REVIEW

NMR and protein folding: Equilibrium and stopped-flow studies

CARL FRIEDEN, SYDNEY D. HOELTZLI, AND IRA J. ROPSON¹

Department of Biochemistry and Molecular Biophysics, Washington University
School of Medicine, St. Louis, Missouri 63110

(RECEIVED July 27, 1993; ACCEPTED August 18, 1993)

Abstract

NMR studies are now unraveling the structure of intermediates of protein folding using hydrogen–deuterium exchange methodologies. These studies provide information about the time dependence of formation of secondary structure. They require the ability to assign specific resonances in the NMR spectra to specific amide protons of a protein followed by experiments involving competition between folding and exchange reactions. Another approach is to use ¹⁹F-substituted amino acids to follow changes in side-chain environment upon folding. Current techniques of molecular biology allow assignments of ¹⁹F resonances to specific amino acids by site-directed mutagenesis. It is possible to follow changes and to analyze results from ¹⁹F spectra in real time using a stopped-flow device incorporated into the NMR spectrometer.

Keywords: ¹⁹F NMR spectra; hydrogen–deuterium exchange; protein folding; protein intermediates; stopped-flow studies

The problem of the mechanism by which a protein spontaneously folds to its correct native structure, given only the information available in its amino acid sequence, has puzzled investigators ever since Anfinsen's classical experiment on ribonuclease more than 30 years ago (Anfinsen et al., 1961). Many reviews have been written on the subject (Kim & Baldwin, 1982, 1990; Alber, 1989; Chothia & Finkelstein, 1990; Creighton, 1990; Baldwin, 1991; Dill & Shortle, 1991; Richards, 1991; Scholtz & Baldwin, 1992; Jaenicke & Buchner, 1993; Rose & Wolfenden, 1993), and most point out that unraveling the mechanisms of protein folding presents some of the most challenging theoretical and experimental problems in biochemistry today. Numerous complex issues are involved and multiple experimental approaches have been used to attempt to answer questions relating to the nature and stability of intermediates in the folding process. Questions to be addressed experimentally include: What is the structure of the intermediates? How rapidly do they form and disappear? Are there incorrectly folded transient species? How

important is steric packing? What are the dominant driving forces (or do different forces become more important at different stages of folding)? Is there a sequential pathway or do multiple paths exist (or both, under different conditions)? How does one detect intermediates that may exist for short times at very low concentrations?

The answers to questions such as these are necessary to provide the basis for sound theoretical work. Yet they remain elusive and difficult to explore because many experimental techniques do not give the needed information. For example, for years investigators have measured the rate of the folding or unfolding process using changes in intrinsic fluorescence, absorbance, or more recently, circular dichroism. These rate measurements, restricted by current mixing devices to time regimes of milliseconds or longer, tell us that intermediates do exist, but very little more, because (1) they reflect global changes in structure rather than yielding information about specific regions of the protein, and (2) they do not clearly distinguish intermediates that may have properties similar to native or denatured states or other intermediates. The kinetically observed changes in these time regimes simply represent metastable states in which the observed property is different from the preceding or succeeding intermediate. Thus, while necessary for kinetic characterization, such

Reprint requests to: Carl Frieden, Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110.

¹ Present address: Department of Biological Chemistry, Penn State University College of Medicine, Hershey, Pennsylvania 17033.

measurements say little about folding pathways, rather being interpreted in terms of general concepts such as the formation of molten globule-like structures, secondary structure, tertiary interactions, or changes in exposure of hydrophobic surfaces. One of the most difficult experimental issues, therefore, is how to examine the nature of those intermediate forms that do occur on the folding pathway. This is especially true for intermediates that may form (and disappear) in very fast (less than millisecond) time regimes. For example, as measured by CD and fluorescent dye binding, much of the hydrophobic core and secondary structure of a domain of the β -chain of tryptophan synthase is formed within the dead time of the mixing device (4 ms) (Chaffotte et al., 1992a). Similar observations have been made with cytochrome *c* (Elove et al., 1992), dihydrofolate reductase (Kuwajima et al., 1991), staphylococcal nuclease (Sugawara et al., 1991), hen lysozyme (Chaffotte et al., 1992b), ribonuclease T1 (Kiefhaber et al., 1992), interleukin-1 β (Varley et al., 1993), and *Escherichia coli trp* aporepressor (Mann & Matthews, 1993). Ropson et al. (1990) have observed that there is a fluorescence phase in the refolding of intestinal fatty acid binding protein (IFABP) that occurs within the dead time of the mixing (<1 ms). These observations indicate that a significant fraction of the folding process occurs faster than can be measured by the usual mixing methods.

While the time regime problem remains to be solved and may involve other techniques such as temperature jump or laser photolysis of caged compounds, NMR techniques have recently given valuable information on intermediate structures in millisecond or longer time regimes.

NMR and hydrogen-exchange labeling

In 1979 Schmid and Baldwin described a method for giving information about structural intermediates in protein folding. The procedure involved labeling exchangeable amide protons of an unfolded protein with tritium, then refolding the protein and examining the number of protons protected from exchange at different times during refolding. The method was based on the assumptions that (1) exchange of amide protons with the solvent protons would be very slow in the folded protein, and (2) there would be competition between folding and tritium exchange. The experimental system used in this first application of the technique was ribonuclease A, and the investigators examined the slow-folding process, a convenient time regime (greater than seconds) for the experiment. From the data obtained, they could only conclude that the hydrogen-bonded backbone of the protein formed during an early stage of folding and that the folding process was sequential. Soon after, Kim and Baldwin (1980) described a pulse-labeling technique in which the entire time course of the folding process was monitored under constant conditions. These authors pointed out that,

for the technique to provide information about structure in intermediates, it would be necessary to assign proton resonances of the NMR spectrum of the protein (or fragments) to specific amino acid residues. Fortunately, at about the same time as these folding experiments were proposed, NMR techniques were being developed to provide just such assignments. In 1982, Wagner and Wüthrich made the first assignments for the small (58 residues) protein bovine pancreatic trypsin inhibitor (BPTI).

In 1988, two seminal papers appeared that related time-dependent refolding processes with structural changes measured by proton NMR (Roder et al., 1988; Udgaonkar & Baldwin, 1988). The proteins used were ribonuclease A (13.7 kDa) and oxidized cytochrome *c* (12.3 kDa). At the time, two-dimensional ^1H -NMR techniques allowed almost complete assignments of backbone amide resonances for both proteins. Combined with rapid mixing techniques and pulse-labeling methods, the authors could determine which protons were protected as a consequence of folding over time after refolding was initiated. This allowed, for the first time, a true characterization of specific structural regions during refolding. Experiments with oxidized cytochrome *c* were performed at 12 different refolding times, the shortest being 4 ms.

Even within this shortest time, all proton occupancies were reduced by about 20% using cytochrome *c*, indicating a very rapid folding process. The first hydrogen bonds that could be detected for this protein were in the N-terminal and C-terminal helices, which formed within 20 ms, with several longer phases present in parallel. For ribonuclease A, the results presented in this and later (Udgaonkar & Baldwin, 1990) experiments gave evidence of one principal early folding intermediate whose properties changed as a function of time. For this protein, the experimental data were consistent with a model in which the stable secondary structure (in this case, β -sheet) formed first and was necessary for tertiary structure formation.

The spectacular advances presented by these workers led to investigations of the folding mechanism of other small proteins by these techniques. These proteins include barnase (Matouschek et al., 1992), hen egg lysozyme (Radford et al., 1992), T4 lysozyme (Lu & Dahlquist, 1992), ubiquitin (Briggs & Roder, 1992), interleukin-1 β (Varley et al., 1993), and ribonuclease T1 (Mullins et al., 1993). Some of the early data and the techniques used are reviewed elsewhere (Roder, 1989; Englander & Mayne, 1992; Baldwin, 1993).

There are, of course, some limitations to these techniques. The assignment of ^1H resonances is not a trivial task and is constrained, at the present time, to proteins (or fragments of proteins) of less than about 20,000 molecular weight. It is unlikely that one will be able to utilize these techniques for large proteins that do not have clearly defined domains. Interestingly, Dyson et al. (1992a,b) have been using ^1H NMR techniques to examine the folding of peptide fragments comprising the complete sequence

of plastocyanin and myohemerythrin as models for initiation sites. Although it is unlikely that driving forces for the folding process differ markedly between small and large proteins, it is probably equally true that the interaction of different segments of a large protein will influence the folding pathway. It is not possible, therefore, to say whether one will be able to predict such pathways for large proteins based on the experimental information obtained from small ones.

There are other experimental considerations. Some have been discussed by Englander and Mayne (1992) and include the question of whether or not non-hydrogen-bonded peptide protons in poorly understood protein structures exchange at the free peptide rate. Slowly exchanging amide protons may suggest native-like structures that do not exist. On the other hand, rapidly exchanging protons may indicate structural fluctuations in a native-like structure rather than an unfolded protein. For refolding, where it is experimentally best to make a large dilution of denaturant, it is possible that the preferred pathway of refolding may depend on the extent of diluting denaturant (Ropson et al., 1990). Special precautions might be needed for highly charged proteins at low pH and salt concentrations where molten globule structures are investigated (Jeng & Englander, 1991). Indeed, not all proteins may be stable at the pH (~5) needed to slow exchange after partial refolding has taken place. Instability and subsequent unfolding will confound any results. Finally, the experimental mixing procedures can be somewhat complex.

Stopped-flow NMR and protein folding

Some experimental problems could be avoided if it were possible to collect ^1H NMR spectra in real time, i.e., to directly follow the folding process in the NMR tube. Some attempts actually have been made along this line by using temperature as the perturbant of structure. Ribonuclease A at low pH (Adler & Scheraga, 1988; Akasaka et al., 1991) and staphylococcal nuclease (Kautz & Fox, 1993) can be unfolded at high temperatures and refolded upon lowering the temperature. For both proteins it was possible, by exchanging the exchangeable hydrogens with deuterium, to assign a few resonances in the one-dimensional NMR spectra (histidine, for example) and follow changes in intensity of either unfolded or folded resonances during refolding. In these cases, one follows the side-chain environment rather than the formation of secondary structure, and the data obtained were consistent with proline isomerization as a slow rate-limiting step. Although this method has advantages over the hydrogen exchange techniques, it also has a number of limitations. First, one must be able to make assignments in one-dimensional ^1H -NMR spectra for both folded and unfolded protein. Because such spectra are quite complex, usually only a few assignments (e.g., histidine, phenylalanine, or tyrosine) can

be made. Second, the protein must undergo a reversible thermal unfolding over a reasonable temperature range, and the unfolded form should not contain any residual structure. Third, one needs rather high (even for NMR) protein concentrations in order to obtain reasonable signal-to-noise ratios (e.g., 70 mg/mL for staphylococcal nuclease). Finally, in the experiments reported, the earliest time points for data collection were at about 10–12 s, too slow for many systems in which major changes occur in the millisecond range or faster.

One solution to some of these limitations is to incorporate a stopped-flow device into the NMR spectrometer. This involves designing an NMR tube that has a mixer directly above it and in which the volume of solution delivered is as small as possible. There have been a number of stopped-flow NMR devices described in the literature (see McGee & Parkhurst [1990] for references) and at least two such designs have appeared recently (Green et al., 1987; McGee & Parkhurst, 1990). A third, which we have used, and which is easily incorporated into the NMR cavity, will be described in detail elsewhere (Hoeltzli et al., *in press*). This latter device has a dead time of less than 100 ms and uses about 0.7–0.9 mL for each new injection of solution into the (nonspinning) NMR tube. An advantage of any stopped-flow device is that it allows a wide variety of conditions to be examined so that, for example, different concentrations of denaturant may be used or one can use a pH jump rather than a denaturant like urea or guanidine. Data from multiple injections can also be averaged to improve the signal-to-noise ratio. Some of the disadvantages of the stopped-flow technique are: (1) the assignment of peaks in a one-dimensional spectrum to specific protons. The complexity of the proton NMR spectrum makes it difficult, if not impossible, to observe the changes in the spectrum for all residues. The key to studying protein folding using proton NMR may be the ability to select, by either selective NMR pulse sequences or chemical labeling, a subset of protons to be examined during folding; (2) the finite time required to collect data; and (3) (not unrelated) a low signal-to-noise ratio. The complexity of a one-dimensional ^1H spectrum is certainly daunting, but new developments in NMR spectroscopy such as the use of shaped pulses (Kessler et al., 1986; Geen & Freeman, 1991) and selective one-dimensional pulse sequences (Zhou, 1993) offer the promise of selective excitation of specific resonances. The proton resonances, however, must still be assigned, the necessary technologies are not always available, and the relatively long time required for selective shaped pulses can allow significant relaxation to occur during the pulse.

As an alternative solution to the problem of chemical-shift degeneracy and spectral overlap, we suggest the use of ^{19}F instead of ^1H to examine the time dependence of structural changes by NMR. In contrast to proton spectra obtained from hydrogen–deuterium (H/D) exchange experiments, ^{19}F spectra reflect the environment around

the ^{19}F nucleus (and thus the environment of the fluorine-labeled side chain), and not the hydrogen-bonding pattern of amide protons (and thus the formation of secondary structure). In the best case, it might be possible to correlate changes upon refolding by performing both H/D exchange experiments and ^{19}F NMR because these two techniques may provide complementary information.

The fluorine nucleus is similar in size to hydrogen, but fluorine is much more electronegative and the C-F bond has a high dipole moment. Because of this difference, the possibility must be considered that incorporating a fluorine analog of an amino acid may result in a structural perturbation of the protein. Several lines of evidence suggest that for the systems studied to date, it is not a perturbing substitution. Unfolding/refolding curves as a function of denaturant concentrations are the same for proteins containing 6- ^{19}F -tryptophan as for the native protein in at least two proteins, dihydrofolate reductase (DHFR) and IFABP (unpubl. data). In addition, 6- ^{19}F -tryptophan *E. coli* DHFR (unpubl. data) and several fluorine-substituted analogs of *Lactobacillus casei* DHFR (Kimber et al., 1977) possess essentially wild-type enzymatic activity. Similar results have been observed for *E. coli* lactate dehydrogenase (Ho et al., 1989) and *E. coli* D-galactose chemosensory receptor (Luck & Falke, 1991).

^{19}F NMR and protein folding

There are several advantages to using ^{19}F NMR to investigate structural questions related to protein folding. First, the ^{19}F NMR spectra will reflect only those ^{19}F atoms incorporated into the protein or ligand, simplifying the spectrum tremendously and eliminating the need for special equipment or complex shaped pulses. This advantage more than compensates for the slight loss of sensitivity (about 17%) between fluorine and proton NMR. The high sensitivity and lack of background signals in ^{19}F NMR allow reasonably low concentrations of protein to be used. Data can be collected in real time at relatively short intervals limited only by the spin-lattice (T_1) relaxation of the nuclei being observed (typically about 0.3–0.8 s). Second, the simplicity of the ^{19}F spectrum of a labeled protein allows the ready use of powerful methods such as Bayesian probability (Bretthorst, 1990) to extract information from the time-domain spectrum. Given a prior model for the frequencies contained in the time-domain spectrum, such methods offer the possibility of improving the quantification and interpretation of relatively noisy stopped-flow NMR data. Third, because of the simplicity of the fluorine spectrum, one is not limited to small (<20 kDa) proteins by the problem of spectral overlap. Finally, fluorine chemical shifts are spread over a larger range of frequencies than proton chemical shifts and are very sensitive to the environment of the ^{19}F nucleus (Sykes & Hull, 1978; Gerig, 1989). Thus, changes

in the fluorine chemical shift can serve as a reporter for conformational changes in the protein.

If specific fluorine chemical shifts can be assigned to specific locations in the protein, it should be possible to correlate changes in those chemical shifts with structural changes in the protein. Early work with ^{19}F -labeled proteins attempted to interpret fluorine chemical shifts in terms of the resonance of “buried” fluorine nuclei being shifted to a lower field relative to the resonance of an unincorporated amino acid or to its chemical shift in denatured protein (Sykes & Hull, 1978). This downfield shift was sometimes attributed to electric-field effects (Kimber et al., 1977). However, upfield shifts also occur in several systems (e.g., Luck & Falke, 1991), and there was no detailed theory to predict the direction or magnitude of a protein-induced change in fluorine chemical shift. Recently, two groups have attempted to develop such a theory using molecular dynamics. Gregory and Gerig (1991) used a ribonuclease A S-protein/S-peptide complex and correctly predicted the direction of the chemical-shift change of the single fluorine incorporated. They suggested that van der Waals interactions may be the most important factor governing protein-induced fluorine chemical shift changes. De Dios et al. (1993), using galactose-binding protein to test their model, correctly predicted the direction and relative magnitude of the five fluorotryptophan chemical shifts and suggested that electrostatic field effects are the dominant term. If these results prove predictive of protein-induced chemical-shift changes in other systems, they will be helpful in interpreting the fluorine chemical shifts of proteins of known structure. Even in the absence of a successful predictive theory, however, structural correlations can still be made because, with current techniques of molecular biology, specific fluorine resonances can be assigned to specific amino acid residues by using site-directed mutagenesis to create mutant proteins.

Incorporation of ^{19}F into a protein can be accomplished in several ways. For example, the protein can be labeled with a fluorine-containing reagent. McGee and Parkhurst (1990), in their stopped-flow studies, attached a ^{19}F probe (3-bromo-1,1,1-trifluoropropanone) to the β -93 cysteine of hemoglobin to measure the time course of alkaline denaturation. Alternatively, ^{19}F amino acids can be biosynthetically incorporated directly into bacterial proteins or proteins expressed in bacteria. This method had been used in the past (i.e., Kimber et al., 1977), but in older studies it was not possible to easily identify a specific fluorine resonance with a particular amino acid residue. The ability to induce expression of a particular protein solves at least two problems: a high efficiency of labeling by the ^{19}F -substituted amino acid and the ability to identify the specific amino acid using site-directed mutagenesis. Data for the incorporation of 6- ^{19}F -tryptophan into IFABP illustrate a method that can be used (Ropson & Frieden, 1992). Large quantities

of protein were produced by the induction of protein synthesis in *E. coli* W3110trpA33, a tryptophan auxotroph, using pMON-IFABP vectors. Overnight cultures were diluted into medium containing tryptophan and grown to an OD₆₀₀ of 3.0. The cells were then harvested and resuspended in the same medium using 6-fluorotryptophan instead of tryptophan. Protein expression was then induced with nalidixic acid. With IFABP, this procedure resulted in protein labeled more than 90% with fluorotryptophan. The resonance peaks of the two tryptophans of IFABP were identified unambiguously by site-directed mutagenesis, substituting tyrosine for tryptophan at each position and repeating the procedure used for the wild-type protein.

Figure 1 shows spectra of native and unfolded IFABP containing 6-¹⁹F-tryptophans at positions 6 and 82. The unexpected presence of two peaks in the unfolded spectrum reflects the extraordinary sensitivity of the fluoro nucleus to local peptide sequence (Ropson et al., 1990). A similar result has been found for *E. coli* DHFR, a protein that contains five tryptophans. In this case, five site-directed mutants were made in which each had a specific tryptophan replaced with phenylalanine, thus again allowing unambiguous identification of the ¹⁹F NMR peaks with the appropriate tryptophan. Not only were the peaks shifted in the presence of substrates, but four separate peaks (one peak had twice the intensity of the others) were observed in the unfolded protein (S.D. Hoeltzli & C. Frieden, unpubl. data).

One is not limited to tryptophan, of course, because other fluoro-substituted amino acids (i.e., phenylalanine, tyrosine) and other auxotrophs are available. Indeed, because fluorescence and absorbance changes are primarily due to tryptophan, the use of other amino acids may provide more useful information.

Fig. 1. ¹⁹F-NMR spectra of [6FTrp] intestinal fatty acid binding protein (IFABP). The upper trace is in the presence of 8 M urea (unfolded) and the lower trace is in the absence of denaturant (native). Each spectrum consists of 512 transients at 0.7 mM protein in 20 mM KPO₄, 0.1 M KCl, 0.1 mM EDTA, 10% D₂O at pH 7.2 and 25 °C. Peaks due to the 6FTrp residues at Trp 6 and Trp 82 are indicated.

Dynamic NMR and protein folding

Once data such as those shown in Figure 1 are obtained, the next step is to determine how these equilibrium data can be analyzed. The measurement at equilibrium of rates of exchange of chemical reactions or conformational changes by determining the effects of those transitions on the spectrum is termed dynamic NMR (DNMR). Protein folding is a case of conformational exchange in which significant spectral changes are found between folded and unfolded forms of the protein. A broad range of rate constants, from 10⁻² to 10⁴ s⁻¹ (Led et al., 1989), can be determined by available NMR techniques. One such technique that has long been used to examine structural transitions and chemical exchange phenomena at equilibrium is magnetization transfer (reviewed by Led et al. [1989]). It is useful in situations where the exchange is slow on the NMR time scale but still more rapid than the spin-relaxation time. In the case of protein folding, the observation window for the rates of exchange observed by these methods would be between 0.01 and 10 s⁻¹. The kinetics of the reversible transition from native to unfolded protein at specific residues have been observed for several proteins using ¹H NMR (Dobson & Evans, 1984; Evans et al., 1989; Roder, 1989). Typically these studies require that the resonance be isolated in the one-dimensional proton spectrum for the accurate determination of rates, thereby limiting their utility to residues with well-resolved resonances, such as histidine or the aromatic amino acids. However, these experiments do measure the actual rates of conversion of protein from one form to another at specific locations in the structure rather than the global changes monitored by spectral properties. The rates measured thus far correlate very well with more global rate measurements, especially for proline isomerization events (Evans et al., 1989). In addition, these methods are very effective for assigning resonances in the unfolded protein by correlating them with the previously assigned resonances in the native protein under solution conditions in which both are present. This condition has identified residues in unfolded protein that have nonrandom chemical shifts for both BPTI (Roder, 1989) and lysozyme (Evans et al., 1991), suggesting the presence of hydrophobic interactions among side chains in an otherwise unfolded protein. Neri et al. (1992a) have assigned the ¹H resonances in the unfolded form of phage 434 repressor protein using these methods and have shown that some of the residues have nonrandom chemical shifts in the unfolded protein (Neri et al., 1992b,c) and that a small hydrophobic cluster of amino acids is present even in 7 M urea (Neri et al., 1992a). Residues involved in a molten globule folding intermediate at low pH have been identified by saturation transfer with the native resonances of α -lactalbumin (Baum et al., 1989).

For somewhat faster reactions (10–10⁴ s⁻¹ [Led et al., 1989]), the actual shape and/or chemical shift of the res-

onances are perturbed by the exchange process. These phenomena occur when the rate of an exchange process between two (or more) conformations is similar to the chemical-shift difference for a resonance that changes frequency during the conformation change (reviewed by Rao [1989]). For example, if there is a difference of 1 ppm in the chemical shift at 500 MHz for a proton in the folded vs. the unfolded state, rates of exchange between 50 and $5,000\text{ s}^{-1}$ would affect the spectra significantly. For two conformers and at the slower exchange rates, two peaks would be present, but the peaks would be broader than in the absence of exchange. At intermediate rates of exchange, a single broad peak would be observed. Under some conditions, the peak might be so broad that it would effectively disappear into the baseline noise. At high rates of exchange, a single sharper peak would be present with a resonance frequency that would be the average of the frequencies of the two conformers. The perturbed spectra thus provide information on the relative populations of the states of the system, the chemical shifts, the shapes of the lines in the absence of exchange, and the rate of exchange between the states. Computer methods for complete line-shape analysis have been developed (LeMaster et al., 1989; Rao, 1989). These methods fit theoretical spectra calculated for a model for conformational exchange describing the various states of the system to real spectral data until a best fit for the model to the data is achieved. Although these methods have not been applied to the changes in the proton spectrum that occur during folding, we have successfully used complete line-shape analysis to examine the folding of IFABP labeled specifically with fluorine (Ropson & Frieden, 1992).

Figure 2 shows the spectrum of partially unfolded IFABP labeled with 6- ^{19}F -tryptophan. At this intermediate concentration of denaturant, there is spectral overlap between the resonances of the Trp 82 in the native protein and Trp 6 in the unfolded protein at -46.4 to -46.6 ppm. Further, some of the expected intensity appears to be missing, a phenomenon observed frequently in systems in which, as indicated above, intermediate

Fig. 2. ^{19}F -NMR spectra of [6FTrp]IFABP in 4.25 M urea. The smooth line through the data was the best fit to a three-state model, $N = I = U$ (Ropson & Frieden, 1992). The residual is the actual intensity minus the predicted intensity for the fit.

conformational exchange is occurring. In this case, it is possible to fit the spectra obtained at several different denaturant concentrations using the program DNMR5 (LeMaster et al., 1989). The simplest model that fits the spectral data at all denaturant concentrations requires a single intermediate involving Trp 82 but not Trp 6. For the data shown in Figure 2, the model predicts that 65% of the protein is in the native form, 18% is in an intermediate structure, and 17% is completely unfolded. The analysis also provides exchange rate constants; for this denaturant concentration, the rate constant for the $I \rightleftharpoons U$ process is $54 \pm 5\text{ s}^{-1}$.

Figure 3 shows an NMR stopped-flow experiment following urea-induced unfolding of IFABP. The concen-

Fig. 3. Stopped-flow ^{19}F NMR spectra during the unfolding of [6FTrp]IFABP at the specified times following mixing. Native and unfolded resonances are labeled according to their chemical shift in 4.5 M urea. Native [6FTrp]IFABP in buffer was mixed with sufficient urea to denature approximately $\frac{2}{3}$ of the protein present as determined by optical methods. Final solution conditions were 50 mM KPO_4 , 100 mM KCl, 0.1 mM EDTA, 15% D_2O , 1.1 mM IFABP, 1 mM 2-fluorophenylalanine (an internal standard), and approximately 4.5 M urea. Proton decoupling was applied only during acquisition. Spectra were collected every 2 s after mixing. The mixing device was modified from an Applied Photophysics rapid mixer (model RX.1000) with a specially constructed NMR tube. The data collection sequence was repeated 12 times with the same time interval for signal averaging. The resulting summed free induction decays were Fourier transformed with 10 Hz of line broadening. The smooth lines are a model calculated by Bayesian analysis of the time-domain spectrum.

tration of urea used was sufficient to unfold the protein to about 65% of the full extent. The frequencies of the native and unfolded resonance are shown for both tryptophan residues. The time resolution is restricted to a collection interval of 2 s because of the T1 relaxation time of the ^{19}F nucleus. The unfolding process under these conditions is completed by 12 s, and in order to determine species concentration, each time interval can be treated as if the system were at equilibrium. The smooth lines at each time interval represent a model calculated by Bayesian analysis of the time-domain data. The rate of loss of native intensity for tryptophan at position 6 is the same as the fast rate of unfolding of IFABP measured by fluorescence changes under these conditions (I. Ropson & C. Frieden, unpubl. data). Changes in intensity for Trp 82 are complex due to spectral overlap with the unfolded Trp 6 resonance and due to missing intensity caused by conformational exchange of Trp 82 as described above. This experiment illustrates the ability to use a stopped-flow device in the NMR spectrometer to measure specific changes in real time. While the data presented here are preliminary, we believe the method holds great promise.

Conclusion

NMR techniques can give information about specific interactions between residues or partial structures during protein folding or unfolding that are not possible using other methods. The information obtained from the use of pulsed hydrogen-deuterium exchange to examine the formation of hydrogen bonds during folding and from fluorine labeling to examine the formation of tertiary contacts in real time may be crucial to eventually solving the folding problem.

Acknowledgments

This work was supported by NIH grant DK13332. We thank Dr. D.A. D'Avignon for help with the experimental work on ^{19}F NMR and Dr. G.L. Bretthorst for help with the Bayesian analysis.

References

- Adler, M. & Scheraga, H.A. (1988). Structural studies of a folding intermediate of bovine pancreatic ribonuclease A by continuous recycled flow. *Biochemistry* 27, 2471–2480.
- Akasaka, K., Naito, A., & Nakatani, H. (1991). Temperature-jump NMR study of protein folding: Ribonuclease A at low pH. *J. Biomol. NMR* 1, 65–70.
- Alber, T. (1989). Mutational effects on protein stability. *Annu. Rev. Biochem.* 58, 765–798.
- Anfinsen, C.B., Haber, M., Sela, M., & White, F.H., Jr. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. USA* 47, 1309–1314.
- Baldwin, R.L. (1991). Experimental studies of pathways of protein folding. *Ciba Found. Symp. Discussion* 161, 190–201.
- Baldwin, R.L. (1993). Pulsed H/D exchange studies of folding intermediates. *Curr. Opin. Struct. Biol.* 3, 84–91.
- Baum, J., Dobson, C.M., Evans, P.A., & Hanley, C. (1989). Characterization of a partly folded protein by NMR methods: Studies on the molten globule state of guinea pig alpha-lactalbumin. *Biochemistry* 28, 7–13.
- Bretthorst, G.L. (1990). Bayesian analysis. I. Parameter estimation using quadrature NMR models. *J. Magn. Reson.* 88, 533–551.
- Briggs, M.S. & Roder, H. (1992). Early hydrogen-bonding events in the folding reaction of ubiquitin. *Proc. Natl. Acad. Sci. USA* 89, 2017–2021.
- Chaffotte, A.F., Cadieux, C., Guillou, Y., & Goldberg, M.E. (1992a). A possible initial folding intermediate: The C-terminal proteolytic domain of tryptophan synthase beta chains folds in less than 4 milliseconds into a condensed state with non-native-like secondary structure. *Biochemistry* 31, 4303–4308.
- Chaffotte, A.F., Guillou, Y., & Goldberg, M.E. (1992b). Kinetic resolution of peptide bond & side chain far-UV circular dichroism during the folding of hen egg white lysozyme. *Biochemistry* 31, 9694–9702.
- Chothia, C. & Finkelstein, A.V. (1990). The classification and origins of protein folding patterns. *Annu. Rev. Biochem.* 59, 1007–1039.
- Creighton, T.E. (1990). Protein folding. *Biochem. J.* 270, 1–16.
- de Dios, A.C., Pearson, J.G., & Oldfield, E. (1993). Secondary and tertiary structural effects on protein NMR chemical-shifts – An ab initio approach. *Science* 260, 1491–1496.
- Dill, K.A. & Shortle, D. (1991). Denatured states of proteins. *Annu. Rev. Biochem.* 60, 795–825.
- Dobson, C.M. & Evans, P.A. (1984). Protein folding kinetics from magnetization transfer nuclear magnetic resonance. *Biochemistry* 23, 4267–4270.
- Dyson, H.J., Merutka, G., Waltho, J.P., Lerner, R.A., & Wright, P.E. (1992a). Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. I. Myohemerythrin. *J. Mol. Biol.* 226, 795–817.
- Dyson, H.J., Sayre, J.R., Merutka, G., Shin, H.C., Lerner, R.A., & Wright, P.E. (1992b). Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. II. Plastocyanin. *J. Mol. Biol.* 226, 819–835.
- Elove, G.A., Chaffotte, A.F., Roder, H., & Goldberg, M.E. (1992). Early steps in cytochrome c folding probed by time-resolved circular dichroism and fluorescence spectroscopy. *Biochemistry* 31, 6876–6883.
- Englander, S.W. & Mayne, L. (1992). Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct.* 21, 243–265.
- Evans, P.A., Kautz, R.A., Fox, R.O., & Dobson, C.M. (1989). A magnetization-transfer nuclear magnetic resonance study of the folding of staphylococcal nuclease. *Biochemistry* 28, 362–370.
- Evans, P.A., Topping, K.D., Woolfson, D.N., & Dobson, C.M. (1991). Hydrophobic clustering in nonnative states of a protein: Interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. *Proteins Struct. Funct. Genet.* 9, 248–266.
- Geen, H. & Freeman, R. (1991). Band-selective radiofrequency pulses. *J. Magn. Reson.* 93, 93–141.
- Gerig, J.T. (1989). Fluorine nuclear magnetic resonance of fluorinated ligands. *Methods Enzymol.* 177, 3–23.
- Green, D.B., Lane, J., & Wing, R.M. (1987). A standard session stopped-flow NMR tube. *Appl. Spectrosc.* 41, 847–850.
- Gregory, D.H. & Gerig, J.T. (1991). Prediction of fluorine chemical shifts in proteins. *Biopolymers* 31, 845–858.
- Ho, C., Pratt, E.A., & Rule, G.S. (1989). Membrane-bound D-lactate dehydrogenase of *Escherichia coli*: A model for protein interactions in membranes. *Biochim. Biophys. Acta* 988, 173–184.
- Hoeltzli, S.D., Ropson, I.J., & Frieden, C. (in press). Application of equilibrium and stopped-flow ^{19}F NMR spectroscopy to protein folding: Studies of *E. coli* dihydrofolate reductase. *Techniques Protein Chem. V*.
- Jaenicke, R. & Buchner, J. (1993). Protein folding: From “unboiling an egg” to “catalysis of folding.” *Chemtracts – Biochem. Mol. Biol.* 4, 1–30.
- Jeng, M.F. & Englander, S.W. (1991). Stable submolecular folding units in a non-compact form of cytochrome c. *J. Mol. Biol.* 221, 1045–1061.
- Kautz, R.A. & Fox, R.O. (1993). NMR analysis of staphylococcal nuclease thermal quench refolding kinetics. *Protein Sci.* 2, 851–858.
- Kessler, H., Oschkinat, H., & Griesinger, C. (1986). Transformation

- of homonuclear two-dimensional NMR techniques into one-dimensional techniques using Gaussian pulses. *J. Magn. Reson.* 70, 106–133.
- Kiefhaber, T., Schmid, F.X., Willaert, K., Engelborghs, Y., & Chaffotte, A. (1992). Structure of a rapidly formed intermediate in ribonuclease T1 folding. *Protein Sci.* 1, 1162–1172.
- Kim, P.S. & Baldwin, R.L. (1980). Structural intermediates trapped during the folding of ribonuclease A by amide proton exchange. *Biochemistry* 19, 6124–6129.
- Kim, P.S. & Baldwin, R.L. (1982). Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* 51, 459–489.
- Kim, P.S. & Baldwin, R.L. (1990). Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59, 631–660.
- Kimber, B.J., Griffiths, D.V., Birdsall, B., King, R.W., Scudder, P., Feeny, J., Roberts, G.C.K., & Burgen, A.S.V. (1977). ^{19}F nuclear magnetic resonance studies of ligand binding to 3-fluorotyrosine and 6-fluorotryptophan-containing dihydrofolate reductase from *Lactobacillus casei*. *Biochemistry* 16, 3492–3500.
- Kuwajima, K., Garvey, E.P., Finn, B.E., Matthews, C.R., & Sugai, S. (1991). Transient intermediates in the folding of dihydrofolate reductase as detected by far-ultraviolet circular dichroism spectroscopy. *Biochemistry* 30, 7693–7703.
- Led, J.J., Gesmar, H., & Abildgaard, F. (1989). Applicability of magnetization transfer nuclear magnetic resonance to study chemical exchange reactions. *Methods Enzymol.* 176, 311–329.
- LeMaster, C.B., LeMaster, C.L., & True, N.S. (1989). *Program QCMPO59*. Quantum Chemistry Program Exchange, Indiana University, Bloomington, Indiana.
- Lu, J. & Dahlquist, F.W. (1992). Detection and characterization of an early folding intermediate of T4 lysozyme using pulsed hydrogen exchange and two-dimensional NMR. *Biochemistry* 31, 4749–4756.
- Luck, L.A. & Falke, J.J. (1991). Open conformation of a substrate-binding cleft: ^{19}F NMR studies of cleft angle in the D-galactose chemosensory receptor. *Biochemistry* 30, 6484–6490.
- Mann, C.J. & Matthews, C.R. (1993). Structure and stability of an early folding intermediate of *Escherichia coli* trp aporepressor measured by far-UV stopped-flow circular-dichroism and 8-anilino-1-naphthalene sulfonate binding. *Biochemistry* 32, 5282–5290.
- Matouschek, A., Serrano, L., Meiering, E.M., Bycroft, M., & Fersht, A.R. (1992). The folding of an enzyme. V. H/2H exchange-nuclear magnetic resonance studies on the folding pathway of barnase: Complementarity to and agreement with protein engineering studies. *J. Mol. Biol.* 224, 837–845.
- McGee, W.A. & Parkhurst, L.J. (1990). A combined nuclear magnetic resonance and absorbance stopped-flow apparatus for biochemical studies. *Anal. Biochem.* 189, 267–273.
- Mullins, L.S., Pace, C.N., & Raushel, F.M. (1993). Investigation of ribonuclease-t1 folding intermediates by hydrogen-deuterium amide exchange-2D-NMR spectroscopy. *Biochemistry* 32, 6152–6156.
- Neri, D., Billeter, M., Wider, G., & Wüthrich, K. (1992a). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* 257, 1559–1563.
- Neri, D., Wider, G., & Wüthrich, K. (1992b). ^1H , ^{15}N and ^{13}C NMR assignments of the 434 repressor fragments 1–63 and 44–63 unfolded in 7 M urea. *FEBS Lett.* 303, 129–135.
- Neri, D., Wider, G., & Wüthrich, K. (1992c). Complete ^{15}N and ^1H NMR assignments for the amino-terminal domain of the phage 434 repressor in the urea-unfolded form. *Proc. Natl. Acad. Sci. USA* 89, 4397–4401.
- Radford, S.E., Buck, M., Topping, K.D., Dobson, C.M., & Evans, P.A. (1992). Hydrogen exchange in native and denatured states of hen egg-white lysozyme. *Proteins Struct. Funct. Genet.* 14, 237–248.
- Rao, B.D.N. (1989). Nuclear magnetic resonance line-shape analysis and determination of exchange rates. *Methods Enzymol.* 176, 279–311.
- Richards, F.M. (1991). The protein folding problem. *Sci. Am.* 264(1), 54–63.
- Roder, H. (1989). Structural characterization of protein folding intermediates by proton magnetic resonance and hydrogen exchange. *Methods Enzymol.* 176, 446–473.
- Roder, H., Elove, G.A., & Englander, S.W. (1988). Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* 335, 700–704.
- Ropson, I.J. & Frieden, C. (1992). Dynamic NMR spectral analysis and protein folding: Identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by ^{19}F NMR. *Proc. Natl. Acad. Sci. USA* 89, 7222–7226.
- Ropson, I.J., Gordon, J.I., & Frieden, C. (1990). Folding of a predominantly beta-structure protein: Rat intestinal fatty acid binding protein. *Biochemistry* 29, 9591–9599.
- Rose, G.D. & Wolfenden, R. (1993). Hydrogen bonding, hydrophobicity, packing and protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 22, 381–415.
- Schmid, F.X. & Baldwin, R.M. (1979). Detection of an early intermediate in the folding of ribonuclease A by protection of amide protons against exchange. *J. Mol. Biol.* 135, 199–215.
- Scholtz, J.M. & Baldwin, R.L. (1992). The mechanism of alpha-helix formation by peptides. *Annu. Rev. Biophys. Biomol. Struct.* 21, 95–118.
- Sugawara, T., Kuwajima, K., & Sugai, S. (1991). Folding of staphylococcal nuclease A studied by equilibrium and kinetic circular dichroism spectra. *Biochemistry* 30, 2698–2706.
- Sykes, B.D. & Hull, W.E. (1978). Fluorine nuclear magnetic resonance studies of proteins. *Methods Enzymol.* 49, 271–295.
- Udgaonkar, J.B. & Baldwin, R.L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* 335, 694–699.
- Udgaonkar, J.B. & Baldwin, R.L. (1990). Early folding intermediate of ribonuclease A. *Proc. Natl. Acad. Sci. USA* 87, 8197–8201.
- Varley, P., Gronenborn, A.M., Christensen, H., Wingfield, P.T., Pain, R.H., & Clore, G.M. (1993). Kinetics of the folding of an all- β sheet protein interleukin-1 β . *Science* 260, 1110–1113.
- Wagner, G. & Wüthrich, K. (1982). Sequential resonance assignments in proton ^1H nuclear resonance spectra: Basic pancreatic trypsin inhibitor. *J. Mol. Biol.* 155, 347–366.
- Zhou, M.-M., Davis, J.P., & van Etten, R.L. (1993). Identification and pK_a determination of the histidine residues of human low-molecular-weight phosphotyrosyl protein phosphatases: A convenient approach using an MLEV-17 spectral editing scheme. *Biochemistry* 32, 8479–8486.