Nature of the Early Folding Intermediate of Ribonuclease A†

Jayant B. Udgaonkar‡ and Robert L. Baldwin*,§

Department of Biochemistry, Beckman Center, Stanford University School of Medicine, Stanford, California 94305-5307, and National Center for Biological Sciences, TIFR Center, P.O. Box 1234, Indian Institute of Science Campus, Bangalore 560 012, India

Received September 16, 1994; Revised Manuscript Received January 6, 1995⊗

ABSTRACT: A previous study of the folding pathway of the major unfolded species of ribonuclease A by pulsed hydrogen exchange [Udgaonkar, J. B., & Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87,* 8197−8201] showed that there is a major early folding intermediate (I_1) that resembles a molten globule species in having stable secondary structure while lacking buried tyrosine side chains. Earlier work showed that there is also a late native-like folding intermediate (I_N) that can bind the specific inhibitor 2′CMP and that has buried tyrosine side chains. Results are reported here indicating that I_1 has a well-developed tertiary structure even though its tyrosine side chains are not buried. First, optical stopped-flow experiments suggest that I_1 binds 2′CMP. Second, the protection against hydrogen exchange is similar in I_1 and I_N for almost all protected amide protons studied. Third, analysis of the mechanism of hydrogen exchange in I_1 confirms the large protection factors reported earlier for probes in the β-sheet of ribonuclease A and indicates that the β-sheet is formed in I_1. Other experiments are also reported that test the interpretation of pulsed hydrogen exchange studies of the folding pathway of ribonuclease A.

The refolding pathway of the major unfolded form of RNase A (U_SII) has two known folding intermediates: the early intermediate I_1[1] (Udgaonkar & Baldwin, 1990) and the late, native-like intermediate I_N (Cook et al., 1979; Schmid & Blaschek, 1981; Schmid, 1983):

$$U_S\text{II} \rightarrow I_1 \rightarrow I_N \rightarrow N \qquad (1)$$

U_SII is a slow-folding species arising from *cis* → *trans* isomerization after unfolding of the two cis proline residues of RNase A (see Discussion); it accounts for about 65% of the total unfolded protein (Schmid, 1983). The final step in the folding of U_SII, $I_N \rightarrow N$, consists of proline isomerization coupled to a conformational change.

Pulsed H-exchange studies on the folding pathways of proteins have yielded useful structural information [reviews in Baldwin (1993) and Dobson et al. (1994)]. Previous work by Udgaonkar and Baldwin (1990) gave the kinetics with which I_1 is formed, the locations of protected peptide NH protons in I_1, and their protection factors. It also gave information about how the protection factors change with time during the formation of I_1.

The first problem considered here is whether or not I_1 is a typical molten globule intermediate without fixed tertiary interactions between side chains [see review in Pititsyn (1992)]. The locations of protected peptide NH protons indicate that I_1 has extensive native-like secondary structure. Stopped-flow absorbance measurements show that the buried tyrosine side chains of native RNase A are not buried in I_1, suggesting that fixed side-chain interactions are not present. Nevertheless, the protection factors of most β-sheet probes are surprisingly large if I_1 is a molten globule intermediate; they are too large to measure (≥1000) by stopped-flow hydrogen exchange. In order to test for specific tertiary structure in I_1, its ability to bind the specific inhibitor 2′CMP was measured by optical stopped-flow experiments, using the change in absorbance of 2′CMP at 254 nm that accompanies its binding to RNase A (Anderson et al., 1968; Garel & Baldwin, 1973; Schmid & Blaschek, 1981). The large protection factors found for β-sheet probes in I_1 have been reinvestigated by testing if they had been calculated for the wrong exchange mechanism. They were calculated for EX2 exchange, which is the mechanism found commonly in studies of proteins [review by Englander and Kallenbach (1984)]. Peptide NH protons which fail to undergo exchange by the EX2 mechanism in a 37 ms pulse at pH 11, 10 °C, can be calculated to have protection factors greater than 1000 (see Discussion). If exchange occurs instead by the uncommon EX1 mechanism, which is not base-catalyzed, then these protons could fail to undergo exchange simply because the pulse length is too short. This possibility is tested here by varying the pulse length.

Another problem considered here is why only 40% of the unfolded U_SII molecules form the early folding intermediate when I_1 is first formed. This failure to show complete formation of a particular intermediate is common in studies of folding intermediates by pulsed hydrogen exchange [see review by Englander and Mayne (1992)]. The approach taken here is to ask if any structure can be detected in the unfolded molecules that fail to form I_1 initially, by using pulse-labeling at various pH values to measure protection factors.

Finally, a basic negative control for the detection of folding intermediates by pulsed hydrogen exchange is reported here. The test is to monitor refolding at a denaturant concentration (2 M GdmCl) close to the transition zone where unfolding

† This research was supported by NIH Grant GM 19988. J.B.U. was a fellow of the Jane Coffin Childs Foundation.

* Correspondence should be addressed to this author.
‡ Indian Institute of Science Campus.
§ Stanford University School of Medicine.

⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1995.

[1] Abbreviations: I_1, early folding intermediate; I_N, late, native-like intermediate; GdmCl, guanidinium chloride; RNase A, ribonuclease A; H-exchange, hydrogen exchange.

0006-2960/95/0434-4088$09.00/0

begins. Folding intermediates are likely to be unstable in these conditions and consequently not observable by pulsed hydrogen exchange. This test was made in earlier studies of folding intermediates detected by $^3H-^1H$ exchange (Schmid & Baldwin, 1979; Kim & Baldwin, 1980). The results of this negative control, applied here to the exchange of individual peptide NH protons, are used to test the accuracy of the pulse-labeling measurements and to determine the rate constants for base-catalyzed hydrogen exchange of individual peptide NH protons in the unfolded protein at the start of refolding in 2 M GdmCl. The rate constants in 2 M GdmCl are then compared to the ones measured for the unfolded protein at the start of refolding in standard folding conditions.

EXPERIMENTAL PROCEDURES

Materials. Bovine pancreatic ribonuclease A was obtained from Sigma (grade XII A) and purified chromatographically (Garel, 1976). Guanidinium chloride was ultrapure grade from Schwarz/Mann. D_2O (99.8%) was from Cambridge Isotope Laboratories; [^{14}C]formaldehyde and ^3H_2O were from New England Nuclear. Deuterated RNase A was prepared by dissolving RNase A in D_2O at pH 3, heating the solution to 60 °C for 20 min, and then lyophilizing the protein, after which the procedure was repeated (Udgaonkar & Baldwin, 1988). ^{14}C-Labeled RNase A was prepared by sodium cyanoborohydride reduction of the protein by [^{14}C]formaldehyde (Jentoft & Dearborn, 1979). Deuterated guanidinium chloride was prepared by repeated evaporation from 99.8% D_2O in a rotary evaporator. The preparation of tritiated RNase A was analogous to that of deuterated RNase A: RNase A was dissolved in H_2O containing 5 mCi/mL ^3H_2O, and the solution was heated to 60 °C for 20 min. All other chemicals were reagent grade.

Rapid Mixing Techniques. All exchange experiments on folding were done on a rapid (millisecond) mixing quench-flow machine that has been previously described in detail (Cash & Hess, 1981). The mixing dead-time was 5 ms for each of the three consecutive mixing events. For optical monitoring of folding, a modified Gibson-Durrum stopped-flow instrument was used (Garel et al., 1976). All experiments were done at 10 °C.

Unfolding of RNase A. The protein was unfolded by dissolving it in an unfolding buffer: 2.65 M deuterated guanidinium chloride (concentration checked by refractometry), 40 mM glycine, in D_2O, pH* 2. pH* refers to the pH measured in the presence of D_2O without correction for isotope effects. (Most exchange experiments are made in 94% H_2O; see below.) The concentration of the unfolded RNase A in solution was 60−70 mg/mL.

$^2H-^1H$ Exchange Experiments. These experiments were carried out in two sets of conditions: (1) *strongly stabilizing conditions* were 0.4 M sodium sulfate, 0.25 M GdmCl, and 0.05 M sodium formate, pH 4; (2) *destabilizing conditions* were 2 M GdmCl, 0.05 M sodium formate, pH 4. Refolding of the unfolded RNase A solution (in D_2O) was initiated by diluting 10.5-fold into a refolding buffer (in H_2O). At different times after the start of refolding, exchange was initiated by mixing with an exchange buffer, involving a further 1.5-fold dilution with H_2O. Because exchange takes place in a solution that is 94% H_2O, 6% D_2O, no account is taken of the deuterium isotope effect on the exchange rate, and the pH values are also not corrected for the isotope effect. The pH during the exchange pulse was fixed at any value between 7 and 11 by suitable adjustment of the pH of the exchange buffer. The exchange pulse was terminated by mixing with a quench buffer, so that the final pH was 2.9. The refolding reaction was then allowed to go to completion (10 min) at this pH. The compositions of the refolding, exchange, and quench buffers for experiments done in the presence of 0.4 M sodium sulfate have been described previously (Udgaonkar & Baldwin, 1990). For experiments done in the presence of 2 M GdmCl, the concentration of that salt in the refolding and exchange buffers was suitably adjusted, and the quench buffer contained sodium sulfate and no GdmCl so that the final quench conditions were 1.5 M GdmCl, 0.15 M sodium sulfate, and 0.1 M sodium formate, pH 2.9. For experiments in which the effect of the exchange pulse on native protein was studied, the fully folded and deuterated protein was dissolved in a D_2O buffer containing 0.4 M sodium sulfate, 0.25 M GdmCl, and 0.05 M sodium formate, pH* 4; the exchange pulse was then applied as in the case of the refolding experiments, as described above. All buffers were freshly prepared for each experiment. For the zero time point, the exchange pulse was applied directly to the unfolded protein solution. The preparation of the NMR sample (pH* 3.5) of the fully folded RNase A has been described (Udgaonkar & Baldwin, 1988, 1990).

The equations used to fit the results when either the pH or the pulse length of the pulse-label experiment was varied are $y = (1 - e^{-k\Delta t})$ and $k = k_o(10^{pH-9})$, where Δt is the pulse length, y is the proton occupancy, and k_o is the exchange rate constant at pH 9.

NMR Assay for $^2H-^1H$ Exchange. A two-dimensional homonuclear *J*-correlated (COSY) spectrum (Aue et al., 1976) of the sample from each time point was recorded at 30 °C on a General Electric GN-500 spectrometer. Data acquisition and processing and normalization of spectra to one another were as described previously (Udgaonkar & Baldwin, 1988, 1990). The intensities of the $C_\alpha H-NH$ cross-peaks (the proton occupancies) in each spectrum were determined by calculating the volume integrals of the cross-peaks, after first setting the base line of the spectrum to zero. The proton occupancy is a direct measure of the extent of labeling by exchange (proton incorporation) that occurs at the specific backbone amide hydrogen site when the exchange pulse is applied at a specific time during the folding process.

$^1H-^3H$ Exchange Experiments. Refolding conditions were 0.4 M sodium sulfate, 0.25 M GdmCl, and 0.05 M sodium formate, pH 4 (strongly stabilizing conditions), and buffers were identical to those used for the $^2H-^1H$ exchange experiments, except that the unfolding buffer was made in H_2O and not D_2O, and that the exchange buffer contained ^3H_2O. The unfolded protein concentration was either 60−70 mg/mL or 5-fold lower. The tritium concentration during the exchange pulse was 4 mCi/mL, and the pH of the exchange pulse was 10. After folding was complete, the ^3H_2O was separated from the tritiated protein by passing the solution over a Sephadex G-25 column equilibrated with 0.05 M sodium formate, pH 3.5. After 20 h of exchange-out at 5 °C, pH 3.5, the extent of labeling by ^3H, which had exchanged-in during the exchange pulse, was determined (Schreier, 1977). The extent of labeling by the exchange pulse was compared to the extent of label retained by a

sample of completely tritiated ribonuclease, again after 20 h of exchange-out at 5 °C. In all $^1H-^3H$ exchange experiments, ^{14}C-RNase A was used as an internal concentration standard.

Optical Experiments. Folding was monitored either by (1) observing the accompanying increase in tyrosine absorbance at 287 nm or by (2) observing the change in absorbance at 254 nm that accompanies the binding of the specific inhibitor 2′CMP or by (3) observing the decrease in tyrosine fluorescence at 305 nm on excitation at 278 nm. The wavelength 254 nm is used to monitor 2′CMP binding because folding monitored by tyrosine absorbance is silent here (Schmid & Blaschek, 1981). As with the exchange experiments described above, refolding at pH 4 was initiated by 10.5-fold dilution of the unfolded protein solution into refolding buffer. Refolding conditions were 0.4 M sodium sulfate, O.25 M GdmCl, pH 4 (0.05 M sodium formate) or pH 6 (0.05 M sodium cacodylate). The protein concentrations used were 12−15 mg/mL. For manual mixing experiments, a Beckman DU-64 spectrophotometer and a Perkin-Elmer LS-5 fluorescence spectrophotometer were used. Absorbance measurements were made using manual mixing as well as stopped-flow mixing, while fluorescence measurements were only made with manual mixing.

RESULTS

Optical Studies of Folding. In the folding conditions used here, which contain the stabilizing sulfate anion, the initial reaction forming I_1 from U_SII occurs within 100 ms (Udgaonkar & Baldwin, 1990). Because the formation of I_1 is monitored by protection against exchange, and because the protection factors increase with time of folding, the relaxation time τ for the U_SII → I_1 step is not known accurately.

For the U_F → N reaction $\tau = 40$ ms, and for the U_SI → N reaction, τ is about 50 s. Thus, the U_SII → I_N folding pathway is isolated kinetically from the faster U_F → N reaction and the slower U_SI → N reaction. The I_N → N step, which is not studied here, is a very slow reaction ($\tau = 67$ s, measured by fluorescence) that can be monitored either by fluorescence (Schmid, 1981; Schmid et al., 1986) or by sequential unfolding−refolding assays that measure the U_F:U_S ratio after unfolding, at various stages in the folding process (Cook et al., 1979).

In Figure 1, the folding of U_S species, in 0.4 M sodium sulfate, 0.25 M GdmCl, and 0.05 M sodium cacodylate, pH 6, is shown as monitored by two optical probes: (1) the increase in tyrosine absorbance at 287 nm, which monitors the formation of the native-like intermediate I_N (Cook et al., 1979; Schmid, 1981), and (2) the decrease in absorbance at 254 nm, which occurs on binding the specific inhibitor 2′CMP. The kinetics for the two probes have been compared at pH 6, and not at pH 4, because the binding of 2′CMP is strongly pH-dependent, with a maximum at pH 5.8 (Anderson et al., 1968), and cannot be measured at pH 4. Figure 1 shows that the change in absorbance at 254 nm, that occurs on binding of 2′CMP, is more than 3 times faster than the change in tyrosine absorbance at 287 nm, that monitors formation of the late-folding intermediate I_N. Thus, 2′CMP binds to a folding intermediate that is formed earlier than I_N. The folding of the fast-folding species U_F cannot be observed at 254 nm because the change in absorbance occurs within the instrumental dead-time (30 ms). The addition of

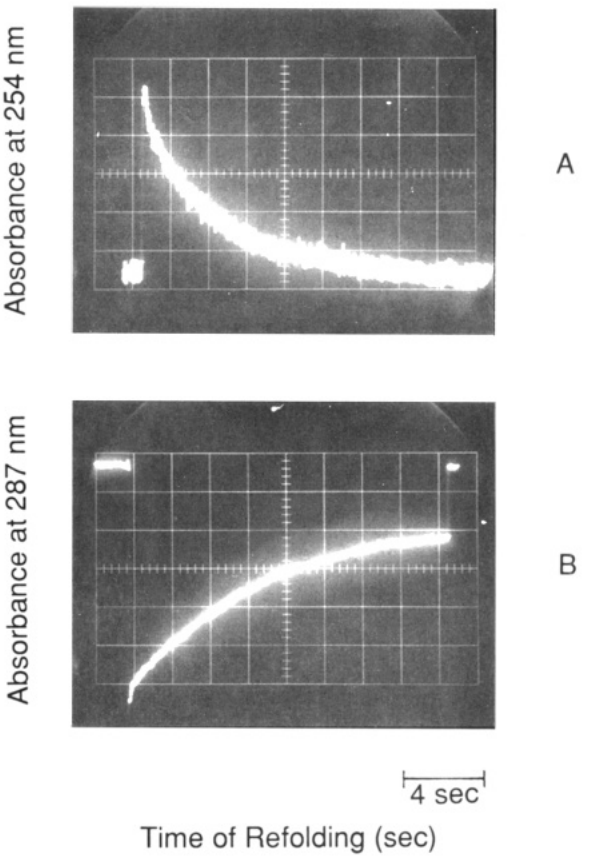

FIGURE 1: Kinetics of folding of U_S species monitored by optical probes. The change in absorbance at 254 nm on binding 2′CMP (A) and the change in tyrosine absorbance at 287 nm (B) were measured on a stopped-flow apparatus. Folding was carried out in 0.4 M sodium sulfate, 0.25 M GdmCl, and 0.05 M sodium cacodylate, pH 6, 10 °C. The concentration of protein during folding was 90 μM in both cases, and the concentration of 2′CMP was 150 μM in (A). The time constant for the folding of U_SII is 14 s when measured by tyrosine absorbance, and 4 s when measured by the absorbance change accompanying 2′CMP binding. The time constant for the folding of U_F, measured by the change in absorbance at 287 nm, was determined to be 15 ms. The folding of U_F was too fast to be monitored by the absorbance change at 254 nm.

2′CMP, at concentrations up to 300 μM, has no effect on the folding kinetics either of U_SII of U_F, as monitored by the change in tyrosine absorbance. The time constant for the I_N → N reaction, measured by fluorescence, is 67 s (data not shown). In the absence of any sodium sulfate but in otherwise identical conditions, the time constants for the U_F → N, U_SII → I_N, and I_N → N reactions are 0.05, 20, and 72 s, respectively (data not shown).

$^1H-^3H$ Pulse-Labeling Studies. In these pulse-labeling experiments for NMR, which utilize $^2H-^1H$ exchange, the initial concentration of RNase A used is high (60−70 mg/mL). High initial concentrations are desirable to obtain a sufficiently high final sample concentration for a good NMR spectrum. To test whether the concentration of RNase A could be a significant variable, pulse-labeling experiments utilizing $^1H-^3H$ exchange were performed. In such experiments, the protein concentration can be easily changed over a wide range, from that used in stopped-flow optical experiments (10−15 mg/mL, initial) to that used in the pulse-labeling experiments for NMR. Two protein concentrations were used, 15 and 75 mg/mL, and a 37 ms $^1H-^3H$ exchange pulse was applied 21 ms after initiation of folding; experi-

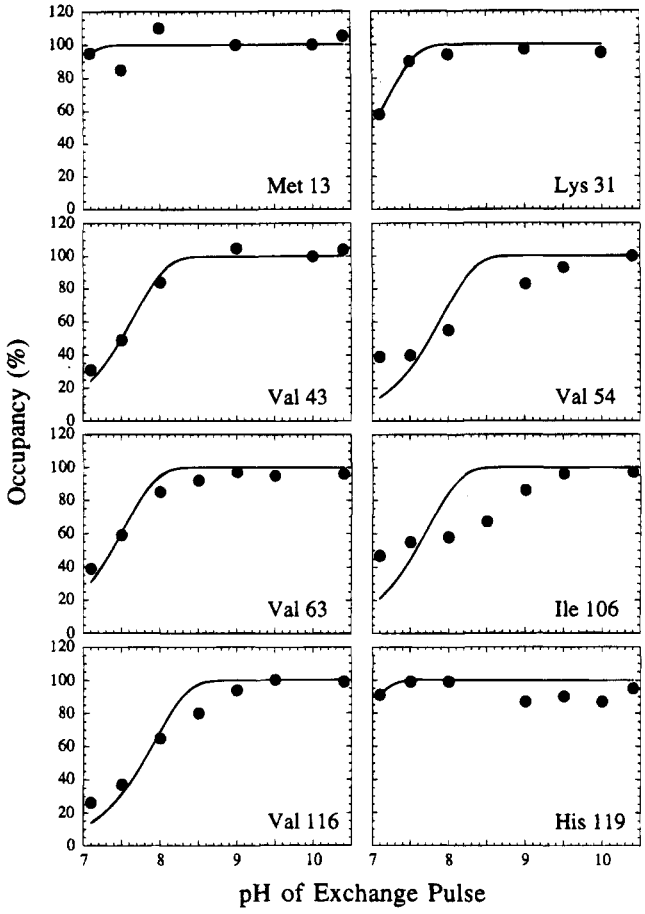

FIGURE 2: Folding in 2 M GdmCl: pH profiles. The proton occupancy, or proton-labeling, is plotted as a function of the pH of the labeling pulse. Refolding was initiated in 2.0 M GdmCl, 0.05 M sodium formate, pH 4. The pH of the 37 ms exchange pulse, applied 400 ms after initiation of folding, was varied between 7.1 and 11. Data for eight different amide protons are shown, fitted to $y = 100\%\ [1 - \exp(-\Delta t k_0 10^{(pH-9)})]$.

mental conditions were otherwise identical to those used in the $^2H-^1H$ pulse-labeling experiments. In both cases, the extent of labeling was 75% of the label retained (50 tritium atoms/RNase A molecule) in a completely tritiated sample of RNase A (after exchange-out, see Experimental Procedures). In a separate experiment in which exchange-out from fully tritiated samples of RNase A was studied, it was observed that the number of stable amide protons in native RNase A was unaffected by the presence of sodium sulfate (in concentrations up to 0.4 M) during the required exchange-out step (see Experimental Procedures).

$^2H-^1H$ Pulse-Labeling of Native RNase A. Thirty-seven millisecond exchange pulses at different pH values were applied to native protein to determine the stability of individual amide protons in the native protein against exchange with solvent protons. Only two amide protons used previously as probes of folding were labeled to any significant extent. The S59 amide proton was significantly labeled from pH 9 upward, and the V43 amide proton was labeled significantly when the pH of the pulse was above 10. The other amide protons that are used as probes of folding (Udgaonkar & Baldwin, 1990) were labeled to less than 10% if at all. The amide protons of M79 and I107 have not been used as probes of folding, but they were measured in this experiment and they were significantly labeled above pH 9.

Table 1: Exchange Rate Constants (s^{-1}) for Unfolded RNase A in Refolding Conditions[a]

residue	standard conditions[b]	2 M GdmCl[c]	k_{pred}[d]
V43		516 ± 62	35
V47	134 ± 49		69
V54	37 ± 15	302 ± 109	40
V63	261 ± 63	788 ± 102	126
I81	157 ± 41		112
Y97	204 ± 175		162
I106	87 ± 17	552 ± 247	78
V108	17 ± 5		35
V116	35 ± 7	314 ± 63	68
V118	38 ± 7		35

[a] Refolding occurs at pH 4.0, 10 °C, and pulse-labeling of the refolding material is made at pH values in the range pH 7−11. Exchange rate constants are determined at pH 9.0, 10 °C (see Experimental Procedures and legends to Figures 2 and 5). [b] 0.4 M Na_2SO_4 0.25M GdmCl; the exchange rate constant is found from the curve showing the extent of pulse-labeling versus pH (see Figure 5). [c] Measured from the pH profile of pulse-labeling (see Figure 2). [d] Predicted from data given by Bai et al. (1993) and Connelly et al. (1993), with $\log k_{pred} = \log k_B + \log [OH^-]$, using $\log k_B = 8.02$ (for ND/H_2O) for poly(DL-alanine) in low salt and $\log [OH^-] = -5.54$, and with side chain corrections added.

Folding in 2 M Guanidinium Chloride. Figure 2 shows the results of varying the pH of an exchange pulse applied 400 ms after initiation of refolding in 2 M GdmCl at pH 4. pH profiles are shown for eight different peptide NH protons. The results are fitted to an equation for base-catalyzed exchange (see Experimental Procedures). The rate constant for exchange at pH 9.0 is computed from the pH midpoint and is given in Table 1. For almost all peptide NH protons, complete labeling occurs under these conditions: a plateau level of 100% labeling is observed for all 27 peptide NH protons except V47 and V108. For each of the peptide NH protons, the extent of labeling decreases from this plateau level as the pH is decreased. This is the result of exchange being base-catalyzed: as the pH during exchange is decreased, the duration of the pulse (37 ms) starts to limit complete labeling of an unprotected amide hydrogen site. Table 1 shows that 2 M GdmCl accelerates the exchange rates of these amide protons to a surprising extent.

Folding in Strongly Stabilizing Conditions. (a) Pulse-Length Dependence. The length of the exchange pulse (at pH 10), applied 400 ms after initiation of folding, was varied between 8 ms and 66 ms. In Figure 4, results for four peptide NH protons are shown. The extent of labeling by a pulse that is 37 ms long is clearly already at the plateau level. Only when the length of the pulse is below 10 ms does the extent of labeling decrease. The extent of labeling does not increase when the length of the pulse is increased to more than 37 ms, as shown in a separate experiment (data not given) in which the exchange pulse at pH 9 was applied 600 ms after initiation of folding; the extent of labeling did not increase even when the length of the pulse was increased to 100 ms.

(b) pH Profiles after 4 s of Folding. The pH profiles were also measured at 4 s after the start of folding, when I_N is the major species present. These data are shown for eight amide protons in Figure 3. All peptide NH proton probes show approximately 15% labeling at pH 9. Those probes that have rising profiles of exchange above pH 9 show measurable labeling of I_N. In the case of Val 43 and Ser 59, any native protein present will also contribute to the observed labeling at high pH.

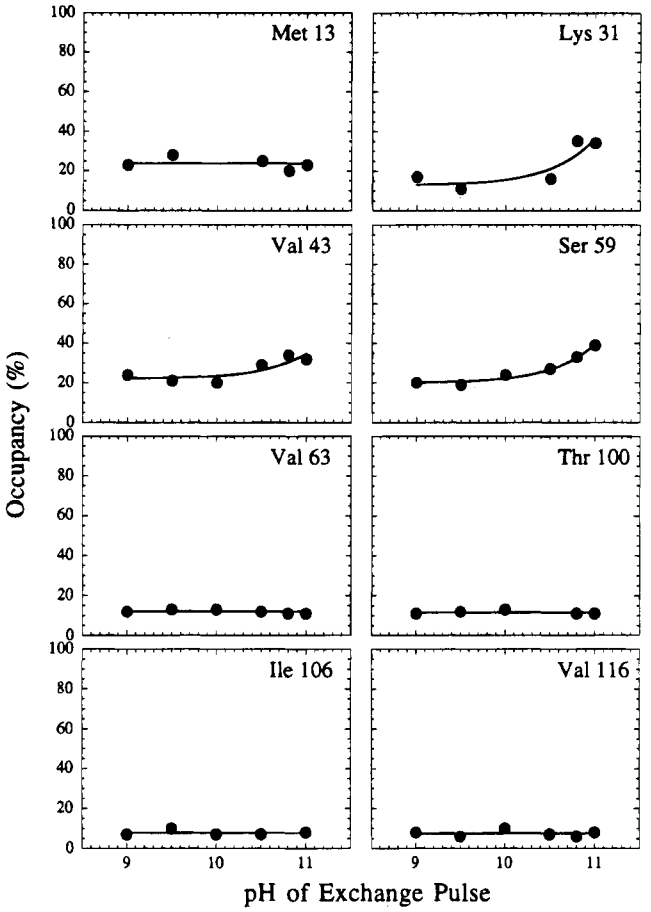

FIGURE 3: Stability of I_N to exchange. The pH profiles measured 4 s after initiation of folding in 0.4 M sodium sulfate were obtained by applying 37 ms exchange pulses at that time. Data for eight different amide protons are shown.

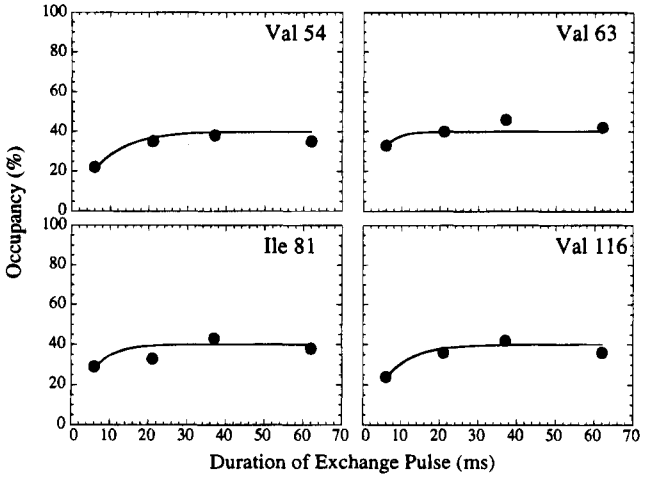

FIGURE 4: Dependence of the extent of labeling on the length of the exchange pulse. The exchange pulses (pH 10) were applied 400 ms after initiation of refolding in 0.4 M sodium sulfate, 0.25 M GdmCl, and 0.05 M sodium formate, pH 4 (strongly stabilizing conditions). Data for four different peptide NH protons are shown, fitted to $y = 40\%\ [1 - \exp(-\Delta t k_o 10^{(pH-9)})]$.

(c) pH Profiles of Exchange for Molecules That Have Not Formed I_1 after 0.4 s of Folding. Exchange profiles were measured between pH 7 and 10 for the fraction of molecules that have not formed I_1 after 0.4 s folding (part of U_SII and all of U_SI, see below: about 40% total). The results were fitted to an equation for base-catalyzed exchange, and the exchange rate constant at pH 9 was determined. pH profiles

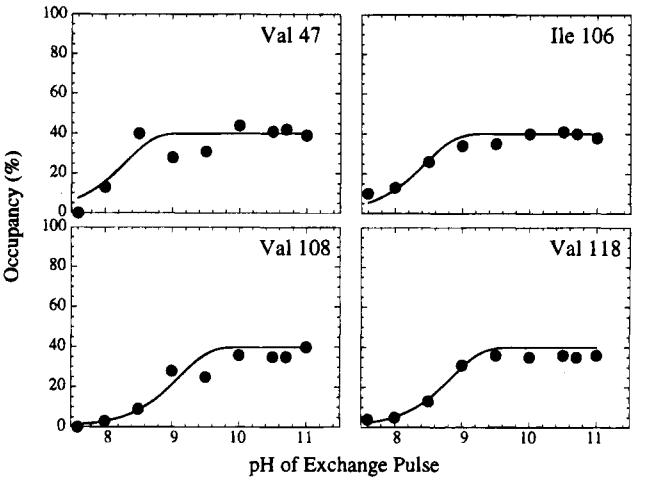

FIGURE 5: Exchange behavior of peptide NH protons in the fraction of protein that fails to form I_1 initially, at 400 ms after the start of folding in standard conditions, including 0.4 M Na_2SO_4. The dependence of exchange on the pH of a 37 ms pulse is shown for four peptide NH protons, fitted to $y = 40\%\ [1 - \exp(-\Delta t k_o 10^{(pH-9)})]$.

are shown for four protons in Figure 5, and the rate constants at pH 9 are summarized in Table 1, where they are compared with predicted values based on model peptide data (Bai et al., 1993).

DISCUSSION

Analysis of Optical Stopped-Flow Experiments in Terms of the Different Unfolded Forms of RNase A. The direct folding pathway of the major unfolded species of RNase A, U_SII, has been represented by a simple sequential mechanism:

$$U_S\text{II} \rightarrow I_1 \rightarrow I_N \rightarrow N \quad (65\%) \qquad (1)$$

where I_1 and I_N are kinetic folding intermediates, formed in observable kinetic steps. Detailed analysis indicates that I_1 is in fact formed in a series of small steps, not yet resolved (Udgaonkar & Baldwin, 1990).

Unfolded RNase A contains a fast-folding species, U_F, and slow-folding species, U_S, in a 20:80 U_F:U_S ratio (Garel & Baldwin, 1973; Schmid, 1982). The U_S species are complex: according to Schmid (1983), the major U_SII species accounts for about 65% of unfolded RNase A, and there is at least one minor U_SI species that accounts for 15%. The refolding reactions of these other unfolded species can be written:

$$U_F \rightarrow N \quad (20\%, \text{fast}) \qquad (2)$$

$$U_S\text{I} \rightarrow N \quad (15\%, \text{very slow}) \qquad (3)$$

Little is known about intermediates in these folding reactions. A new analysis by Houry et al. (1994) indicates that U_F itself is complex, and can be resolved into two species with different refolding rates (see below).

Brandts et al. (1975) suggested that the *cis*–*trans* isomerization of proline residues after unfolding accounts for the different unfolded forms. This suggestion was confirmed by Schultz et al. (1992), who used directed mutagenesis to replace the two *cis* proline residues of RNase A, Pro 93 and Pro 114, and who found that Pro 93 and Pro 114 collectively account for the existence of U_SII and U_SI in wild-type RNase

A. The problem of accounting for the refolding behavior of U_F, U_SII, and U_SI by assigning appropriate *cis* and *trans* isomers of Pro 93 and Pro 114 has remained unsolved, however. Recently Houry et al. (1994) proposed such a model on the basis of new data for the unfolding and refolding kinetics of RNase A. They assume that a *trans* isomer of Pro 93 has a major effect in slowing down the refolding kinetics while a *trans* isomer of Pro 114 has only a minor effect. They find a new, minor (7%), very fast-folding species, and they suggest that U_F has a *trans* isomer of Pro 114 while U_SI has a *trans* isomer of Pro 93. According to their model, U_SII has *trans* isomers of both Pro 93 and Pro 114.

The relaxation times that characterize the three folding reactions (eq 1−3, above) are obtained by optical monitoring of the folding process (see Results). The early intermediate I_1 is not detected by optical methods. Two optical probes of folding were used to obtain the time constants for the three parallel folding reactions: tyrosine absorbance and tyrosine fluorescence. There is some controversy in the literature [see, for instance, Schmid (1983) and Lin and Brandts (1983)] regarding the analysis of the kinetic curves for the absorbance change and for the fluorescence change.

A standard kinetic test of whether I_1 is the direct precursor to I_N is to find out if the rate of formation of I_N is maximal when the population of I_1 is maximal, in other words to find out if a lag in the formation of I_N occurs as the population of I_1 first builds up. No such lag is observed, but I_1 forms very early, within 100 ms, at the same time that U_F is folding to native protein (N). Since tyrosine absorbance does not distinguish between I_N (forming from I_1) and N (forming from U_F), an extremely high level of accuracy in its measurement would be required to detect a lag in the formation of I_N. Moreover, this lag should be small because the precursor, I_1, is populated initially to only 40%. Given the accuracy both in the measurements of tyrosine absorbance and in the measurements of the formation of I_1 using $^2H-^1H$ exchange measurements, it is not surprising that the lag in the formation of I_N is not observed.

Pulse-labeling can be used to find if an intermediate such as I_1 is on the productive pathway (Baldwin, 1991). If the structure that is present when I_1 is formed initially breaks down again later, before the intermediate enters the productive folding pathway, these molecules will be rapidly labeled when I_1 breaks down, provided that the labeling pulse extends past the time that breakdown occurs. No evidence for this kind of breakdown and re-formation of structure was found in our initial study (Udgaonkar & Baldwin, 1988) when the labeling pulse extended throughout the folding process.

Control Experiments. It is important to demonstrate that in refolding conditions where folding intermediates should be undetectable because they are unstable, application of an exchange pulse does in fact label completely all protein molecules. This basic control experiment has now been made (Figure 2). Native ribonuclease A remains fully folded in 2 M GdmCl at pH 4, 10 °C, but kinetic intermediates, including I_N, are greatly destabilized under these folding conditions (Schmid, 1983), as indicated by the observation that absorbance-detected kinetics and fluorescence-detected kinetics can no longer be distinguished. As expected, this destabilization of kinetic intermediates leads to a drastic decrease in the folding rates of both the fast-folding species, U_F, and the slow-folding species, U_S, compared to folding in stabilizing conditions. The relaxation time for the folding of U_F is increased from 0.04 s in 0.25 M GdmCl to 5 s in 2 M GdmCl (Schmid, 1983). Thus, 400 ms after initiation of folding in 2 M GdmCl, a negligible fraction of the unfolded protein molecules should fold, and a 37 ms exchange pulse applied then should completely label all peptide NH sites. With the possible exception of the V47 and V108 peptide NH protons, for which approximately 85−90% labeling is observed, all other amide proton sites show 100% proton occupancy. The exchange rate constants measured in 2 M GdmCl are discussed separately below.

A second important control experiment is to check whether the folding kinetics are dependent on the concentration of the protein used. It has been reported (Hall & Frieden, 1989) that a large C-terminal fragment of dihydrofolate reductase (DHFR) slows down the folding of DHFR. Also, at high concentrations, many proteins are known to aggregate either in the unfolded state or during the actual folding process, e.g., bovine growth hormone (Havel et al., 1988). Although no such aggregation was detected in the course of our experiments with RNase A, it is nevertheless important to demonstrate concentration independence, because the kinetic curves for acquisition of protection from exchange are directly compared to the kinetic curve for the formation of the native-like intermediate I_N, and an approximately 5-fold higher concentration of protein is used in the determination of the former than is used in the determination of the latter (see Experimental Procedures). Thus, a direct comparison of the results of the exchange measurements and the optical measurements is dependent on both results being independent of protein concentration. The kinetics monitored by tyrosine absorbance are limited to protein concentrations below 15 mg/mL, however, and the exchange measurements are generally limited to concentrations 4- or 5-fold higher because of the need to obtain sufficient amounts of protein for a good NMR spectrum. To overcome this problem of nonoverlapping concentration ranges for the two types of measurements, it was decided to carry out a $^1H-^3H$ pulse-labeling experiment, which can cover both concentration ranges. Identical levels of labeling were obtained for both concentration ranges (see Results), indicating that protein concentration was not a relevant factor; moreover, the extent of labeling (75%) observed in the $^1H-^3H$ labeling experiment was what was predicted from the corresponding $^2H-^1H$ labeling experiment.

It is also necessary, for a correct interpretation of pulse-labeling experiments, to demonstrate that native protein is not labeled by an exchange pulse, at all peptide NH sites under study. Results from tritium-labeling experiments had previously shown that at pH values greater than 9, exchange pulses greater than 10 s in duration are capable of labeling even native RNase A (Brems & Baldwin, 1986), with the extent of labeling being dependent on both the length and the pH of the exchange pulse. This result was also confirmed in the first study in which $^2H-^1H$ exchange was used to study folding (Udgaonkar & Baldwin, 1988), where it was noticed that a 10 s exchange pulse at pH 9 labeled 3 out of the 40 peptide NH sites that are stable to exchange at pH 4 in native RNase A. After this first report, the methodology was changed so as to be able to use much shorter exchange pulses, typically 37 ms (Udgaonkar & Baldwin, 1990). This in turn made it possible to extend the range over which the pH of the exchange pulse could be varied up to pH 11. Only a

few peptide NH protons in native RNase A are labeled by 37 ms exchange pulses, even at pH 11 (see Results). The observation that the H12, M13 and E49 peptide NH protons are not labeled in N, even by an exchange pulse of pH 11, implies that these protons, which are very weakly protected (1−10-fold) in I_1, are strongly protected (>1000-fold) in N, by the criteria used previously (Udgaonkar & Baldwin, 1990) (see, however, the discussion below). In the context of the stability of native RNase to exchange at high (greater than 9) pH, it is important to remember that equilibrium unfolding studies indicate that native RNase A is more stable to denaturation by GdmCl at pH 9 than it is at pH 4 (Pace et al., 1990).

The capability to apply exchange pulses that are short in duration compared to the time of application of the pulse makes it possible to check whether the extent of labeling I_1 is dependent on the duration of the pulse. A 37 ms pulse, at a pH value greater than 9, will label all unstructured molecules, but 37 ms may not be enough to label peptide NH sites in I_1. If peptide NH sites in I_1 are being labeled by the EX1 mechanism (Englander & Kallenbach, 1984), in which the rate of exchange into I_1 is given by the rate of opening I_1 to exchange, then the extent of labeling of peptide NH sites in I_1 is expected to increase with an increase in the length of the exchange pulse, but not with an increase in the pH of the pulse. The results in Figure 4 clearly show that the extent of labeling at 400 ms does not increase when the pulse-length is increased from 37 ms to 66 ms, even though the pH of the pulse is such (pH 10) that exchange of a free peptide NH proton would be complete within 0.2 ms. Therefore, the length of the exchange pulse cannot be limiting the extent of exchange, provided that exchange is occurring in the time range studied here, and the failure to label by exchange the β-sheet probes in I_1 must be ascribed instead to the large protection factors found when protection is computed for the EX2 (base-catalyzed) mechanism.

2′CMP Binding by I_1. Optical stopped-flow experiments (Figure 1) show that 2′CMP is bound well before I_N is formed during refolding at pH 6.0, 10 °C, in these refolding conditions. The experiment is made at pH 6.0, rather than pH 4.0 used for analyzing the folding pathway, because 2′CMP binding is strongly pH-dependent, with an optimum near pH 6 (Anderson et al., 1968). The presence of a 2′CMP binding site is clear evidence of a well-developed tertiary structure in I_1. Earlier RNase A refolding experiments, in other conditions, indicate that 2′CMP binding occurs with the same kinetics as burial of tyrosine side chains, as I_N is formed (Schmid & Blaschek, 1981). Their refolding conditions did not include the stabilizing salt Na_2SO_4, in contrast to the 0.4 M Na_2SO_4 used here, and this difference probably explains the contrast between their results and ours.

Comparison between Protection Factors in I_1 and I_N. pH profiles of exchange have been measured at 4 s after the start of folding, when I_N is the major species populated according to optical stopped-flow data. Because the proton intensities were low, the exchange results could not be evaluated for the less intense cross-peaks. The results (Figure 4) are compared with the pH profiles of exchange in I_1, at 0.4 s after the start of folding, given earlier (Udgaonkar & Baldwin, 1990). We reported earlier that about 85% of RNase A is highly protected against exchange at this time, as expected from optical stopped-flow data, and indicating that 85% of the protein has formed either I_N or N. The present results show that peptide NH protons H12 and M13 in helix 1, which are only weakly protected (1−10-fold) in I_1 (Udgaonkar & Baldwin, 1990), are strongly protected in I_N (data for H12 are not shown). The results show further that Lys 31, Val 43, and Ser 59, which are only moderately protected and show measurable exchange in I_1, also show measurable exchange in I_N between pH 10 and 11 (Figure 3). This result shows that, in part, the structures of I_1 and I_N have similar stabilities to exchange. The results for the other moderately protected protons of I_1 could not be evaluated for I_N because of low proton intensities in this experiment. The β-sheet probes Val 63, Thr 100, Ile 106, and Val 116, which are highly protected in I_1 (Udgaonkar & Baldwin, 1990), are also highly protected in I_N (Figure 3). In the case of these probes, the similarity between the behavior of I_1 and I_N orginates in the surprising fact that almost all β-sheet probes are highly protected in the early intermediate I_1. Met 13 shows a higher base line level of exchange than the other probes in Figure 3. This effect may be caused by some anomaly in processing the data for His 12 and Met 13, which have relatively low intensities. If so, this could affect the classification of His 12 and Met 13 as being weakly protected in I_1.

Nature of the Early Intermediate I_1. These results suggest that I_I is not a typical molten globule intermediate but rather has some fixed side chain structure. First, reinvestigation of the protection factors confirms that most β-sheet probes have protection factors that are too large to measure (>1000). This behavior is unusual for a molten globule species, and it shows that I_1 has a very stable structure. Second, those probes that have protection factors in a measurable range have very similar protection factors in I_1 and I_N, except for His 12 and Met 13 (see above). This observation suggests that the stability of I_1 is not very different from that of I_N, which is known to be highly structured. I_N has RNase catalytic activity (Schmid & Blaschnek, 1981). Third, I_1 probably binds the specific inhibitor 2′CMP: at any rate, 2′CMP binding occurs well before I_N is formed.

A major difference between I_1 and I_N is that the buried tyrosine side chains of RNase A (Tyr 25, 73, and 97; Wlodawer & Sjölin, 1983) are buried in I_N but not in I_1, according to absorbance measurements. Figure 6 shows that 2′CMP is bound at the center of RNase A whereas the buried tyrosine residues are found at either end. A similar observation concerning separate kinetic steps for 2′CMP binding and tyrosine side chain burial has been made for RNase S (Laurents et al., 1993). The two buried tyrosine residues of RNase S, Tyr 73 and Tyr 97, are buried only as the native protein is formed, when proline isomerization occurs. Nevertheless, there is an earlier folding intermediate that is highly structured according to the 1D NMR spectrum of the four histidine residues, and this intermediate binds the specific inhibitor 2′CMP.

Protection Factors Early in Refolding before I_1 Is Formed. Exchange rates of individual peptide NH protons have been measured in unfolded RNase A in refolding conditions by two approaches: (1) by taking pH profiles of exchange at 0.4 s after the start of folding in 2 M GdmCl, before any detectable folding has occurred, and (2) by examining the pH profiles of exchange in the 40% of the protein that remains unfolded at 0.4 s after the start of folding in standard conditions, when the kinetic curve for the formation of I_1 has leveled off. Representative pH profiles are given in

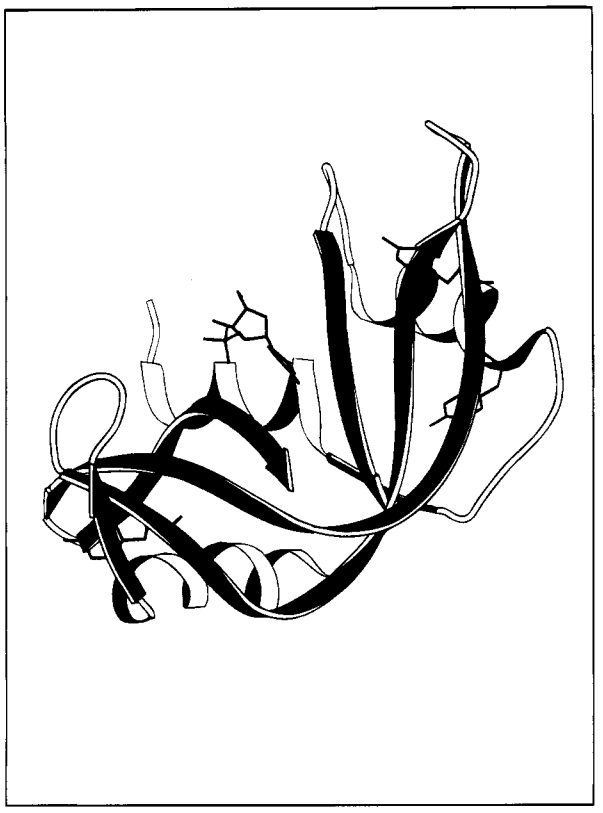

FIGURE 6: Ribbon diagram of RNase A, showing the location at which the specific inhibitor 2′CMP is bound (center, 10 o'clock) and the locations of the buried tyrosine residues 25 (3 o'clock), 73 (8 o'clock), and 97 (1 o'clock). The protein Data Base file containing the structure of RNase A complexed with 2′CMP is 1 ROB. The ribbon diagram was drawn using Molscript (Kraulis, 1991).

Figure 2 and Figure 5, respectively, and the exchange rate constants at pH 9 are given in Table 1, where they are compared to predicted values for model peptides (Bai et al., 1993). Some of the experimental pH profiles show unsatisfactory fits to a theoretical curve for base-catalyzed exchange of a single class of exchanging protons (see Val 54 and Ile 106 in Figure 2).

In the standard folding conditions used here, seven of the rate constants measured for nine protons agree within a factor of 2 with the model compound predictions from Bai et al. (1993), and the protection factors do not differ significantly from 1. Thus, there is no evidence from these results for stable secondary structure in the fraction of RNase A that fails to form I_1 initially. The problem of a transient kinetic barrier to the initial folding of part of the unfolded protein remains a mystery. We speculated previously (Udgaonkar & Baldwin, 1990) that the transient presence of a *cis* non-prolyl peptide bond, distributed at random positions in different unfolded molecules, might explain the transient barrier to folding in the $U_S II \rightarrow I_1$ step. Recently Odefey et al. (1994) measured the unfolding−refolding kinetics of a mutant RNase T1 containing a *cis* non-prolyl peptide bond. They find that *cis*−*trans* isomerization is considerably faster for a non-prolyl bond than for a prolyl peptide bond. Thus, the effect is probably in the right time range to present a transient barrier to the $U_S II \rightarrow I_1$ step. The next question is whether the frequency of occurrence of *cis* non-prolyl peptide bonds is high enough (e.g., about 1%) to explain the refolding results.

During refolding in 2 M GdmCl, the exchange rates of several NH protons (chiefly Val and Ile residues) are much faster than predicted from peptide data taken in the absence of GdmCl. Loftus et al. (1986) report an increase in the base-catalyzed exchange rate of poly(DL-alanine) in GdmCl, but the increase is less than 2-fold at 2 M GdmCl. The reason for the curious behavior reported here is not known, but it seems possible that we are observing general base catalysis of exchange by Gdm^+. General base catalysis of exchange has been observed in dioxane−H_2O mixtures by Klotz and Frank (1965).

Applications of the pH Profile Method. The pH profile method (Udgaonkar & Baldwin, 1990; Elöve & Roder, 1991) was introduced as a method for determining protection factors when a folding pathway is investigated by pulsed hydrogen exchange. In order to analyze folding intermediates at consecutive stages in the folding process, the pulse lengths are limited by necessity to very short times, and pH then provides the variable needed to obtain varying extents of exchange. The pH profile method is limited to the case when exchange occurs by the EX2 (base-catalyzed) mechanism. Use of the pH profile method has shown a second important application: when two separate folding intermediates are present that have different protection factors, the pH profile method is capable of resolving them in favorable cases. The protection factors of a given proton in two different folding intermediates should differ by 10-fold, in order to resolve the pH profiles of the two intermediates, and at least one protection factor must lie in the measurable range for this method to work. Miranker et al. (1993) have shown recently that combining mass spectrophotometric analysis with NMR analysis of pulsed hydrogen exchange is able to resolve two different folding intermediates. The pH profile method can provide a useful cross-check of the results.

ACKNOWLEDGMENT

We thank Carol Rohl, Doug Laurents, Andrew Robertson, and David Schultz for help and useful discussions, Virginia Robbins for purifying the protein, and Franz Schmid for sending us the manuscript by Odefey et al. (1994).

REFERENCES

Anderson, D. G., Hammes, G. G., & Walz, F. G., Jr. (1968) *Biochemistry 7*, 1637−1645.
Aue, W. P., Bartoldi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229−2246.
Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
Baldwin, R. L. (1991) *CIBA Found. Symp. 161*, 190−205.
Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84−91.
Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953−4963.
Brems, D. N., & Baldwin, R. L. (1984) *J. Mol. Biol. 180*, 1141−1156.
Brems, D. N., & Baldwin, R. L. (1985) *Biochemistry 24*, 1689−1693.
Cash, D. J., & Hess, G. P. (1981) *Anal. Biochem. 112*, 39−51.
Connelly, G. P., Bai, Y., Jeng, M.-F., & Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 87−92.
Cook, K. H., Schmid, F. X., & Baldwin, R. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157−6161.
Dobson, C. M., Evans, P. A., & Radford, S. E. (1994) *Trends Biochem. Sci. 19*, 31−37.
Elöve, G. A., & Roder, H. (1991) *Am. Chem. Soc. Symp. 470*, 50−63.

Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521−655.
Englander, S. W., & Mayne, L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 243−265.
Garel, J.-R. (1976) *Eur. J. Biochem. 70*, 179−189.
Garel, J.-R., & Baldwin, R. L. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3347−335.
Garel, J.-R., Nall, B. T., & Baldwin, R. L. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1853−1857.
Hall, J. G., & Frieden, C. (1989) *Proc Natl. Acad. Sci. U.S.A. 86*, 3060−3064.
Havel, H. A., Kaufman, E. W., Plaisted, S. M., & Brems, D. N. (1986) *Biochemistry 25*, 6533−6538.
Houry, W. A., Rothwarf, D. M., & Scheraga, H. A. (1994) *Biochemistry 33*, 2516−2530.
Jentoft, N., & Dearborn, D. G. (1979) *J. Biol. Chem. 254*, 4359−4365.
Kim, P. S., & Baldwin, R. L. (1980) *Biochemistry 19*, 6124−6129.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
Klotz, I. M., & Frank, B. H. (1965) *J. Am. Chem. Soc. 87*, 2721−2728.
Kraulis, J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Labhardt, A. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7674−7678.
Laurents, D. V., Doig, A. J., Schultz, D. A., & Baldwin, R. L. (1993) *Proc. R. Soc. London, A 345*, 131−140.
Lin, L.-N., & Brandts, J. F. (1983) *Biochemistry 22*, 564−573.
Loftus, D., Gbenle, G. O., Kim, P. S., & Baldwin, R. L. (1986) *Biochemistry 25*, 1428−1436.
Molday, R. S., Englander, S. W., & Kallen, R. G. (1972) *Biochemistry 11*, 150−158.
Montelione, G. T., & Scheraga, H. A. (1989) *Acc. Chem. Res. 22*, 70−76.
Odefey, C., Mayr, L. M., & Schmid, F. X. (1994) *J. Mol. Biol.* (in press).
Pace, C. N., Laurents, D. V., & Thomson, J. A. (1990) *Biochemistry 29*, 2564−2572.
Ptitsyn, O. B. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 243−300, W. H. Freeman and Co., New York.
Ptitsyn, O. B., & Rashin, A. A. (1975) *Biophys. Chem. 3*, 1−20.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700−704.
Schmid, F. X. (1981) *Eur. J. Biochem. 114*, 105−109.
Schmid, F. X. (1982) *Eur. J. Biochem. 128*, 77−80.
Schmid, F. X. (1983) *Biochemistry 22*, 4690−4696.
Schmid, F. X., & Baldwin, R. L. (1979) *J. Mol. Biol. 135*, 199−215.
Schmid, F. X., & Blaschek, H. (1981) *Eur. J. Biochem. 114*, 111−117.
Schmid, F. X., Grafl, R., Wrba, A., & Beintema, J. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 872−876.
Schreier, A. A. (1976) *Anal. Biochem. 83*, 178−184.
Schultz, D. A., Schmid, F. X., & Baldwin, R. L. (1992) *Protein Sci. 1*, 917−924.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694−699.
Udgaonkar, J. B., & Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8197−8201.
Wagner, G. (1983) *Q. Rev. Biophys. 16*, 1−57.
Wlodawer, A., & Sjölin, L (1983) *Biochemistry 22*, 2720−2728.

BI9421937