Strategy for Trapping Intermediates in the Folding of Ribonuclease and for Using ^{1}H-NMR to Determine Their Structures

KUNIHIRO KUWAJIMA, PETER S. KIM, and ROBERT L. BALDWIN, *Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305*

Synopsis

The major unfolded form of ribonuclease A is known to show well-populated structural intermediates transiently during folding at 0°–10°C. We describe here how the exchange reaction between D_2O and peptide NH protons can be used to trap folding intermediates. The protons protected from exchange during folding can be characterized by ^{1}H-nmr after folding is complete. The feasibility of using ^{1}H-nmr to resolve a set of protected peptide protons is demonstrated by using a specially prepared sample of ribonuclease S in D_2O in which only the peptide protons of residues 7–14 are in the ^{1}H-form. All eight of these protected peptide protons are H-bonded. Resonance assignments made on isolated peptides containing these residues have been used to identify the protected protons. Other sets of protected protons trapped in the ^{1}H-form can also be isolated by differential exchange, using either ribonuclease A or S. Earlier model compound studies have indicated that H-bonded folding intermediates should be unstable in water unless stabilized by additional interactions. Nevertheless, peptides derived from ribonuclease A that contain residues 3–13 do show partial helix formation in water at low temperatures. We discuss the possibility that specific interactions between side chains can stabilize short α-helixes by nucleating the helix, and that specific interactions may also define the helix boundaries at early stages in folding.

INTRODUCTION

The problem of how to determine the folding pathway of a protein has been a challenge of long standing. Detailed information on the structures of proteins has come chiefly from x-ray crystallography, but no one has yet determined how to adapt x-ray crystallography to the study of folding intermediates. Nmr spectra of proteins in solution contain a wealth of structural information, and the spectra can be correlated with x-ray structures. Recently it has been suggested that two-dimensional (2D) nmr techniques may even be capable of determining the spatial structures of small proteins *ab initio.*[1]

When refolding is coupled to the reoxidation of S—S bonds, the S—S intermediates can be trapped covalently and the locations of the S—S bonds can be found, as shown by studies of bovine pancreatic trypsin inhibitor (BPTI) and of ribonuclease A (RNase A).[2,3] These are kinetic intermediates that are practically not detectable at equilibrium. The same sit-

Biopolymers, Vol. 22, 59–67 (1983)
© 1983 John Wiley & Sons, Inc. CCC 0006-3525/83/010059-09$01.90

uation prevails in studies of the folding of typical small proteins with S—S bonds left intact: equilibrium intermediates are present only at very low levels, but kinetic intermediates are well populated for short times during folding. One possible reason is that kinetic intermediates can accumulate transiently without being populated at equilibrium. Another reason that certainly applies is that kinetic intermediates can be studied outside the equilibrium transition zone: in "strongly native" folding conditions, kinetic intermediates can be stable relative to the unfolded protein while being unstable relative to the final product of folding, the native protein. The study of kinetic intermediates introduces two new problems, not found with equilibrium intermediates: (i) there are multiple forms of an unfolded protein, consisting of both fast-folding (U_F) and slow-folding (U_S) species, and the folding pathway of each species must be determined separately; (ii) kinetic intermediates are populated only for short times, at most a few seconds, whereas high-resolution nmr methods require much longer times.

Earlier work has shown that the major unfolded form of RNase A is suitable for studying folding intermediates.[4–8] It is a slow-folding species (U_S^{II})[8] that forms well-populated structural intermediates during folding at 0–10°C. The following section summarizes what is known about the folding of U_S^{II}.

In this paper we describe a rapid method for trapping structural intermediates in such a way that the structures present at the time of trapping can be determined later, after folding is complete. The principle is to use the exchange reaction between the amide protons of the peptide bonds and water (H_2O or D_2O). Exchange from those peptide protons that are fully exposed to D_2O occurs rapidly, relative to the folding of U_S^{II} at neutral pH, while other peptide protons that are protected by structure in a folding intermediate are trapped in the ^1H-form before exchange can occur.

REFOLDING BEHAVIOR OF THE MAJOR UNFOLDED FORM OF RNase A

Three species of unfolded RNase A have been found: the fast-folding species U_F and two slow-folding species, U_S^I and U_S^{II}.[8] The presence of both U_F and U_S species was first detected by stopped-flow experiments on the binding of a specific inhibitor (2′-CMP) during folding[9]: folded species able to bind 2′-CMP are formed in a fast reaction ($U_F \rightarrow N$) as well as in a slow reaction ($U_S \rightarrow N$). In strongly native folding conditions, the U_S^I and U_S^{II} species fold at different rates, which allows their relative concentrations to be measured (80:20 U_S^{II}:U_S^I).[8] A precise method of determining the U_F:U_S ratio has been worked out recently, and the ratio has been found to be 18:82, independent of temperature.[10] Thus, U_S^{II} is the major unfolded species of RNase A.

The evidence is good that the *cis-trans* isomerization of proline residues after unfolding[11] accounts for the U_F and U_S species of unfolded RNase A. Refolding assays for U_S as a function of time after unfolding ("double-jump experiments") show that U_F is the initial product of unfolding ($N \rightarrow U_F$) and that U_S is formed slowly from U_F.[11,12] At low temperatures almost 100% U_F can be obtained transiently.[5] In unfolded RNase A, the $U_F \rightleftharpoons U_S$ reaction has the kinetic properties characteristic of proline isomerization: it is catalyzed by strong acid (>5*M* $HClO_4$)[13] and shows the large activation enthalpy of 21 kcal/mol.[13] Moreover, its rate is independent of the concentration of the denaturant GuHCl (guanidinium chloride),[14] which indicates that unfolding in the sense of breaking structure is not a property of the $U_F \rightleftharpoons U_S$ reaction. Protease digestion experiments (Lin and Brandts, to be published) correlate the formation of one U_S species after unfolding chiefly with the *cis* → *trans* isomerization of Pro 93. This proline residue is *cis* in native RNase A. The role of the other three proline residues of RNase A in producing U_S species after unfolding is still under study.

Proline isomerization is the last known step in the folding of U_S^{II} at 0–10°C, as judged by an assay for RNase molecules that contain a wrong proline residue.[5] The assay makes use of rapid unfolding in 5.1*M* GuHCl, pH 1.9, at 0°C, where proline isomerization is slow, followed by a refolding step to determine the ratio of U_F:U_S in the unfolded sample. Refolding of U_S^{II} at 0–10°C yields a nativelike intermediate I_N that is able to bind 2′-CMP[5,8] and is enzymatically active[8] but still contains a wrong proline isomer.[5]

STRATEGY FOR TRAPPING AND NMR CHARACTERIZATION OF FOLDING INTERMEDIATES

Because nmr experiments take a long time compared to events in folding, it is necessary to adopt one of the following strategies in order to use nmr to characterize folding intermediates. (1) Find unusual folding conditions, or find an atypical small protein, in which equilibrium folding intermediates are well populated. (2) Find kinetic folding conditions in which transient intermediates are populated for long times (e.g., at subzero temperatures) (Biringer and Fink, to be published). (3) Find a method of trapping kinetic intermediates in such a way that their structures can be determined later, after folding is complete. We follow the third route here. It is possible to describe the H-bonded structures of the intermediates at successive stages in folding provided that H-bonding is the major determinant in inhibiting peptide proton exchange in the folding intermediate. In any case, trapping will show what parts of the protein molecule are involved in structure that retards peptide proton exchange.

Evidence has been presented[4,6] that there are one or more early intermediates in the folding of U_S^{II} that protect several peptide protons against exchange with solvent. In the earlier experiment,[4] the exchangeable protons of the unfolded protein (U_S form) were ^3H-labeled; then a competition between exchange-out and folding was allowed to occur once refolding was initiated. The later experiment[6] introduced a pulse-labeling method that could be used at any time during folding. A short (10-s) pulse of (^3H)-H_2O was given at different times, and the pulse was quenched at pH 3 while folding continued. The pulse-label method measures exclusion of ^3H label from protected peptide protons.

The considerations governing the exchange method are as follows. Exchange of amide protons is both acid- and base-catalyzed, with a minimum rate near pH 3. Exchange may be rapid (e.g., 0.1 s, 10°C, pH 7.5) or slow ($>10^3$ s, 0°C, pH 3), depending on pH. This makes it feasible either to label rapidly a transient folding intermediate or to stop exchange. Since amide proton exchange occurs by standard proton-transfer mechanisms (addition or abstraction of a proton at the peptide NH group),[16] exchange of H-bonded protons necessitates breaking these H-bonds. Also, water must penetrate to the peptide NH groups in order for exchange to occur. These two mechanisms of preventing exchange, namely, H-bonding and limiting access of solvent, are thought to be chiefly responsible for the enormous reduction in exchange rate ($>10^9$) for some peptide protons in native proteins.[16,17]

To turn this trapping procedure into a method for characterizing the structures of folding intermediates, it is necessary to change from H_2O to D_2O during folding, then allow folding to go to completion and to analyze the ^1H-nmr spectra of the peptide protons protected during folding. There are two stages in the analysis: (i) resolution of trapped proton resonances and (ii) assignment of these resonances. For proteins as small as BPTI (58 residues), 2D nmr methods at 500 MHz can be used both to resolve and to assign the resonances.[18] These methods may also prove to be successful with larger proteins such as RNase A. Meanwhile, we show here that 1D nmr is capable of characterizing trapped peptide protons, as judged by results for a model system.

Figure 1 shows the resonance lines of eight peptide protons belonging to residues 7–14 of RNase S: seven resonances are resolved in Fig. 1 and the eight can be resolved at pH 6. This spectrum, in which only these eight peptide protons are in the ^1H-form, was obtained as follows (Kuwajima and Baldwin, to be published). Freeze-dried S-peptide (residues 1–20) in the ^1H-form was added at 4°C, pH 3.0 to 30% mol excess S-protein (residues 21–124) in the ^2H-form in D_2O, 0.1*M* NaCl; the final S-peptide concentration was 0.1 m*M*. In these conditions, combination between S-peptide and S-protein occurs rapidly, before there is substantial exchange between D_2O and peptide protons of the S-peptide.[19] The pH of the solution was immediately adjusted to pH 5.2, 4°C, and exchange was allowed to occur for 5 h. The half-time for exchange-out of most of these eight protons is

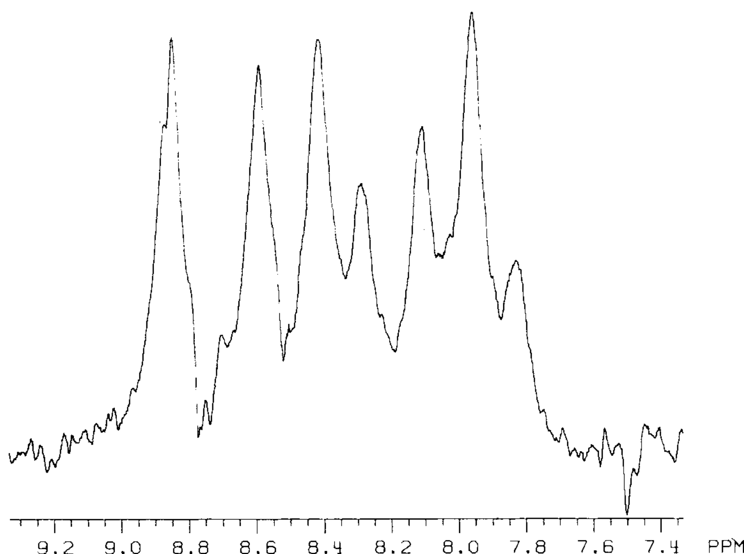

Fig. 1. An ^{1}H-nmr difference spectrum of the peptide protons of residues 7–14 of RNase S in D_2O at 20°C, pH 4.57, 2 mM. The sample was prepared by allowing S-peptide (^{1}H-form) to combine with S-protein (^{2}H-form) in D_2O, followed by 5-h exchange-out at pH 5.2, 0.1M NaCl, 4°C (see text). The spectrum taken after complete exchange-out of the peptide protons was subtracted from the spectrum taken after 5-h exchange to give this difference spectrum. See text for the procedure used to identify the protons protected against exchange.

more than one day in these conditions. The sample was then concentrated by ultrafiltration, the pH was adjusted to pH 4.6, and the spectrum shown in Fig. 1 was taken at 20°C.

The identity of the protected protons was found (Kuwajima and Baldwin, to be published) by dissociating RNase S. In identifying the protected protons, S-peptide (residues 1–20) was replaced with the simpler peptide 1–15: the same eight peptide protons are protected in both, and peptide 1–15 binds almost as tightly to S-protein as does peptide 1–20.[20] Dissociation of S-protein from peptide 1–15 occurs at or above 3M urea at −4°C, pH 2.3, in D_2O: these are conditions in which exchange between amide protons of peptide 1–15 and D_2O is slow enough to get a satisfactory nmr spectrum. The resonance assignments of free peptide (1–15) are known from spin-decoupling experiments and from studies of shorter peptides (Kuwajima and Baldwin, to be published). Differences between the exchange rates of these eight protons have been used to complete the assignments of the resonances when peptide 1–15 is bound to S-protein.

What can we conclude from this study of a model system about using nmr to characterize the structures of trapped folding intermediates? First, it is feasible to use 1D ^{1}H-nmr at 360 MHz to resolve about eight amide protons at a time for a protein the size of RNase A. By using differential

exchange,[21] the process can be repeated with a different set of protected protons. For example, S-peptide bound to S-protein also shows an intermediate set of less protected protons that can be resolved and analyzed. Second, the results show that the highly protected protons are all H-bonded protons, which supports the hypothesis that the H-bonded structures of folding intermediates can be studied in this way. Third, the fact that the H-bonded protons of the 3–13 α-helix are all highly protected protons means that it should be possible to determine when this helix is formed during folding (i.e., when all of these protons become protected). Later steps in folding that increase the degree of protection of this set of protons can also be detected. Finally, the fact that the peptide proton of Asp 14, which is H-bonded to the peptide C=O group of Val 47, is also a highly protected proton means that we should be able to determine whether the 3–13 helix is formed before or at the time that it is H-bonded to the β-sheet.

Preliminary experiments (Bierzynski, Kallenbach, Kim, and Baldwin, unpublished) indicate that about 30 highly protected protons of RNase A can be resolved without difficulty by using differential exchange to divide them into classes. It should be feasible to "track" these protected protons during folding, by measuring the degree of protection for each proton at different stages in folding. The problem of assigning these resonances is still under study. A neutron diffraction study of partially exchanged RNase[22] gives the locations of 28 partially protected protons and demonstrates the possibility of assigning protected protons in RNase A by correlating ^{1}H-nmr and neutron-diffraction studies.

HELIX FORMATION BY PEPTIDE FRAGMENTS OF RNase A

The strategy discussed here for characterizing the structures of folding intermediates is essentially a strategy for determining their H-bonded structures, as judged by recent evidence[22–24] that H-bonding is the major factor in retarding H exchange in native proteins. Neutron-diffraction studies of partially exchanged RNase A[22] and of trypsin[23] support this conclusion, and so does a 2D ^{1}H-nmr study of exchange rates in solution for BPTI.[24] As mentioned above, ^{3}H-exchange experiments[4,6] have shown that numerous peptide protons in RNase A are protected early in the folding of U_S^{II} at 0–10°C.

These studies[4,6] raise the question of how it is possible for H-bonded structures to be stable in water. Model-compound studies[25,26] indicate that the peptide H-bond has little stability in water because of competing H-bonds to water, and studies of α-helix formation in random copolymers by the host–guest technique[27] indicate that short α-helices (<20 residues) are always unstable in water, regardless of temperature and composition, when random copolymers are studied.[28] Consequently, H-bonded folding

intermediates should be stable only if additional interactions are present. It is important for understanding the mechanism of folding to find out whether these additional interactions must be tertiary interactions.

Recent studies (Ref. 28; Bierzynski and Baldwin, to be published; and Kim, Bierzynski, and Baldwin, to be published) of α-helix formation in water by peptides containing residues 3–13 of RNase A show that interactions between side chains within the helix can stabilize a short α-helix. These studies confirm the earlier report[29] that peptide 1–13 (lactone) does show partial helix formation at low temperatures (0–20°C) and that the helix is formed intramolecularly. The evidence for one or more stabilizing interactions between side chains is as follows. Peptide 1–13 (lactone) shows CD spectra characteristic of partial α-helix content near 0°C and nmr studies indicate that all, or nearly all, residues participate in the postulated helix-forming reaction.[28] The helix can be melted out by GuHCl (Bierzynski and Baldwin, to be published) or by urea (Kuwajima, unpublished), as well as by increasing temperature. Surprisingly, helix stability is strongly dependent on pH: it follows a bell-shaped curve with an optimum near pH 5.[28] The bell-shaped curve can be fitted by pKs, indicating that protonated His 12 and also a deprotonated glutamate residue (Glu 2 or Glu 9) are needed for optimal stability.

The C-terminal residue of peptide 1–13 is homoserine-13 lactone, which is formed by cyanogen bromide cleavage of RNase A at Met 13. The lactone is easily converted to the α-carboxylic acid by mild alkaline hydrolysis and peptide 1–13 (carboxylate) has also been prepared and studied. It fails to show α-helix formation in conditions where peptide 1–13 (lactone) does form a helix (Kim, Bierzynski, and Baldwin, to be published). Moreover, peptide 1–13 (carboxylate) contains the H-bonded salt bridge His $12^{+}\cdots$HSer 13 COO^{-}, as shown by pH titration of the peptide protons in this and simpler peptides (Kim, Bierzynski, and Baldwin, to be published). The lactone ring of HSer 13 lactone is not needed for α-helix formation because peptide 1–20, which has neither the lactone ring nor an α-carboxylate group on residue 13, also shows pH- and temperature-dependent helix formation (Kim, Bierzynski, and Baldwin, to be published) resembling that of peptide 1–13 (lactone).

The simplest interpretation of these results is that a Glu $9^{-}\cdots$His 12^{+} salt bridge is present in the helix formed by peptide (1–13) (lactone) and stabilizes the helix.[28] Model building shows that the H-bonded salt bridge can be made when the helix is present and that it stabilizes one turn of the helix, thereby nucleating the helix. In peptide 1–13 (carboxylate) the competing salt bridge His $12^{+}\cdots$HSer 13 COO^{-} is assumed to be stronger (it closes a smaller ring) and to prevent helix formation because the Glu $9^{-}\cdots$His 12^{+} salt bridge is necessary for stable helix formation.

In one current model for folding, sequence determines secondary structure and secondary structure determines the tertiary fold of a protein. This implies either that α-helices and β-sheets are formed initially at specific locations in an otherwise unfolded polypeptide chain or else that

many such combinations preexist at the first stage in folding and that formation of the tertiary fold selects one of these. Studies of α-helix formation in model systems raise two unresolved problems concerning this mechanism. (1) What additional interactions can stabilize a short α-helix in water? We discuss above the possibility that a specific side-chain interaction might nucleate a helix. (2) What terminates the helix? The average length of helical segment at the transition midpoint is about $1/\sigma^{1/2}$ (Ref. 30) or 100 residues if $\sigma = 10^{-4}$, as indicated by studies of homopolymers and random copolymers.[30] Nevertheless, an ^{1}H-nmr study of helix formation in RNase A unfolded by GuHCl at low temperatures (Bierzynski and Baldwin, to be published) indicates that the 3–13 helix does not extend as far as His 48. The C2H resonances of His 12 and His 48 are used as probes of helix formation.

One possible explanation is that specific interactions also terminate a helix. The example of peptide 1–13 (carboxylate), in which a specific salt bridge apparently blocks helix formation, illustrates this possibility. Another possible explanation is that a side-chain interaction that nucleates a helix also decreases the average helical length by increasing σ. Then the presence of residues with low s values, such as glycine or proline, on either side of the helix might terminate it (s is the stability constant for helix formation). A striking result of the host–guest studies[27] is that the range in s values is small: there is only about a 3-fold difference between the residues that most strongly stabilize and those that destabilize a helix. This three-fold range can be compared to the 1000-fold greater helix stability of peptide 1–13 (lactone) found experimentally[28] as compared to prediction from host–guest results.

We are grateful for the discussion of Drs. A. Bierzynski and N. R. Kallenbach. K.K. is an International Fellow of the U.S. National Institutes of Health. P.S.K. is a predoctoral fellow of the PHS Medical Scientist Training Program. This research was supported by Grant 2 R01 GM 19988-22 from the National Institutes of Health and Grant PCM 77-16834 from the National Science Foundation. Use of the Stanford Magnetic Resonance Facility (supported by NSF Grant GP 23633 and NIH Grant RR 00711) is gratefully acknowledged.

References

1. Wüthrich, K., Wider, G., Wagner, G. & Braun, W. (1982) *J. Mol. Biol.* **155,** 311–319.
2. Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol.* **33,** 231–298.
3. Konishi, Y., Ooi, T. & Scheraga, H. A. (1981) *Biochemistry* **20,** 3945–3955.
4. Schmid, F. X. & Baldwin, R. L. (1979) *J. Mol. Biol.* **135,** 199–215.
5. Cook, K. H., Schmid, F. X. & Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 6157–6161.
6. Kim, P. S. & Baldwin, R. L. (1980) *Biochemistry* **19,** 6124–6129.
7. Schmid, F. X. (1981) *Eur. J. Biochem.* **114,** 105–109.
8. Schmid, F. X. & Blaschek, H. (1981) *Eur. J. Biochem.* **114,** 111–117.
9. Garel, J.-R. & Baldwin, R. L. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 3347–3351.
10. Schmid, F. X. (1982) *Biochemistry* (in press).
11. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975) *Biochemistry* **14,** 4953–4963.
12. Nall, B. T., Garel, J.-R. & Baldwin, R. L. (1978) *J. Mol. Biol.* **118,** 317–330.
13. Schmid, F. X. & Baldwin, R. L. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4764–4768.

14. Schmid, F. X. & Baldwin, R. L. (1979) *J. Mol. Biol.* **133,** 285–287.
15. Tsong, T. Y. & Baldwin, R. L. (1978) *Biopolymers* **17,** 1669–1678.
16. Englander, S. W., Donner, N. N. & Teitelbaum, H. (1972) *Annu. Rev. Biochem.* **41,** 903–924.
17. Woodward, C., Simon, I. & Tüchsen, E. (1982) *Mol. Cell. Biochem.* (in press).
18. Wagner, G. & Wüthrich, K. (1982) *J. Mol. Biol.* **155,** 347–366.
19. Schreier, A. A. & Baldwin, R. L. (1976) *J. Mol. Biol.* **105,** 409–426.
20. Niu, C.-H., Shindo, H., Matsuura, S. & Cohen, J. S. (1980) *J. Biol. Chem.* **255,** 2036–2038.
21. Englander, S. W., Calhoun, D. B., Englander, J. J., Kallenbach, N. R., Liem, L. K. H., Malin, E. L., Mandal, C. & Rogero, J. R. (1980) *Biophys. J.* **32,** 577–589.
22. Wlodawer, A. & Sjölin, L. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1418–1422.
23. Kossiakoff, A. A. (1982) *Nature* **296,** 713–721.
24. Wagner, G. & Wüthrich, K. (1982) *J. Mol. Biol.* **160,** 343–361.
25. Klotz, I. M. & Franzen, J. S. (1962) *J. Am. Chem. Soc.* **84,** 3461–3466.
26. Susi, H., Timasheff, S. N. & Ard, J. S. (1964) *J. Biol. Chem.* **239,** 3051–3054.
27. Scheraga, H. A. (1978) *Pure Appl. Chem.* **50,** 315–324.
28. Bierzynski, A., Kim, P. S. & Baldwin, R. L. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 2470–2474.
29. Brown, J. E. & Klee, W. A. (1971) *Biochemistry* **10,** 470–476.
30. Poland, D. & Scheraga, H. A. (1970) *Theory of Helix-Coil Transitions in Biopolymers*, Academic Press, New York, p. 85.

Received June 20, 1982
Accepted August 24, 1982