Biochemistry

 Volume 42, Number 43 November 4, 2003

Current Topics

The Kinetics of Side Chain Stabilization during Protein Folding†

Carl Frieden*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 South Euclid Avenue, St. Louis, Missouri 63110

Received August 20, 2003; Revised Manuscript Received September 24, 2003

ABSTRACT: The rate of stabilization of side chains during protein folding has never been carefully studied. Recent developments in labeling proteins with ^{19}F-labeled amino acids coupled with real-time NMR measurements have allowed such measurements to be made. This paper describes the application of this method to the study of several proteins using 6-^{19}F-tryptophan as the reporting group. It is found that these side chains adopt their final stable state at the last stages of the folding process and that the stabilization of side chains into their final conformation is a highly cooperative process. It is also possible to show the presence of intermediates in which the side chains are not correctly packed. The technique should be applicable to many systems.

It is probably a truism to say that the side chains of a protein define its final structure. Yet, despite intensive investigation using a wide variety of methods, the mechanism by which a protein folds to that final structure remains elusive. There are multiple reasons for such uncertainty including the facts that, in general, the concentration of any intermediate is low, that it has been difficult to measure very fast processes, that the forces between residues have not been clearly defined, that the role of water is unclear, that many experimental methods measure global, rather than residue-specific, effects, that complex computational calculations, such as molecular dynamics, only cover short times, and so on. Considerable progress, both computationally and experimentally, has been made on the basis of the folding of small proteins or peptides that may be considered as the starting point for larger proteins which fold by more complex mechanisms. Current thinking, based on studies with several proteins (e.g., refs *1*−*3*), suggests that small clusters of residues preexist in the unfolded state and that these serve as nuclei for the folding process. Daggett and Fersht (*4*) have recently discussed the various proposals that have been made over the years and have proposed a unifying mechanism for folding involving a nucleation−condensation process.

Left out of much of the discussion about mechanisms of protein folding is the issue of the time dependence of the stabilization of side chains. Yet this is critical, especially in vivo, because it defines not only the process by which the protein reaches its final stable state but also the rates of protein−protein interactions, of protein−ligand interactions, and possibly the rate of any posttranslational modification.

In this review, I examine the role of the later steps in the folding process, specifically those related to the stabilization of the side chains. As will be discussed, these processes can be studied using proteins containing fluorine-labeled amino acids in conjunction with NMR measurements made in real time.

Burial, Packing, and Stabilization of Side Chains

Side chain stabilization differs from side chain packing or burial because stabilization implies that the side chain

† This work was supported by NIH Grant DK13332.

* To whom correspondence should be addressed. Phone: (314) 362-3344. Fax: (314) 362-7183. E-mail: frieden@biochem.wustl.edu.

10.1021/bi030192l CCC: $25.00
Published on Web 10/10/2003

attains the configuration that occurs in the final native form of the protein. Packing and burial only suggest that the side chain occupies some, perhaps transient, stable form. For example, an intermediate in the folding process may have side chains either buried or packed, but repacking may be required to reach the final stable form. Indeed, we shall see an example of that later.

There is, of course, a large literature concerning side chain packing (*5*, *6*). Matthews and co-workers have extensively investigated the role of packing and stability for T4 lysozyme (*7*−*11*). Colon et al. have used site-directed mutagenesis to evaluate the importance of side chain packing in the folding and stability of cytochrome *c* (*12*). Studies such as these deal with the critical issues of side chain packing and thermodynamic stability. In such experimental studies, however, it may be difficult to distinguish between packing and final stabilization.

The purpose of this paper is severalfold: first, to examine the available data on the rates of side chain stabilization; second, to show that such stabilization is the last, or essentially the last, step in the folding process; third, to show that this stabilization occurs coincident with forming the final stable protein; fourth, to show that such stabilization is normally a highly cooperative process; and finally, to show that early side chain stabilization may reflect misfolded intermediates.

Previous Studies

There are a variety of techniques used to measure protein folding. Many, such as changes in absorbance, tryptophan fluorescence, far- or near-UV, CD, and IR, measure global changes rather than changes in specific regions of the protein. Of the techniques to measure protein folding, only a few measure side chain stabilization and fewer still measure the rate of that process for specific residues. By far, the most commonly used technique to measure structural stability in specific regions of a protein is hydrogen/deuterium exchange, as pioneered by Englander (*13*, *14*). This important and useful technique measures the changes in backbone amide protection. It is likely, however, that amide protection and side chain stabilization do not reflect the same process since secondary structure formation, resulting in amide proton protection, may occur early in the folding process while side chain stabilization appears to occur late.

There have been studies that specifically target side chain accessibility during folding. Loh and co-workers, for example, monitored side chain accessibility of apomyoglobin by following the extent of thiol−disulfide exchange (*15*, *16*). Ha and Loh (*15*), using a thiol−disulfide exchange method, examined several cysteine mutants to discern residues that became inaccessible as a function of time. Two (at positions 108 and 110) in the AGH helical interface showed 70−80% protection prior to formation of the native structure while two others (at positions 65 and 134, reflecting the packing of helix E) were buried at a slower rate consistent with the formation of the native structure. For those side chains that pack early, these experiments might not be able to differentiate between being buried or obtaining the correctly stabilized structure of the native state.

Lyon et al. used chemically induced dynamic nuclear polarization (CIDNIP) to probe the accessibility of aromatic side chains during the folding of bovine α-lactoglobulin and lysozme (*17*). An advantage of this method is the time resolution as shown, for example, by Hore et al. (*18*), who utilized stopped-flow photo-CIDNP to observe changes in aromatic side chain accessibility for times as short as 100 ms. With respect to lysozyme, these authors assumed that in the denatured state both tyrosine and tryptophan were largely solvent accessible. In an intermediate state, however, tryptophan side chains were protected. As folding proceeded to the native state, there was a rearrangement in which tyrosines became partially buried while two tryptophans became solvent exposed. This latter rearrangement appeared to be coincident with the formation of the active site. This study clearly showed that side chain burial could not be equated with the final stabilization of the side chain.

Udgaonkar and co-workers examined side chain behavior for both folding and unfolding of the small (89 aa) protein, barstar. For folding, they used a combination of steady-state and time-resolved fluorescent methods, particularly lifetime measurements (*19*), while in the unfolding direction they used time-dependent changes in FRET efficiency (*20*). They concluded that during folding tight packing of side chains in the hydrophobic core occurs before substantial secondary and tertiary structure formation. In agreement with these results, they found that in the unfolding direction structural changes occur faster at the surface than in the protein core (*20*). Again, since they used fluorescence and absorbance methods, they could not distinguish between side chain packing and final stabilization.

In experiments similar to those using ^{19}F-labeled amino acids, as discussed later, van Nuland et al. (*21*) examined the folding of muscle acylphosphatase, monitoring the 1D ^{1}H NMR spectrum in real time. By selecting a small region of the spectrum, they could plot changes in peak intensity as a function of time, showing that the majority of the NMR peak intensity was obtained within about 20 s after dilution from 7 M urea. Interestingly, recovery of activity under the same conditions was much slower. The authors attributed the slow activity recovery to proline isomerization (*22*). Nevertheless, from the time dependence of the appearance of methyl and methylene resonances in a 1D NMR experiment, the authors did conclude that side chain stabilization was an early event relative to activity recovery.

This observation raises the important issue of the relation of the rate of side chain stabilization during folding to the rate of recovery of enzymatic activity. That is, can a structure appear to be folded (by circular dichroism or fluorescence changes) yet require further adjustment for gain of function? Unfortunately, few experiments have been able to answer this question, yet it deals directly with the rate of side chain stabilization. The acylphosphatase results mentioned earlier, for example, suggest that activity recovery is slow relative to structure formation. These are, however, not easy experiments. First, activity and side chain stabilization are performed at very different protein concentrations, and second, the addition of substrate(s) to measure activity may influence the folding rate.

Finally, there have been some experiments related to destabilization of side chains during unfolding. Laurents and Baldwin, for example, have used 1D ^{1}H NMR to examine the kinetics of the slow unfolding of egg white lysozyme (*23*). Lysozyme unfolding is slow, thus allowing these

measurements. With the *Escherichia coli* dihydrofolate reductase (DHFR), Hoelztli and Frieden (*24*) showed that the tryptophan side chains became destabilized very early during unfolding. With lysozyme, however, solvent-exposed tryptophans become destabilized early while those buried in the core, along with the histidine, become destabilized very slowly, simultaneously and concomitant with CD changes. The authors suggest that the difference between their study and that of Hoeltzli and Frieden may lie in the greater sensitivity of ^{19}F to detect motional differences.

Why Use Fluorine-Labeled Amino Acids?

As noted in the two recent reviews (*25*, *26*), there are many properties of fluorine that make it a useful probe for studies of protein structure, function, and folding. For example, the fluorine nucleus is small, only slightly larger than the hydrogen nucleus. Using a fluorine-labeled amino acid might be expected to result in minimal perturbation to the structure, stability, and functionality of the protein even though there is a large dipole moment for the C−F bond. Because fluorine is extraordinarily sensitive to its environment and to local shielding effects, the NMR peaks are typically well resolved from one another in a 1D NMR spectrum. Even in the denatured state the peaks are frequently resolved (*24*, *27*, *28*). Proteins of high molecular weight can also be examined since the spectral resolution is large. Fortunately, a number of fluorine-labeled amino acids are available. In all, data obtained with fluorine-labeled amino acids are useful because they measure exactly what we wish to determine, i.e., the stabilization of the environment around the fluorine nucleus.

Taking advantage of the sensitivity of ^{19}F chemical shifts, fluorine has been used extensively to follow structural changes (e.g., ref *29*). Theory relating these chemical shifts to structure is difficult (*30*). Pearson et al. (*31*) suggested that electric field strength effects were major contributors to ^{19}F shielding and found reasonably good agreement between experimental and calculated chemical shifts for the *E. coli* galactose binding protein (*31*). However, in a later paper, Feeney et al. stated that additional contributions to shielding were needed for aromatic amino acids labeled with fluorine (*29*).

We initially undertook the study of protein folding using ^{19}F-labeled amino acids with the idea that there would be differential stabilization of the side chains in different regions of the protein and therefore would allow one to follow the folding pathway. Indeed, this is the case but since NMR peaks of the side chains (or at least aromatic side chains) appeared to occur very late in the folding process, what we are actually following are the last stages of the folding process. The results, however, give important insights into the folding process.

One should not ignore the ability to use 1D ^{1}H NMR spectra in conjunction with the fluorine data for studying the kinetics of side chain stabilization. While a 1D ^{1}H NMR protein spectrum has a large number of overlapping resonances, some peaks are associated with the side chains of specific residues. These include histidine and the indole of tryptophan, as well as methyl and methylene groups which can be identified and assigned. The increasing availability of triple resonance cryoprobes and the use of selective ^{15}N or ^{13}C enrichment make this approach attractive. If the NMR assignments have already been made, identification of these peaks in the 1D ^{1}H NMR spectrum is simple. Alternatively, resonances could be identified by site-directed mutagenesis, especially for larger proteins. Van Nuland et al. (*21*), using muscle acylphosphatase, as described above, did follow the refolding using a 1D ^{1}H NMR spectrum of the methyl, methylene region. The spectra were complex, and the observed peaks were not assigned to specific amino acid residues, but the assignments should be possible. This appears to be the only example using 1D ^{1}H NMR to follow side chain stabilization during folding. The method, however, has great potential. It is not discussed further here because there is much more data using ^{19}F-labeled proteins.

Incorporating Fluorine-Labeled Amino Acids into Proteins

While there are a number of fluorine-labeled amino acids that can be incorporated into proteins, the most commonly used are the aromatic amino acids. Methods used are described elsewhere (*32*). Of particular interest is the ability to incorporate a single *p*-^{19}F-phenylalanine into any position in the protein (*33*). However, only results with 6-^{19}F-tryptophan will be discussed here.

Results Using ^{19}F-Labeled Tryptophan

Fluorine-labeled amino acids have been used for many years to examine the effect of ligand addition on protein structure (*34*−*40*). The sensitivity of the chemical shift of 6-^{19}F-typtophan in the folding process was first experimentally described by Ropson and Frieden (*27*), using the intestinal fatty acid binding protein that contains two tryptophan residues. They noted that the chemical shifts of the 6-^{19}F-tryptophan residues were a function of the urea concentration. Further, the dependence of one of the two tryptophans was the same as the denaturation curve (followed by circular dicroism or fluorescence) while the other was not. From these data, the authors could calculate the fractional concentration of species as a function of urea and could estimate the concentration of an intermediate in the presence of the folded and unfolded forms (*27*). They also showed that the intermediate persisted at high denaturant concentrations and concluded that it probably reflected a hydrophobic cluster of amino acids close to one of the two tryptophans. Obviously, it would be useful to know the time dependence of the formation of this intermediate. For that purpose, one needs to collect data as a function of time and as quickly as possible. To accomplish this, we designed and built a stopped-flow NMR tube (*41*).

One disadvantage of using 6-^{19}F-labeled amino acids is that the relaxation time of the fluorine nucleus is sufficiently slow (∼0.5 s) that the fastest time at which data can be collected after starting the folding or unfolding process is about 1 s. Unfortunately, the intestinal fatty acid binding protein folds so rapidly (<5 s) that stopped-flow NMR experiments were not feasible. Thus, to determine the kinetics of side chain stability, experiments have to be performed with proteins that fold more slowly. Several examples are given here.

E. coli Dihydrofolate Reductase. The folding properties of the *E. coli* dihydrofolate reductase have been extensively studied. It is a relatively small protein (18 kDa) consisting

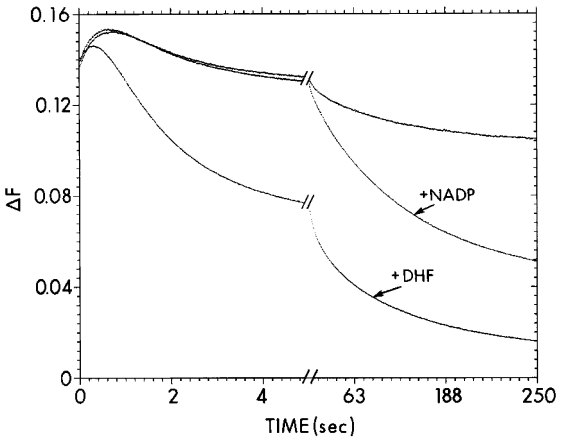

FIGURE 1: Stopped-flow fluorescence changes (arbitrary units) observed on the refolding of *E. coli* dihydrofolate reductase from 4.5 M urea to 1.29 M urea. For the curves labeled NADP and DHF, the ligands (NADP and dihydrofolate) were added at the start of the refolding process. No ligand was added in the top curve. The experiments were performed at 10 °C in 0.02 M phosphate buffer, pH 7.2, containing 1 mM dithiothreitol and 0.1 mM EDTA. The initial enzyme concentration was 8.5 μM. Data were taken from Frieden (*44*).

of eight β-strands and four α-helices (*42*, *43*). The C-terminal region comprising about a third of the protein consists primarily of β-strands. Our studies have always been consistent with a mechanism in which there is a sequential series of events (*44*−*46*) while Matthews and co-workers have proposed a channel mechanism of at least four different paths to form the native protein (*47*−*50*). Both groups agree that there are at least four phases observed by fluorescence changes on refolding from denaturant. A typical renaturation curve, in the presence and absence of substrates, is shown in Figure 1 (*44*). Refolding is a reasonably slow process, taking several minutes with an initial increase in fluorescence followed by a decrease which can be deconvoluted into three phases. Of interest is that quenching of the fluorescence by NADP appears only at later stages of folding while that by dihydrofolate occurs earlier. These results suggest that the NADP binding site is formed later in the folding process than the dihydrofolate binding site. Figure 2 shows an NMR refolding experiment under similar conditions using 6-^{19}F-tryptophan as the reporting group in which the unfolded protein is diluted into buffer containing NADP. The numbers at the top and bottom in this figure show the assignments of the resonance peaks to the specific residues as determined by site-directed mutagenesis (*51*). There is a considerable amount of information that can be obtained from these data. First, there is a loss of intensity from the denatured peaks with no equivalent appearance of intensity in the native peaks. Not shown is the observation that in this experiment about 20% of the intensity of the denatured peaks disappears before the first time point is obtained (1.5 s). Second, when nondenatured peaks appear, they always appear at chemical shifts of the native protein. Interestingly, peaks associated with the apoprotein appear before those of the NADP-bound protein, suggesting, again, that the NADP binding site forms late and reinforces the view that the gain of structure and that of function may have very different time scales. Third, the appearance of the native peaks is slow and correlates with the last phases of folding as shown in Figure 1, and fourth, their appearance is essentially highly cooperative. These observations will be discussed below.

PapD. PapD is a chaperone essential for the formation of P pili in pathogenic bacteria (*52*). A two-domain protein of molecular mass 25 kDa, it is composed primarily of β-strands

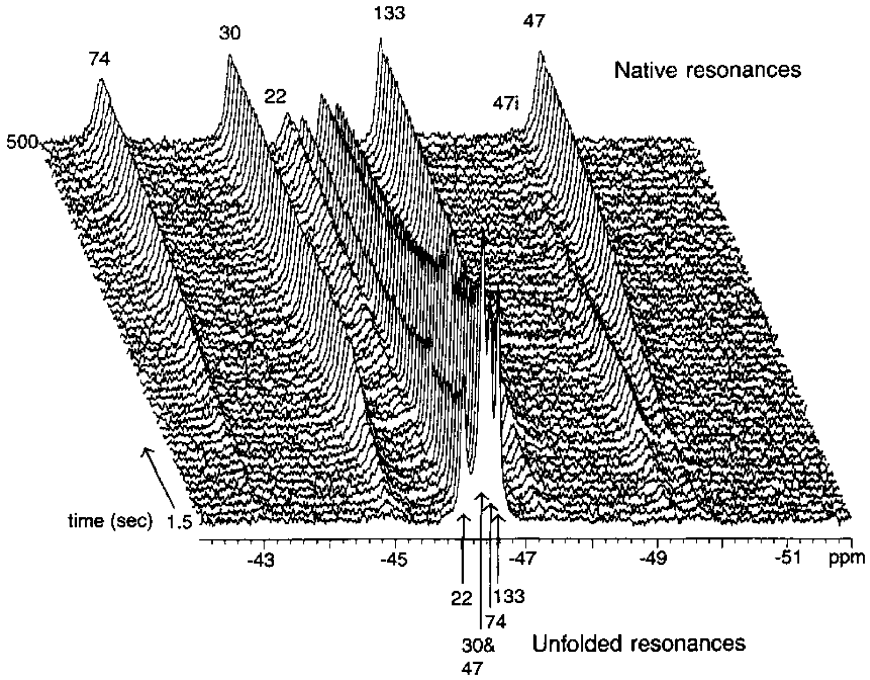

FIGURE 2: Stopped-flow ^{19}F NMR spectra for the refolding of *E. coli* dihydrofolate reductase, labeled with 6-^{19}F-tryptophan, from 5.5 M urea to 2.75 M urea in the presence of NADP. The numbers at the top and bottom of the figure identify the five native and unfolded tryptophan residues in the protein. The peak labeled 47i is associated with the apoenzyme. The disappearance of this peak correlates with the formation of the NADP binding site. Experiments were performed using a Varian superconducting fluorine probe. Forty-one separate injections were summed for each time point. The experiment was performed at 5 °C in 0.05 M phosphate buffer, pH 7.2, containing 15 mM dithiothreitol and 0.1 mM EDTA. The final protein concentration was 0.61 mM. Data were taken from Hoeltzli and Frieden (*46*).

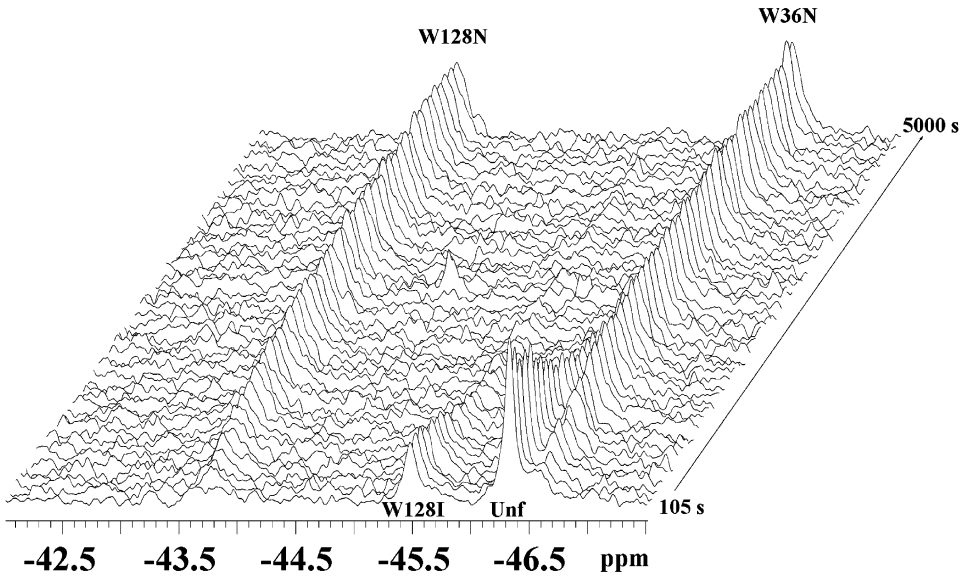

FIGURE 3: Real-time ^{19}F NMR spectra of PapD labeled with 6-^{19}F-tryptophan on refolding from 4.5 M urea to 2.25 M urea. The numbers at the top of the figure identify the two tryptophan residues in the protein. Note that there is an peak representing an intermediate in the folding process that appears early (W128i) and disappears as the native peaks form. The peak labeled Unf corresponds to the unfolded protein. Experiments were performed on a Varian fluorine Cryo-Q probe with a final protein concentration of 70 μM. Only one refolding experiment was needed to collect the data. The experiment was performed at 20 °C in 0.03 M MOPS buffer at pH 7.0. The experiment is more fully described by Bann et al. (*28*).

with each domain having a immunoglobulin-like fold (*53*). There is a single tryptophan in each domain, and therefore one can follow the folding of each domain by following the NMR spectrum of the protein labeled with 6-^{19}F-tryptophan. An NMR refolding experiment is shown in Figure 3, and it is immediately apparent that there are some similarities and some differences between the data for this protein compared with dihydrofolate reductase. The most obvious similarity is that the native peaks appear slowly and cooperatively with the same rate constants. The obvious difference is that a non-native peak appears early and disappears as the native peaks appear. We have interpreted this result to mean that the C-terminal domain, represented by Trp128, collapses early to a structure with the side chain stabilized in a non-native environment. Then as the N-terminal domain folds, the C-terminal domain readjusts to its correct configuration as a consequence of the domain−domain interaction (*28*). The slow folding of the N-terminal domain is almost certainly controlled by proline isomerization since the folding process is enhanced in the presence of proline isomerase (Bann and Frieden, unpublished results).

Adenosine Deaminase. Murine adenosine deaminase (mADA) is a 40 kDa protein containing a tightly bound Zn atom essential for activity (*54*). In contrast to dihydrofolate reductase and PapD, the protein is highly helical with a $(\beta/\alpha)_8$ topology. The folding kinetics are complicated by the presence of the tightly bound Zn. The metal, however, can be removed to form the apoenzyme. Figure 4 shows the ^{19}F NMR spectrum of the holoprotein, apoprotein, and denatured protein. As with dihydrofolate reductase and PapD, the peaks for the denatured protein are sharp and close to that of tryptophan itself. The spectrum of the apoprotein differs from that of the native and the denatured protein in that it is broad and unstructured (Shu and Frieden, unpublished results), yet the total amplitude is quite similar. Thus in this case, the side chains are not freely rotating and at the same time are not sensing their native environment. Refolding from this state is extremely slow (data not shown), suggesting that this form may be misfolded. Again, however, the appearance of resonances corresponding to the native resonances occurs slowly and cooperatively (Shu and Frieden, unpublished results).

Interpreting the Results

In the experiments using ^{19}F-labeled amino acids, we define side chain stabilization as the appearance of NMR resonances with the same chemical shift as in the native protein. The rate of side chain stabilization is measured as the change in peak intensity as a function of time. There are several possible scenarios for stabilization of side chains. Those described here make some assumptions as to the mechanism of the folding process. Thus, let us assume that there is transient nativelike residual structure in the unfolded state (*1*, *3*, *55*−*58*), that there is a collapsed state, and that the final folding process occurs after that collapsed state. Under these conditions, side chain stabilization might occur at any one of these steps: in those nativelike regions of the unfolded protein, in the collapsed state, or during the final steps.

Appearance of Stabilized Side Chains. The NMR data clearly show that no appreciable side chain stabilization occurs in the denatured state even if that state has regions that flicker between unfolded and nativelike structures. Unfolded resonances are sharp and always occur close to or at the same chemical shift as the isolated amino acid. The small differences in chemical shift from the isolated amino acid may simply reflect a dependence on residues nearby in the sequence.

It is generally agreed that some sort of condensation step occurs early in the folding process. With respect to the NMR data, this step may be associated with the partial loss of

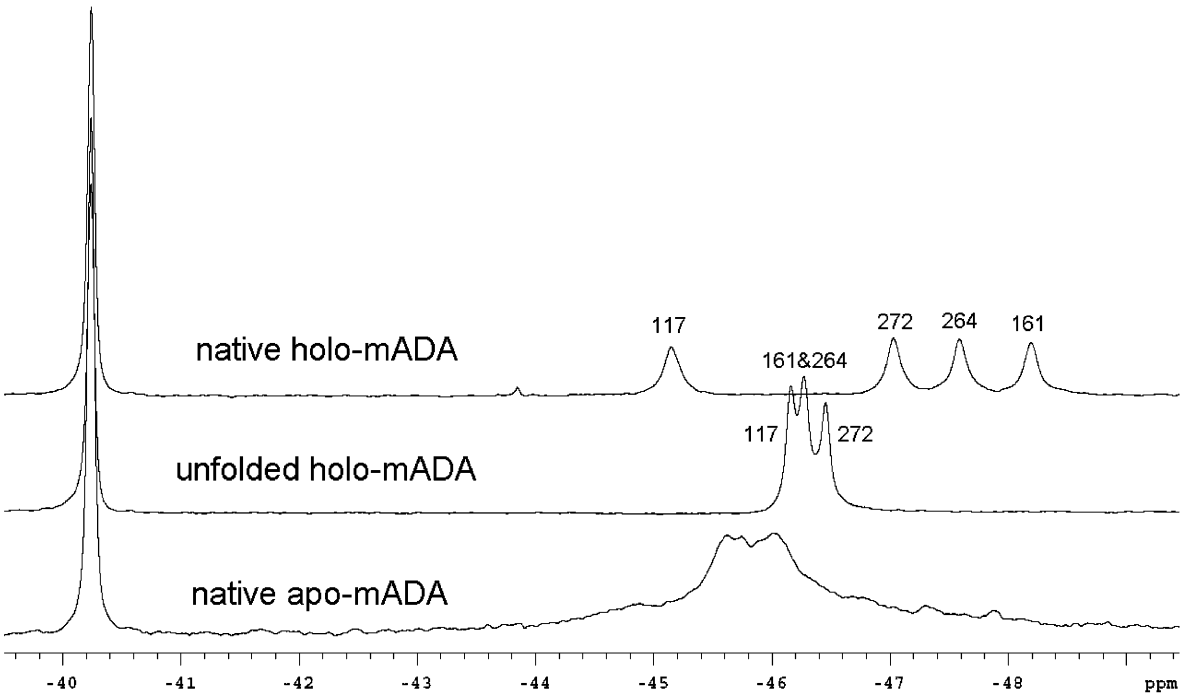

FIGURE 4: NMR spectra of murine adenosine deaminase labeled with 6-^{19}F-tryptophan. Spectra shown are for the native protein, the denatured protein in 8 M urea, and the apoprotein in which the tightly bound Zn has been removed. The numbers in the figure identify the four tryptophan residues where possible. Spectra were collected with a Varian fluorine Cryo-Q probe. All spectra were collected at 20 °C in 20 mM Tris-HCl, pH 7.4, 2 mM dithiothreitol, 5% D_2O, and 0.2 mM *p*-^{19}F-phenylalanine as an internal reference. The protein concentration used for the holo-mADA at 0 or 8 M urea was 85 μM, while that for the apoprotein was 100 μM. Data are from Shu and Frieden (unpublished results).

intensity of the denatured peaks without an equivalent appearance of nativelike peaks. In this collapsed state, there may be a considerable amount of secondary structure, but the side chains have not become stabilized and therefore appear to be sampling large areas of conformational space. In the case of PapD, one of the two tryptophans gets packed early but in a different environment than that of the final structure. Because there is no good theory relating fluorine chemical shifts to structure, it is not yet possible to define what the intermediate structure may be. It disappears, however, at the same rate as the formation of the final stabilized form. In all cases that we have examined stabilization of side chains occurs late in the folding process. Furthermore, the data show that side chain stabilization is a highly cooperative process.

In the native state of a protein the side chains are closely packed. Stabilization of the side chains appears to be concomitant with the loss of entropy associated with packing. The high degree of cooperativity observed for side chain stabilization suggests that there is a final compaction of the structure with concomitant loss of internal water. One could argue, for example, that native peaks do not appear early because the residues are in a liquid environment sensing numerous microenvironments. Unfortunately, the results with fluorine only allow speculation on this issue.

What Is the Role of Proline Isomerization? The proteins discussed here fold slowly, most likely as a consequence of slow *cis*−*trans* proline isomerization. Most proteins, especially larger ones, do contain proline so that, in a practical sense, one cannot ignore the role of the isomerization in attempting to characterize the mechanism of protein folding. For proteins that contain multiple proline residues, kinetic simulations show that the folding rates can become extraordinarily slow if several proline residues isomerize at similar slow rates. The slow kinetics of side chain stabilization may also reflect proline isomerization. This suggests the critical role that proline isomerization can play in the cooperative stabilization of side chains. The role of proline isomerization as a molecular switch has recently been thoroughly discussed by Andreotti (*59*).

It should be noted, however, that slow folding need not be a consequence of proline isomerization. The half-time of folding of the intestinal fatty acid binding protein (which does not contain proline) can be increased to about 50 s just by mutating residues in turns (*60*). Indeed, a number of proteins contain non-prolyl *cis* bonds (http://www.imb-jena.de/ImgLibDoc/cispep/non_proline/IMAGE_CISPEP2.html). Pappenberger et al., for example, have assigned a slow phase of the folding of tendamistat to the *cis*−*trans* isomerization of a non-prolyl peptide bond (*61*). Dihydrofolate reductase contains a Gly−Gly *cis* bond (*62*), and it is not known what role this bond plays in protein folding. It is of interest that mutating the Gly−Gly to Ala−Gly leads to loss of enzymatic activity (*62*).

Why Is There an Initial Loss of Intensity in the Unfolded Resonances? In 1973, Anfinsen (*55*) suggested that portions of a protein chain may serve as nucleation sites and might "flicker" in and out of the conformation that they occupy in the final protein. As mentioned above, there is accumulating evidence for residual structure in the unfolded state and that this residual structure is nativelike (*1*, *3*, *56*−*58*). The fact that only denatured peaks are observed at high denaturant concentrations suggests that if there are nativelike structures, their concentration must be extremely low. With dihydrofolate reductase, there is a rapid partial loss of denatured peaks without appearance of the native peaks observed on

refolding. This is also seen with PapD when *p*-^{19}F-phenylalanine is used as the probe (Bann and Frieden, unpublished results). While we cannot characterize this early step, it must represent an intermediate in which the side chains are no longer in the unfolded state but sense a large number of microenvironments such that the peak is broadened into the baseline. This would be the so-called collapsed state about which the ^{19}F NMR data can currently say very little.

Final Comments. The method described here should be applicable to a wide variety of systems. All too often, investigators find techniques that they believe will help to define the mechanism of protein folding. One is tempted to go down that same path with the type of data obtained in these experiments. The advantage here is the ability of follow specific residues. On the other hand, there remain a number of issues that need to be addressed. The results here say almost nothing about the early stages of the folding process. Because there is no good theory for fluorine chemical shifts and structure, we know little about the nature of any intermediate forms. At this time, we do not know what the major determinants that lead to stabilization of the side chains are. The ^{19}F NMR studies discussed here have dealt only with tryptophan. Examination of the kinetics of side chain stabilization using other amino acids may lead to different results and should be a fruitful area of investigation. Finally, the results from these experiments should be useful in determining the mechanism of protein folding particularly for those investigators interested in the computational approaches to the later stages of the folding process.

ACKNOWLEDGMENT

The author thanks Drs. George Rose, David Cistola, Rohit Pappu, James Bann, Sydney Hoeltzli, and Qin Shu for helpful discussions.

REFERENCES

1. Shortle, D., and Ackerman, M. S. (2001) *Science 293*, 487−489.
2. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) *Science 295*, 1719−1722.
3. Hodsdon, M. E., and Frieden, C. (2001) *Biochemistry 40*, 732−742.
4. Daggett, V., and Fersht, A. R. (2003) *Trends Biochem. Sci. 28*, 18−25.
5. Ponder, J. W., and Richards, F. M. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 421−428.
6. Chothia, C., Levitt, M., and Richardson, D. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4130−4134.
7. Zhang, X. J., Baase, W. A., and Matthews, B. W. (1991) *Biochemistry 30*, 2012−2017.
8. Gassner, N. C., Baase, W. A., and Matthews, B. W. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12155−12158.
9. Eriksson, A. E., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1992) *Nature 355*, 371−373.
10. Liu, R., Baase, W. A., and Matthews, B. W. (2000) *J. Mol. Biol. 295*, 127−145.
11. Xu, J., Baase, W. A., Baldwin, E., and Matthews, B. W. (1998) *Protein Sci. 7*, 158−177.
12. Colon, W., Elove, G. A., Wakem, L. P., Sherman, F., and Roder, H. (1996) *Biochemistry 35*, 5538−5549.
13. Englander, S. W. (2000) *Annu. Rev. Biophys. Biomol. Struct. 29*, 213−238.
14. Englander, S. W., Sosnick, T. R., Englander, J. J., and Mayne, L. (1996) *Curr. Opin. Struct. Biol. 6*, 18−23.
15. Ha, J. H., and Loh, S. N. (1998) *Nat. Struct. Biol. 5*, 730−737.
16. Feng, Z., Butler, M. C., Alam, S. L., and Loh, S. N. (2001) *J. Mol. Biol. 314*, 153−166.
17. Lyon, C. E., Suh, E. S., Dobson, C. M., and Hore, P. J. (2002) *J. Am. Chem. Soc. 124*, 13018−13024.
18. Hore, P. J., Winder, S. L., Roberts, C. H., and Dobson, C. M. (1997) *J. Am. Chem. Soc. 119*, 5049−5050.
19. Sridevi, K., Juneja, J., Bhuyan, A. K., Krishnamoorthy, G., and Udgaonkar, J. B. (2000) *J. Mol. Biol. 302*, 479−495.
20. Sridevi, K., and Udgaonkar, J. B. (2003) *Biochemistry 42*, 1551−1563.
21. van Nuland, N. A., Chiti, F., Taddei, N., Raugei, G., Ramponi, G., and Dobson, C. M. (1998) *J. Mol. Biol. 283*, 883−891.
22. Chiti, F., Taddei, N., Giannoni, E., van Nuland, N. A., Ramponi, G., and Dobson, C. M. (1999) *J. Biol. Chem. 274*, 20151−20158.
23. Laurents, D. V., and Baldwin, R. L. (1997) *Biochemistry 36*, 1496−1504.
24. Hoeltzli, S. D., and Frieden, C. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9318−9322.
25. Danielson, M. A., and Falke, J. J. (1996) *Annu. Rev. Biophys. Biomol. Struct. 25*, 163−195.
26. Gerig, J. T. (2001) in *Biophysical textbook online* (http://www.biophysics.org/btol/).
27. Ropson, I. J., and Frieden, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7222−7226.
28. Bann, J. G., Pinkner, J., Hultgren, S. J., and Frieden, C. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 709−714.
29. Feeney, J., McCormick, J. E., Bauer, C. J., Birdsall, B., Moody, C. M., Starkmann, B. A., Young, D. W., Francis, P., Havlin, R. H., Arnold, W. D., and Oldfield, E. (1996) *J. Am. Chem. Soc. 118*, 8700−8706.
30. Oldfield, E. (2002) *Annu. Rev. Phys. Chem. 53*, 349−376.
31. Pearson, J. G., Oldfield, E., Lee, F. S., and Warshel, A. (1993) *J. Am. Chem. Soc. 115*, 6851−6862.
32. Frieden, C., Hoeltzli, S. D., and Bann, J. G. (2003) *Methods Enzymol*. (in press).
33. Furter, R. (1998) *Protein Sci. 7*, 419−426.
34. Sykes, B. D., and Hull, W. E. (1978) *Methods Enzmol. 49*, 270−295.
35. Luck, L. A., and Falke, J. J. (1991) *Biochemistry 30*, 4248−4256.
36. Luck, L. A., and Falke, J. J. (1991) *Biochemistry 30*, 4257−4261.
37. Falke, J. J., Luck, L. A., and Scherrer, J. (1992) *Biophys. J. 62*, 82−86.
38. Sun, Z. Y., Pratt, E. A., Simplaceanu, V., and Ho, C. (1996) *Biochemistry 35*, 16502−16509.
39. Duewel, H., Daub, E., Robinson, V., and Honek, J. F. (1997) *Biochemistry 36*, 3404−3416.
40. Vaughn, M. D., Cleve, P., Robinson, V., Duewel, H. S., and Honek, J. F. (1999) *J. Am. Chem. Soc. 121*, 8475−8478.
41. Hoeltzli, S. D., Ropson, I. J., and Frieden, C. (1994) in *Technology in Protein Chemistry*, pp 455−465, Academic Press, New York.
42. Bystroff, C., Oatley, S. J., and Kraut, J. (1990) *Biochemistry 29*, 3263−3277.
43. Bystroff, C., and Kraut, J. (1991) *Biochemistry 30*, 2227−2239.
44. Frieden, C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4413−4416.
45. Hoeltzli, S. D., and Frieden, C. (1996) *Biochemistry 35*, 16843−16851.
46. Hoeltzli, S. D., and Frieden, C. (1998) *Biochemistry 37*, 387−398.
47. Touchette, N. A., Perry, K. M., and Matthews, C. R. (1986) *Biochemistry 25*, 5445−5452.
48. Garvey, E. P., Swank, J., and Matthews, C. R. (1989) *Proteins 6*, 259−266.
49. Iwakura, M., Jones, B. E., Falzone, C. J., and Matthews, C. R. (1993) *Biochemistry 32*, 13566−13574.
50. Jones, B. E., Jennings, P. A., and Matthews, C. R. (1993) *Biophys. J. 64*, A176.
51. Hoeltzli, S. D., and Frieden, C. (1994) *Biochemistry 33*, 5502−5509.
52. Lindberg, F., Tennent, J. M., Hultgren, S. J., Lund, B., and Normark, S. (1989) *J. Bacteriol. 171*, 6052−6058.
53. Holmgren, A., and Branden, C. I. (1989) *Nature 342*, 248−251.
54. Sharff, A. J., Wilson, D. K., Chang, Z., and Quiocho, F. A. (1992) *J. Mol. Biol. 226*, 917−921.
55. Anfinsen, C. B. (1973) *Science 181*, 223−230.
56. Evans, P. A., Topping, K. D., Woolfson, D. N., and Dobson, C. M. (1991) *Proteins 9*, 248−266.
57. Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) *Science 257*, 1559−1563.

58. Blanco, F. J., Serrano, L., and Forman-Kay, J. D. (1998) *J. Mol. Biol. 284*, 1153−1164.
59. Andreotti, A. H. (2003) *Biochemistry 42*, 9515−9524.
60. Kim, K., and Frieden, C. (1998) *Protein Sci. 7*, 1821−1828.
61. Pappenberger, G., Aygun, H., Engels, J. W., Reimer, U., Fischer, G., and Kiefhaber, T. (2001) *Nat. Struct. Biol. 8*, 452−458.
62. Matthews, D. A., Alden, R. A., Bolin, J. T., Freer, S. T., Hamlin, R., Xuong, N., Kraut, J., Poe, M., Williams, M., and Hoogsteen, K. (1977) *Science 197*, 452−455.

BI030192L