJ, Spencer J, Clarke AR. J Mol Biol 1995;253:771. [PubMed: 7473751]
201. Houry WA, Rothwarf DM, Scheraga HA. Nat Struct Biol 1995;2:495. [PubMed: 7664113]
202. Brandts JF, Halvorsen HR, Brennan M. Biochemistry 1975;14:4953. [PubMed: 241393]
203. Kiefhaber T, Kohler H, Schmid FX. J Mol Biol 1992;224:217. [PubMed: 1548700]
204. Sauder JM, MacKenzie NE, Roder H. Biochemistry 1996;35:16852. [PubMed: 8988024]
205. Gianni S, Travaglini-Allocatelli C, Cutruzzola F, Bigotti MG, Brunori M. J Mol Biol 2001;309:1177. [PubMed: 11399087]
206. Walkenhorst WF, Green SM, Roder H. Biochemistry 1997;63:5795. [PubMed: 9153420]
207. Oliveberg M, Tan YJ, Silow M, Fersht AR. J Mol Biol 1998;277:933. [PubMed: 9545382]
208. Park SH, O'Neil KT, Roder H. Biochemistry 1997;36:14277. [PubMed: 9400366]
209. Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. Science 1991;253:657. [PubMed: 1871600]
210. Alexander P, Orban J, Bryan P. Biochemistry 1992;31:7243. [PubMed: 1510916]

211. Benson, SW. McGraw-Hill. Advanced Topics in Chemistry. 1. McGraw-Hill; New York: 1960.
212. Pogliani L, Terenzi M. J Chem Educ 1992;69:278.
213. Krantz BA, Mayne L, Rumbley J, Englander SW, Sosnick TR. J Mol Biol 2002;324:359. [PubMed: 12441113]
214. Krantz BA, Sosnick TR. Biochemistry 2000;39:11696. [PubMed: 10995237]
215. Laub PB, Khorasanizadeh S, Roder H. Protein Sci 1995;4:973. [PubMed: 7663353]
216. Went HM, Benitez-Cardoza CG, Jackson SE. 2004submitted
217. Brems DN, Stellwagen E. J Biol Chem 1983;258:3655. [PubMed: 6300051]
218. Muthukrishnan K, Nall BT. Biochemistry 1991;30:4706. [PubMed: 1851434]
219. Hammack B, Godbole S, Bowler BE. J Mol Biol 1998;275:719. [PubMed: 9480763]
220. Hagen SJ, Hofrichter J, Szabo A, Eaton WA. Proc Natl Acad Sci USA 1996;93:11615. [PubMed: 8876184]
221. Hagen SJ, Eaton WA. J Mol Biol 2000;297:781. [PubMed: 10731428]
222. Qiu L, Zachariah C, Hagen SJ. Phys Rev Lett 2003;90:168103. [PubMed: 12732017]
223. Barrick D, Baldwin RL. Biochemistry 1993;32:3790. [PubMed: 8466917]
224. Ballew RM, Sabelko J, Gruebele M. Proc Natl Acad Sci USA 1996;93:5759. [PubMed: 8650166]
225. Ballew RM, Sabelko J, Gruebele M. Nat Struct Biol 1996;3:923. [PubMed: 8901868]
226. Eliezer D, Jennings PA, Wright PE, Doniach S, Hodgson KO, Tsuruta H. Science 1995;270:487. [PubMed: 7570004]
227. Gilmanshin R, Williams S, Callender RH, Woodruff WH, Dyer RB. Proc Natl Acad Sci USA 1997;94:3709. [PubMed: 9108042]
228. Gulotta M, Gilmanshin R, Buscher TC, Callender RH, Dyer RB. Biochemistry 2001;40:5137. [PubMed: 11318635]
229. Jamin M, Yeh SR, Rousseau DL, Baldwin RL. J Mol Biol 1999;292:731. [PubMed: 10497035]
230. Tsui V, Garcia C, Cavagnero S, Siuzdak G, Dyson HJ, Wright PE. Protein Sci 1999;8:45. [PubMed: 10210182]
231. Cocco MJ, Lecompte JT. Protein Sci 1994;3:267. [PubMed: 8003963]
232. Eliezer D, Wright PE. J Mol Biol 1996;263:531. [PubMed: 8918936]
233. Roder H. Proc Natl Acad Sci U S A 2004;101:1793. [PubMed: 14769941]
234. Yao J, Chung J, Eliezer D, Wright PE, Dyson HJ. Biochemistry 2001;40:3561. [PubMed: 11297422]
235. Nishimura C, Wright PE, Dyson HJ. J Mol Biol 2003;334:293. [PubMed: 14607120]
236. Nymeyer H, Garcia AE, Onuchic JN. Proc Natl Acad Sci U S A 1998;95:5921. [PubMed: 9600893]
237. Sabelko J, Ervin J, Gruebele M. Proc Natl Acad Sci USA 1999;96:6031. [PubMed: 10339536]
238. Ma H, Gruebele M. Proc Natl Acad Sci U S A 2005;102:2283. [PubMed: 15699334]
239. de Gennes PG. J Phys (Paris) Lett 1985;46:L639.
240. Hagen SJ. Proteins: Struct, Funct, Genet 2003;50:1. [PubMed: 12471594]
241. Shortle D. Adv Protein Chem 2002;62:1. [PubMed: 12418099]
242. Tanford C, Kawahara K, Lapanje S. J Amer Chem Soc 1967;89:729.
243. Agashe VR, Shastry MCR, Udgaonkar JB. Nature 1995;377:754. [PubMed: 7477269]
244. Sosnick TR, Mayne L, Englander SW. Proteins: Struct, Funct, Genet 1996;24:413. [PubMed: 9162942]
245. Sridevi K, Udgaonkar JB. Biochemistry 2003;42:1551. [PubMed: 12578368]
246. Sosnick TR, Shtilerman MD, Mayne L, Englander SW. Proc Natl Acad Sci USA 1997;94:8545. [PubMed: 9238013]
247. Qi PX, Sosnick TR, Englander SW. Nat Struct Biol 1998;5:882. [PubMed: 9783747]
248. Choe SE, Matsudaira PT, Osterhout J, Wagner G, Shakhnovich EI. Biochemistry 1998;37:14508. [PubMed: 9772179]
249. Ferguson N, Capaldi AP, James R, Kleanthous C, Radford SE. J Mol Biol 1999;286:1597. [PubMed: 10064717]
250. Gorski SA, Capaldi AP, Kleanthous C, Radford SE. J Mol Biol 2001;312:849. [PubMed: 11575937]

251. Ivankov DN, Garbuzynskiy SO, Alm E, Plaxco KW, Baker D, Finkelstein AV. Protein Sci 2003;12:2057. [PubMed: 12931003]
252. Burton RE, Huang GS, Daugherty MA, Fullbright PW, Oas TG. J Mol Biol 1996;263:311. [PubMed: 8913309]
253. Wittung-Stafshede P, Lee JC, Winkler JR, Gray HB. Proc Natl Acad Sci U S A 1999;96:6587. [PubMed: 10359755]
254. Jacob M, Holtermann G, Perl D, Reinstein J, Schindler T, Geeves MA, Schmid FX. Biochemistry 1999;38:2882. [PubMed: 10074340]
255. Creighton TE. Nat Struct Biol 1994;1:135. [PubMed: 7656027]
256. Fersht AR. Proc Natl Acad Sci USA 1995;92:10869. [PubMed: 7479900]
257. Roder, H.; Elöve, GA. Mechanisms of Protein Folding: Frontiers in Molecular Biology. Pain, RH., editor. Oxford University Press; New York: 1994.
258. Silow M, Oliveberg M. Proc Natl Acad Sci USA 1997;94:6084. [PubMed: 9177173]
259. Matouschek, A.; Serrano, L.; Fersht, AR. Mechanisms of protein folding: Frontiers in molecular biology. Pain, RH., editor. Oxford University Press; Oxford: 1994.
260. Raschke TM, Kho J, Marqusee S. Nat Struct Biol 1999;6:825. [PubMed: 10467093]
261. Heidary DK, Gross LA, Roy M, Jennings PA. Nat Struct Biol 1997;4:725. [PubMed: 9303000]
262. Jennings P, Roy M, Heidary D, Gross L. Nat Struct Biol 1998;5:11. [PubMed: 9437419]
263. Capaldi AP, Kleanthous C, Radford SE. Nat Struct Biol 2002;9:209. [PubMed: 11875516]
264. Plaxco KW, Simons KT, Baker D. J Mol Biol 1998;277:985. [PubMed: 9545386]
265. Myers JK, Oas TG. Nat Struct Biol 2001;8:552. [PubMed: 11373626]
266. Bushnell GW, Louie GV, Brayer GD. J Mol Biol 1990;214:585. [PubMed: 2166170]

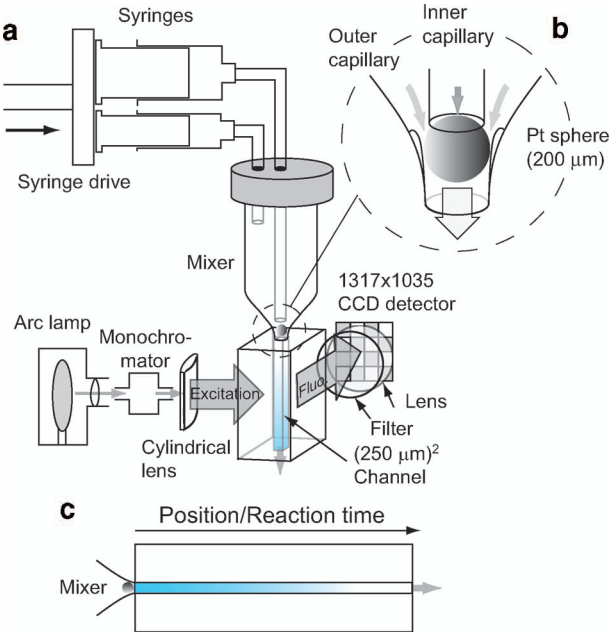


Figure 1.

Continuous-flow capillary mixing apparatus in fluorescence mode. ***a*** Schematic of the solution delivery system, mixer, observation cell and optical arrangement. ***b*** Expanded view of the mixer. ***c*** Diagram illustrating continuous-flow measurement.

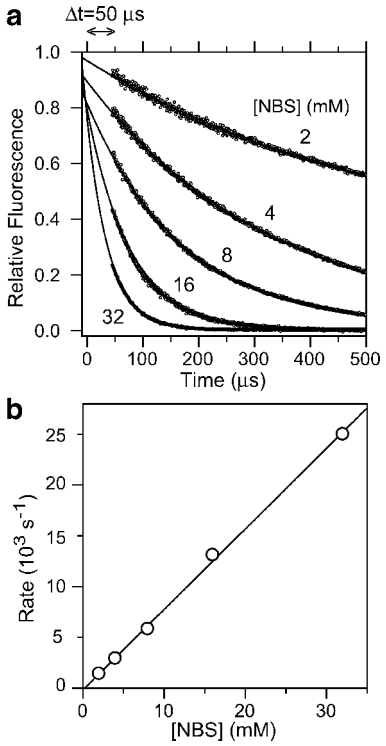


Figure 2.

Continuous-flow measurements of the quenching of NATA fluorescence by NBS used to determine the experimental dead-time. ***a*** Plot of NATA fluorescence (>324 nm) *vs*. time at several NBS concentrations. ***b*** NATA-NBS reaction rates from exponential fitting of the data in panel *a vs*. NBS concentration. Linear regression (line) yields a second-order rate constant of 7.9 10^5 M^{-1} s^{-1}.

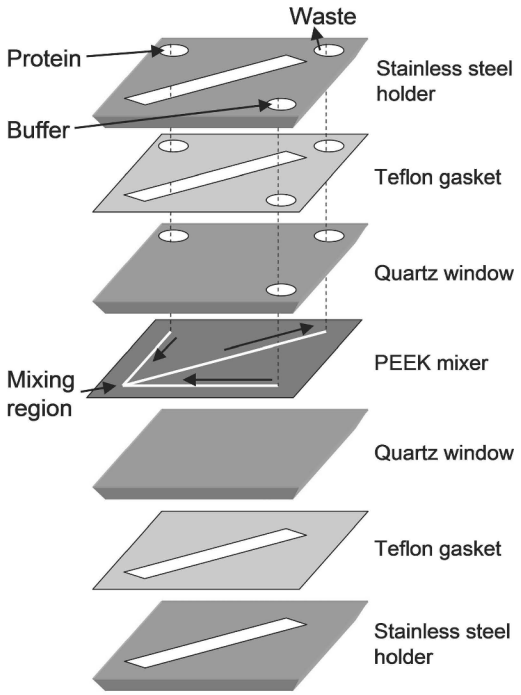


Figure 3.

Schematic of a microfluidic mixer designed by Bilsel et al.[91]. The 127-μm-thick mixer is sandwiched between quartz windows and is sealed in a stainless-steel holder using Teflon gaskets (1.6 mm thick). Solutions are delivered to and from the mixing region through holes in the upper layers (arrows). Reprinted with permission from ref [91]. Copyright 2005 American Institute of Physics.

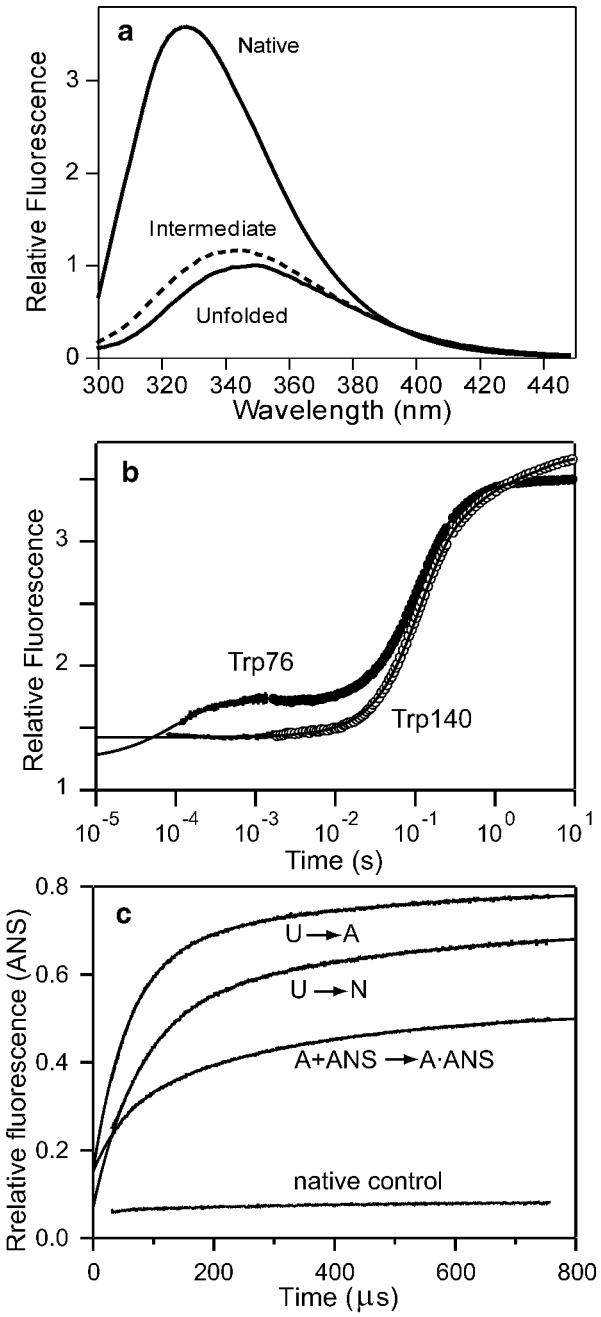


Figure 4.

Folding mechanism of SNase probed by tryptophan fluorescence. ***a*** Fluorescence emission spectra of the Trp76 variant of SNase under native and denaturing conditions (solid) and a folding intermediate populated at equilibrium (dashed). The spectrum of the intermediate was determined by global analysis of the fluorescence spectra as a function of urea concentration (pH 5.2, 15 °C). ***b*** Time-course of folding (triggered by a pH jump from 2 to 5.2) for wild-type SNase (Trp140) and a single-tryptophan variant (Trp76) measured by continuous-flow (< 10^{-3} s) and stopped-flow (> 10^{3} s) fluorescence. ***c*** ANS fluorescence changes during ANS binding/folding of Trp76 SNase measured by continuous-flow experiments at 15°C in the presence of 160 μM ANS. U → A: Salt concentration jump from 0 M to 1 M KCl at pH2.0; U → N: refolding induced by a pH-jump from 2.0 to 5.2; A+ANS → A•ANS: ANS binding

kinetics in the presence of 1 M KCl at pH 2.0; native control: ANS binding kinetics under the native condition (pH 5.2). Adapted from ref [87] (Figures 3 and 5).

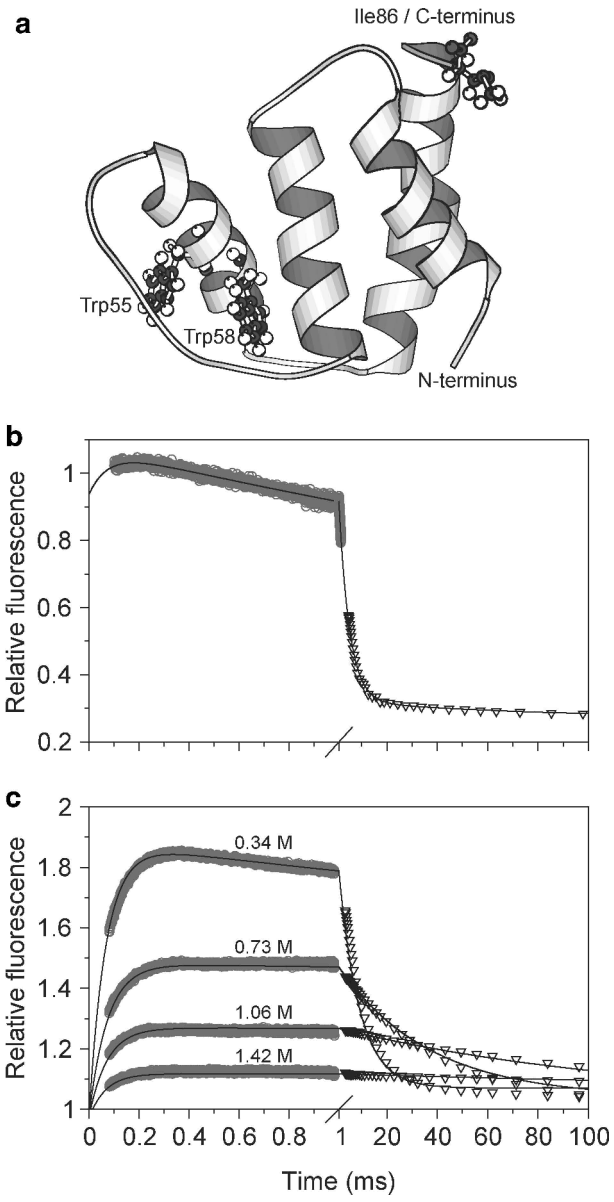


Figure 5.

FRET-detection of an early folding intermediate in a helix-bundle protein, ACBP. ***a*** Ribbon diagram of ACBP, based on an NMR structure. The two tryptophan residues and the mutated C-terminal isoleucine are shown in ball and stick. The two lower panels show refolding kinetics of unmodified ACBP (***b***) and AEDANS-labeled ACBP,I86C (***c***) in pH 5.3 buffer containing 0.34 M GuHCl at 26 °C. In both panels data from continuous-flow (○) and stopped-flow (▽) experiments were matched and combined. Reprinted with permission from ref [83]. Copyright 2002 National Academy of Sciences of the USA.

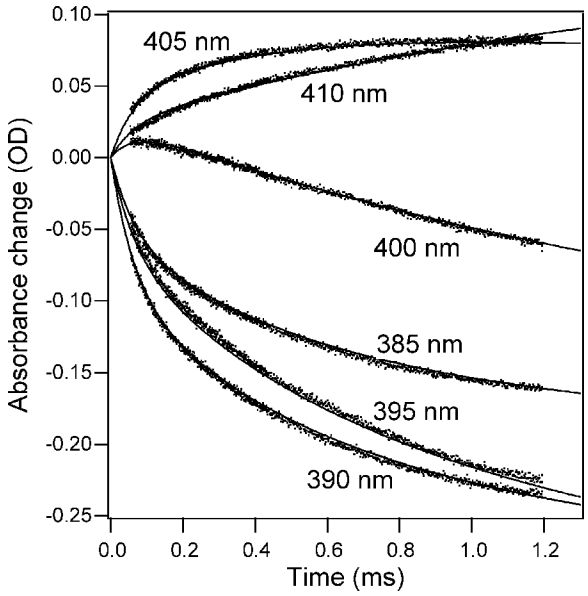


Figure 6.

Initial stages of refolding of acid-denatured oxidized cyt *c* at pH 5 monitored by continuous-flow absorbance measurements at different wavelengths spanning the Soret heme absorbance band. The lines represent a global fit of a four-state folding mechanism to the family of kinetic traces. Reprinted with permission from ref [138]. Copyright 2004 Elsevier Inc.

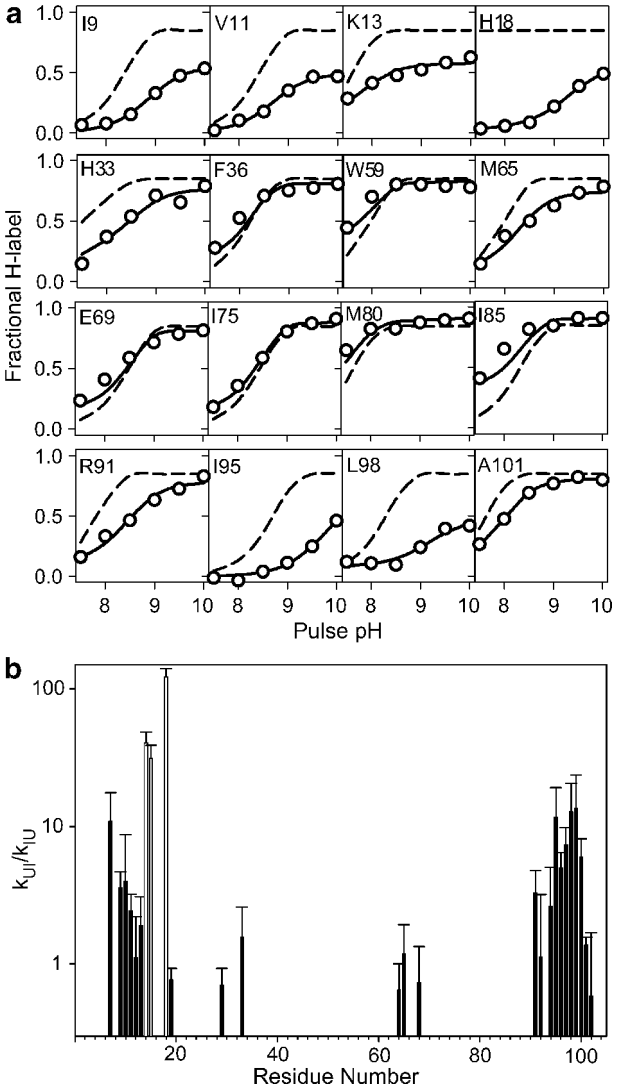


Figure 7.

Characterization of an intermediate populated during folding of cytochrome *c* (pH 6, 10 °C).[173] ***a*** Fractional degree of labeling *vs.* pulse pH for a representative set of NH groups measured by 2D NMR analysis of refolding cytochrome *c* samples that were exposed to a 50 ms labeling pulse of increasing pH at a folding time of 100 ms. Solid lines represent a fit of the data yielding the rates of formation and unfolding of the intermediate state (see text). Dashed lines indicate the labeling profiles expected in the absence of structure. ***b*** Equilibrium constant for formation of the intermediate, $K_{UI} = k_{UI}/k_{IU}$, based on the fits of the labeling results in ***a***. Values of $K_{UI} \geq 1$ indicative of persistent hydrogen bonded structure are mainly found for residues in the N- and C-terminal helices. Cys 14, Ala15 and His 18 (gray bars) are protected even in the unfolded state. Adapted from ref [173] (Figures 3 and 4), with permission by the authors.

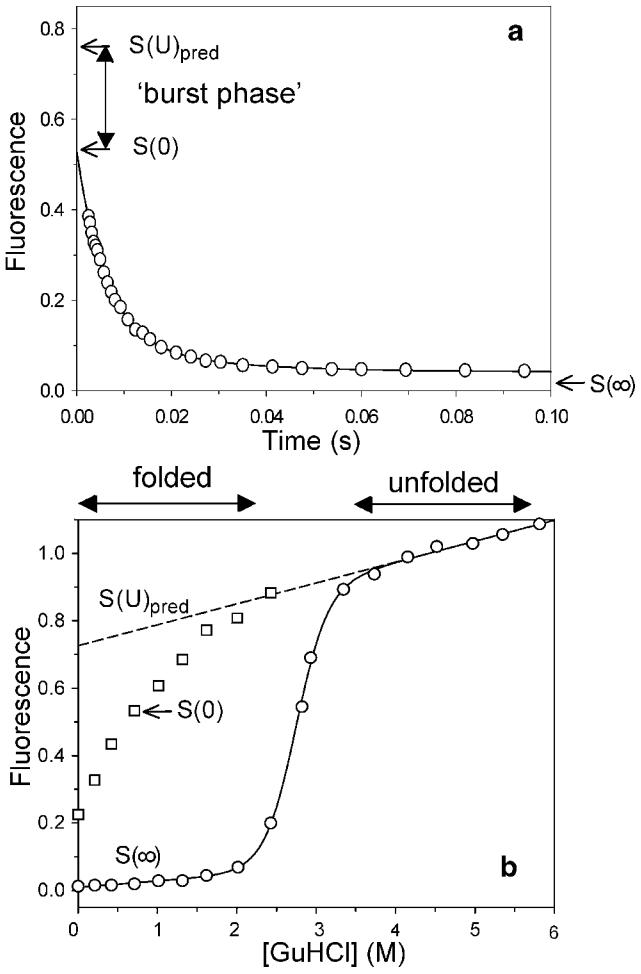


Figure 8.

Stopped-flow fluorescence evidence for an unresolved rapid process (burst phase) during folding of cyt *c* (pH 5, 10 °C). ***a*** Tryptophan fluorescence changes during refolding of acid-unfolded cytochrome c (pH 2, ~15 mM HCl) at a final GuHCl concentration of 0.7 M. The initial signal S(0) at t=0 (determined on the basis of a separate dead-time measurement) falls short of the signal for the unfolded state under refolding conditions, $S_{pred}(U)$, obtained by linear extrapolation of the unfolded-state baseline (see dashed line in *b*). ***b*** Effect of the denaturant concentration on the initial (squares) and final (circles) fluorescence signal, S(0) and S(∞), measured in a series of stopped-flow refolding experiments at different final GuHCl concentration. Reprinted with permission from ref [138]. Copyright 2004 Elsevier Inc.

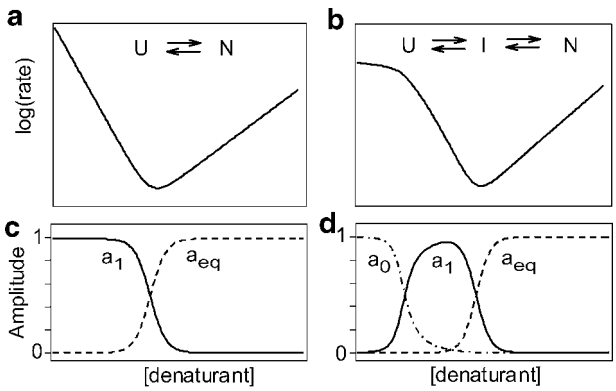


Figure 9.

Schematic log(rate) *vs*. [denaturant] plots (chevrons) for a two-state (***a***) and a three-state (***b***) folding/unfolding mechanism. The lower panels show the predicted amplitude for the main folding phases (a_1), the burst phase predicted for a three-state process (a_0 in panel ***d***) and the equilibrium unfolding transition (a_{eq}).

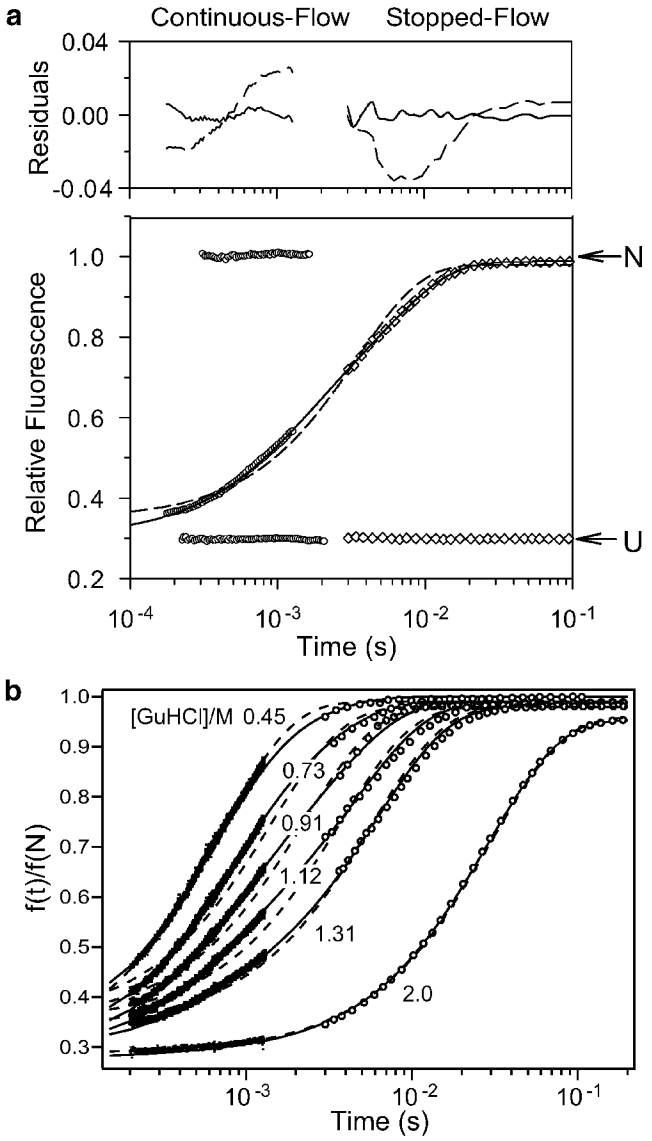


Figure 10.

Folding kinetics of GB1 at pH 5.0, 20 °C, in the presence of 0.4 M sodium sulfate. Panel ***a*** shows a representative kinetic trace at 1.12 M GHCl monitored by continuous-flow (circles) and stopped-flow (diamonds) fluorescence along with controls for fully unfolded (U) and folded (N) solutions. Single- and double-exponential fits and residuals are shown with solid and dashed lines, respectively. Reprinted with permission from ref [78]. Copyright 1999 Nature Publishing Group. Panel ***b*** shows a family of refolding traces at the final GuHCl concentrations indicated. Solid lines show the time course predicted on the basis of a three-state model (Scheme 3), using the following parameters: $k_{UI}° = 2300\ s^{-1}$, $m_{UI}^{‡} = -0.6$ kcal $mol^{-1}M^{-1}$, $k_{IU}° = 70\ s^{-1}$, $m_{IU}^{‡} = 1.15$ kcal $mol^{-1}M^{-1}$, $k_{IN}° = 600\ s^{-1}$, $m_{IN}^{‡} = 0$, $k_{NI}° = 0.14$, $m_{NI}^{‡} = 0.3$ kcal $mol^{-1}M^{-1}$ (see Figure 11*a*, ref [78]). The relative signals for the N-, I- and U-states were $s_N = 0.29 \pm 0.01$, $s_I = s_U = 0.98 \pm 0.02$. Dashed lines indicate a best "fit" of a two-state mechanism, using rate constants falling on a linear extrapolation of the chevron plot between 1.5 and 3 M GuHCl to lower denaturant concentrations (cf. ref [213]).

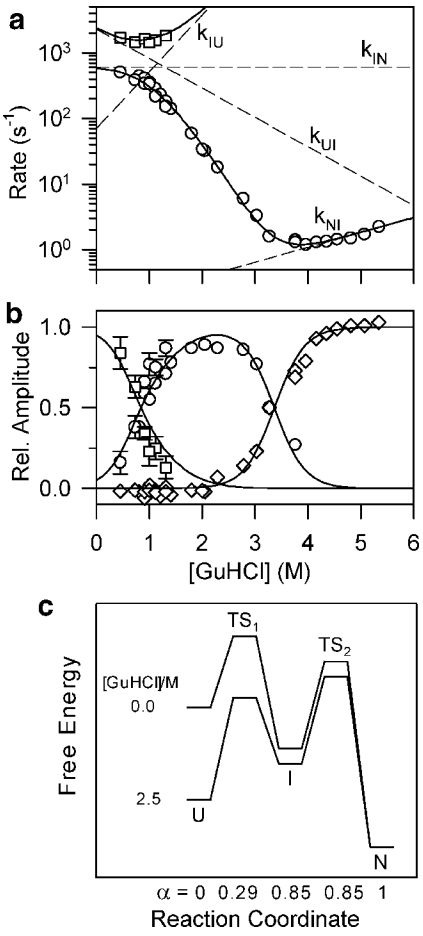


Figure 11.

GuHCl-dependence of the rate constants (***a***) and kinetic amplitudes (***b***) of the fast (squares) and slow (circles) kinetic phases observed during folding of GB1. ***c*** Free energy diagrams for folding of GB1 under conditions where the intermediate, I is well populated (0 M) and unstable (2.5 M GuHCl). α represents the change in solvent-accessible surface area relative to the unfolded state U. Reprinted with permission from ref [78]. Copyright 1999 Nature Publishing Group.

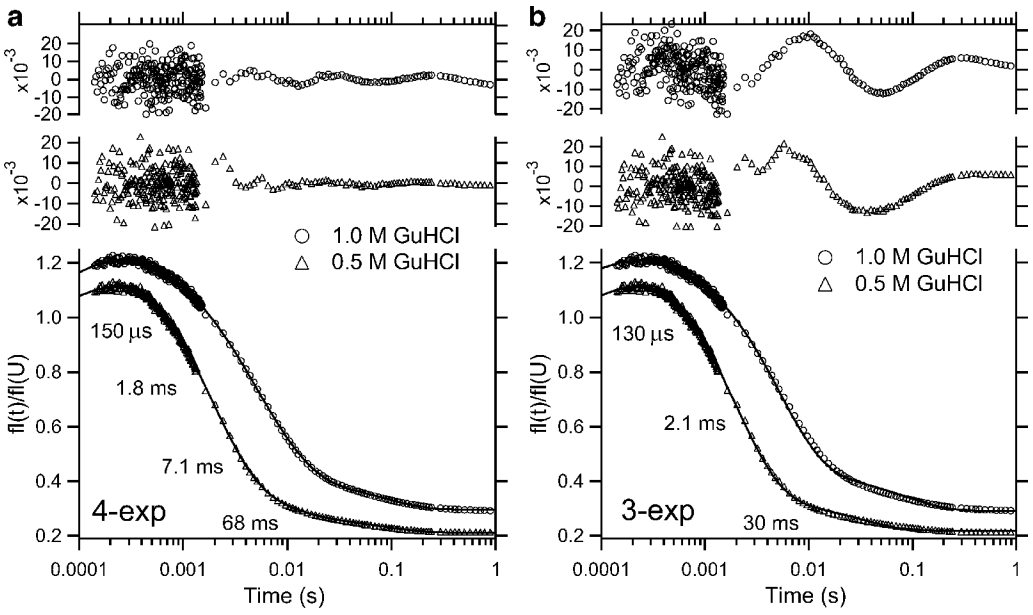


Figure 12.

Comparison of quadruple (***a***) and triple (***b***) exponential fitting of the kinetics of refolding of F45W ubiquitin at final GuHCl concentrations of 0.5 and 1.0 M (pH 5, 25 °C). Fluorescence traces measured in continuous- and stopped-flow experiments were normalized with respect to the unfolded protein in 6 M GuHCl. The residuals (top two traces in each panel) indicate that four exponentials are required to obtain a satisfactory fit of the data over the time window shown. Reprinted with permission from ref [14]. Copyright 2005 Wiley-VCH.

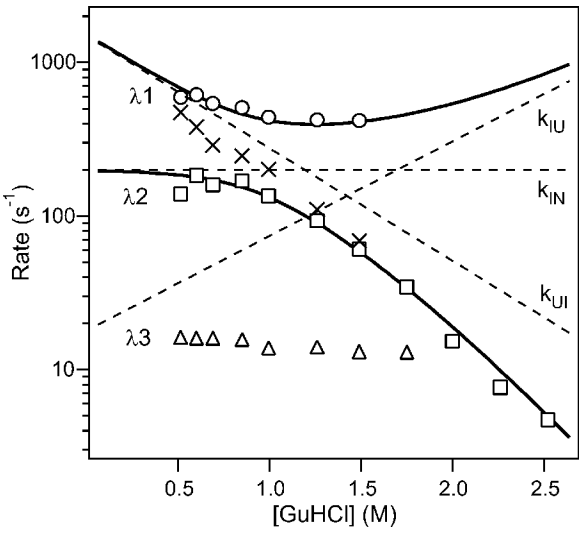


Figure 13.

Expanded region of the rate profile (log rate vs. [GuHCl[) for the two main phases (circles and squares) and a minor slower phase (dtriangles) observed during folding of F45W ubiquitin (pH 5, 25 °C). The solid lines show the rates predicted by a three-state model, and the dashed lines indicate the elementary rate constants used. The X-symbols show the apparent rates obtained by triple-exponential fitting (Figure 10*b*). Reprinted with permission from ref [14]. Copyright 2005 Wiley-VCH.

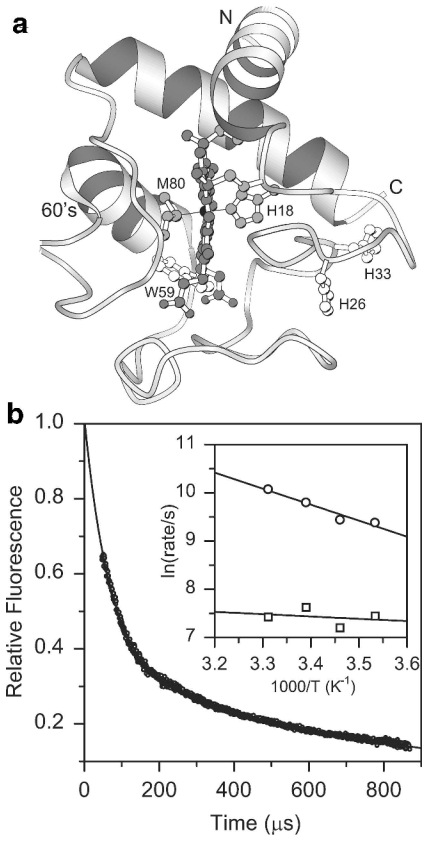


Figure 14.

a Ribbon diagram of horse cytochrome *c*, based on the crystal structure[266]. ***b*** Refolding kinetics of acid-unfolded cyt *c* at pH 4.5, 22 °C measured by continuous-flow mixing, indicating heme-induced quenching of Trp59 fluorescence associated with chain collapse.[76] Inset: Arrhenius plot for the rates of the major (circles) and the minor (squares) submillisecond phases. Adapted from ref [76] (Figure 2).

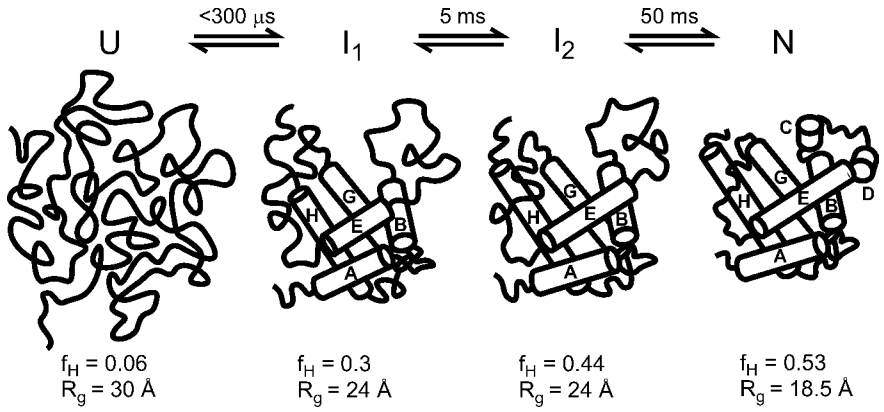


Figure 15.

Sequential four-state mechanism involving two partially structured ensembles, I_1 and I_2, in addition to the acid-unfolded (U) and native (N) states. Cartoons of various states apomyoglobin consistent with the fraction of helical residues (f_H) and radius of gyration (R_g) reported in ref [89]. Cylinders indicate the approximate position of α-helices in the native state and possible arrangement of core helices in the intermediates consistent with amide protection data.[171,174]

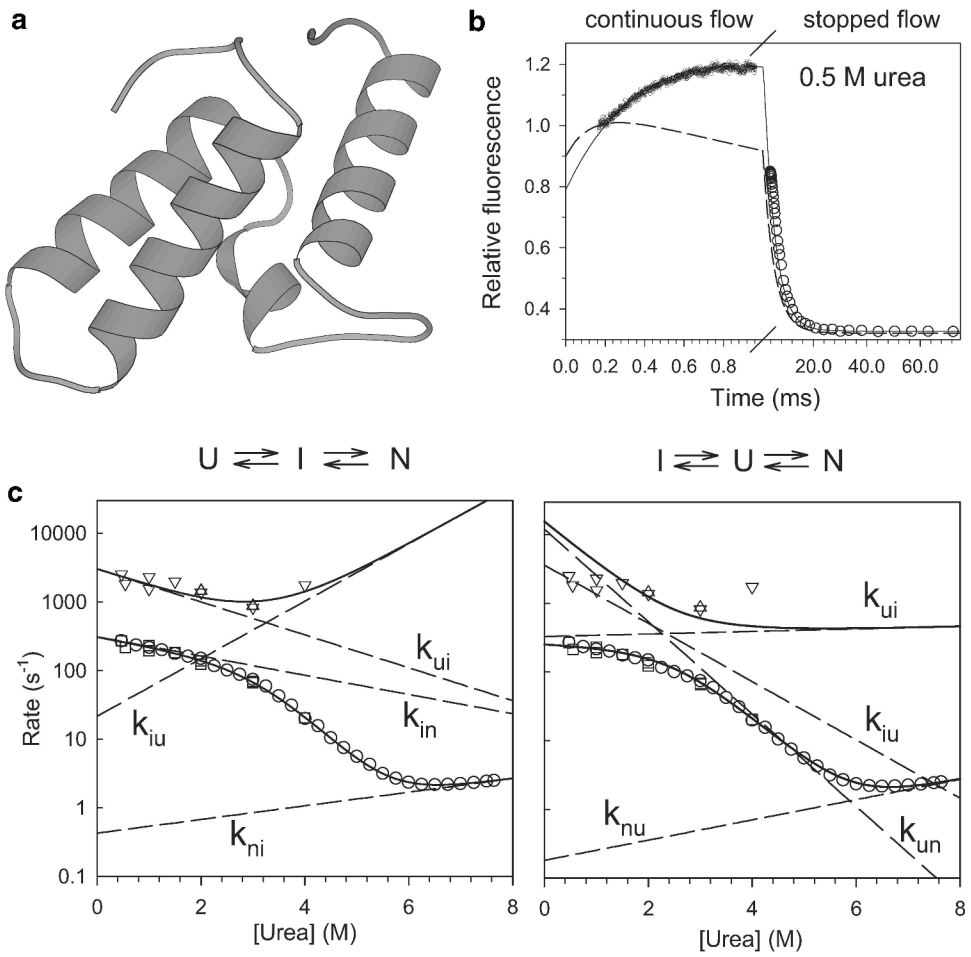
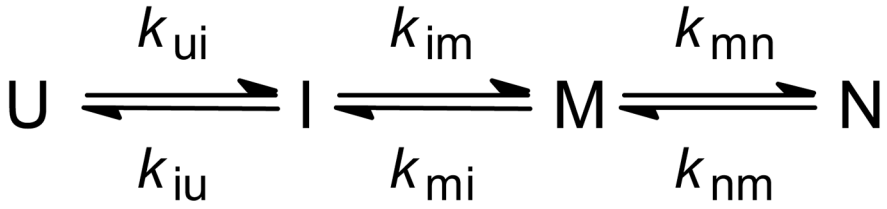
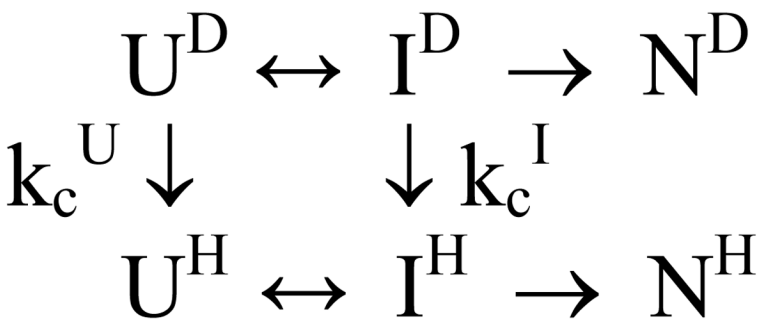


Figure 16.

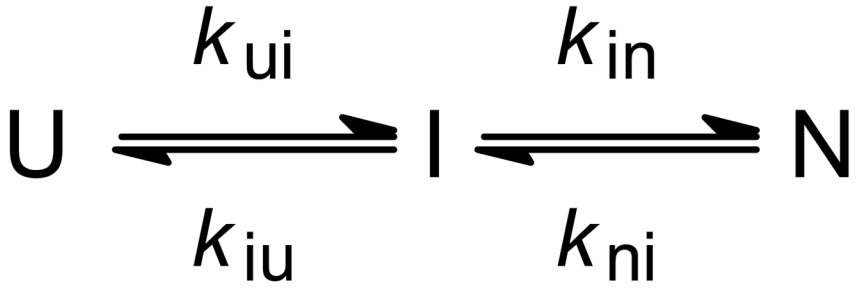
Kinetic mechanism of Im7 folding. ***a*** Ribbon diagram of Im7. ***b*** Representative kinetic trace measured by continuous-flow (●) and stopped-flow (○) fluorescence. The kinetics at this and all other urea concentrations measured is accurately predicted by on on-pathway mechanism (solid line) while schemes with off-pathway intermediates fail to reproduce the data (dashed line). ***c*** Observed (symbols) and predicted (solid lines) rates of folding and unfolding, based on mechanisms with on-pathway (left) and off-pathway (right) intermediates. Dashed lines indicate the corresponding elementary rate constants. Adapted from ref [81] (Figures 2 and 3).



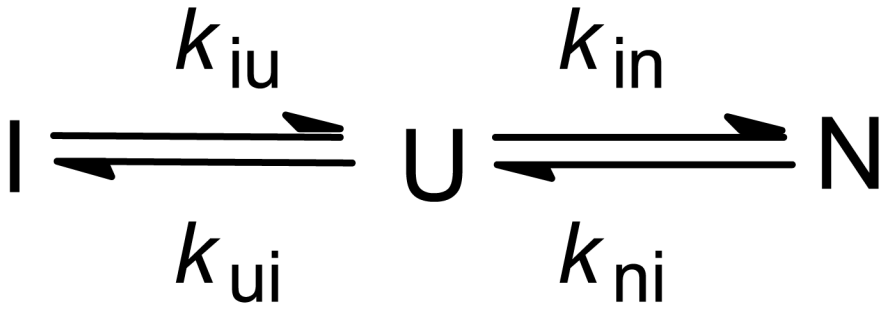
Scheme 1.



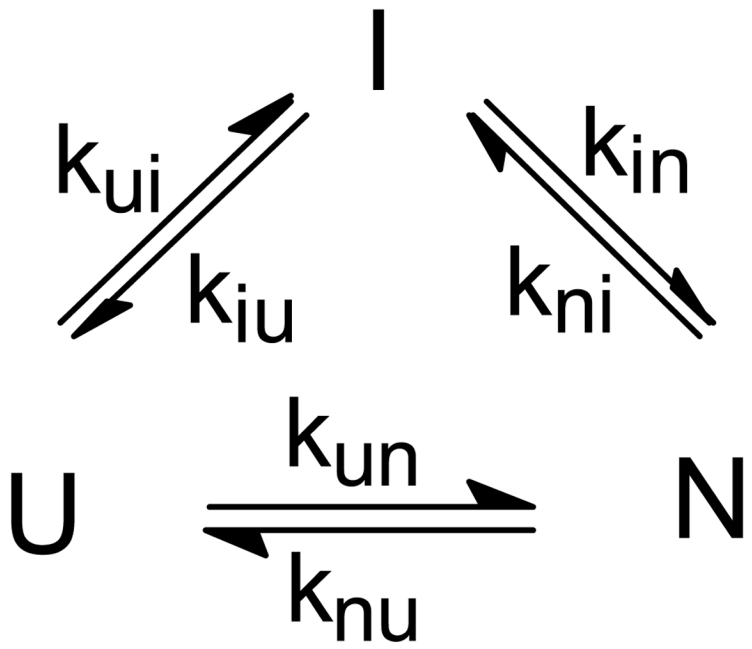
Scheme 2.



Scheme 3.



Scheme 4.



Scheme 5.

Table 1

Common detection methods used in rapid mixing studies of protein folding

Method	Probe	Properties Probed	Sensitivity
Fluorescence	Trp, Tyr	solvent shielding tertiary contacts (quenching)	+++
	ANS	Hydrophobic clusters, collapse	+++
	FRET	donor-acceptor distance	++
Absorbance	Trp, Tyr, cofactor	polarity, solvent perturbation	++
far-UV CD	peptide bond	secondary structure	−−
near-UV CD	Tyr, Trp, co-factor	side-chain packing, mobility	−−
IR, Raman	peptide bond cofactor	secondary structure metal coordination	+
SAXS	heavy atoms	size (Rg), shape	−
NMR, H/D exchange	labile hydrogens	H-bonding, solvent accessibility	−
EPR	unpaired electrons	environment of free radicals, paramagnetic metals	−

Table 2

Comparison of the time constant of the initial folding phase with protein size and secondary structure type for

Protein	No. of Residues	Structure type	Relative Contact Order[1]	Time Constant (μs)[2]	Reference
GB1	57	α/β	17.2%	600	78
ACBP	86	α	14.0%	80	83
Im7	87	α	10.6%	450	81
cyt c	104	α	11.1%	60	76
SNase	149	α/β	10.0%	75[3]	87

[1]The relative contact order was calculated using a web-based program at http://depts.washington.edu/bakerpg/contact_order/ (ref 264).

[2]Time constant of earliest kinetic phase detected in continuous-flow fluorescence measurements.

[3]Value for the F76W/W140H variant of SNase.