preceded by TIPs with a duration of 25% of the total interval. There were three false alarms.

A similar test was performed on an even wider suite of seismic zones with another version of the algorithm that differs from the above only in the values of certain parameters[28]. Twenty-three out of 29 strong earthquakes that occurred in these zones were preceded by TIPs occupying 25% of the total elapsed time.

For brevity, we bypass a detailed description[28] of the assignment of a regional TIP to one of the smaller areas within a region. The most frequently occurring trait is Q2, which is large in D features and small in N features. The number of exceptions to regularity of traits is significantly larger among the features in this case than among the diagnostics for regional prediction. Except for trait T2, there are no significant dissimilarities between the areas and the regions. The total space-time occupied by the 15 TIPs in the areas is now only 13%, but at the cost of two more failures-to-predict (the Imperial Valley earthquakes of 1942 and 1956). In some cases TIPs are diagnosed in two or three areas simultaneously instead of only one. Conclusions for areas are speculative at this time since we have not subjected the algorithm for areas to the tests of stability and of transfer to other parts of the world.

The statistical significance of the observations cannot be estimated because of the small sample size. Furthermore, we have not formalized methods for the choice of regionalization nor for the choices of many of the fixed parameters. We do not claim that the algorithm is optimal. But the success of the geographical transfers for the regional studies suggests that the set of traits we have identified is intrinsically relevant to the approach of a strong earthquake in diverse tectonic settings and that the algorithm deserves the decisive test of the prediction of future strong earthquakes.

Received 12 April; accepted 16 August 1988.

1. Gelfand, I. M. *et al. Phys. Earth planet. Inter.* **11,** 227-283 (1976).
2. Briggs, P., Press, F & Guberman, Sh.A. *Bull. geol. Soc. Am.* **88,** 161-173 (1977).
3. Caputo, M. *et al. Phys. Earth planet. Inter.* **21,** 305-320 (1980).
4. Gvishiani, A. D. & Soloviev, A. A. *Earthquake Prediction Res.* **2,** 237-243 (1984).
5. Jones-Cecil, M., Wheeler, R. L. & Dewey, J. D. *U.S. geol. Sur. Open File Rep. 81-195* (1981).
6. McCann, W. R., Nishenko, S. P., Sykes, L. R. & Krause, J. *PAGEOPH* **117,** 1082-1147 (1979).
7. Sykes, L. R. & Nishenko, S. P. *J. geophys. Res.* **89,** 5905-5927 (1984).
8. Jones, L. M. *Bull. seismol. Soc. Am.* **75,** 1669-1679 (1985).
9. Kagan, Y. Y. & Knopoff, L. *Science* **236,** 1563-1567 (1987).
10. Raleigh, C. B. *EOS* **58,** 236-272 (1977).
11. Mogi, K. *Earthquake Prediction* (Academic, Tokyo, 1985).
12. Rikitake, T. *Earthquake Prediction* (Elsevier, Amsterdam, 1976).
13. Molchan, G. M., Dmitrieva, O. E. & Rotvain, I. M. *Comput. Seismol.* **21** (in the press).
14. Keylis-Borok, V. I. & Malinovskaya, L. N. *J. geophys. Res.* **69,** 3019-3024 (1964).
15. Keilis-Borok, V. I., Knopoff, L., Rotvain, I. M. & Sidorenko, T. M. *J. geophys. Res.* **85,** 803-811 (1980).
16. Keilis-Borok, V. I., Knopoff, L. & Rotvain, I. M. *Nature* **283,** 259-263 (1980).
17. Sobolev, G. A. & Zavialov, A. D. *Dokl. Akad. Nauk SSSR* **252,** 69-71 (1980).
18. Kagan, Y. Y. & Knopoff, L. *J. geophys. Res.* **86,** 2853-2861 (1981).
19. Prozorov, A. G. *Comput. Seismol.* **14,** 20-26 (1983).
20. Jennings, C. W. *Calif. Div. Mines Geol.* Geologic Data Map No. 1 (1975).
21. Jennings, C. W. *Calif. Div. Mines Geol.* Geologic Map of California (1977).
22. Real, C. R., Toppozada, T. R. & Parke, D. L. *Calif. Div. Mines Geol.* Map Sheet 39 (1978).
23. Slemmons, D. B., Gimlett, J. I., Jones, A. E., Greensfelder, R. & Koenig, J. *Nevada Bur. Mines* Map 29 (1965).
24. Stover, C. W. *U.S. Geol. Sur.* Geophys. Inv. Map 984 (1986).
25. *Earthquakes in Southern California, 1932-1985* (Seismol. Lab., Calif. Inst. Tech., 1976-1986).
26. *World hypocenter data file 1932-1985* (US geol. Surv., natn Earthquake Inf. Serv., 1986).
27. Keilis-Borok, V. I., Lamoreaux, R., Johnson, C. & Minster, B. *Earthquake Prediction Res.* **1,** 135-151 (1982).
28. Gabrielov *et al.*, Preprint, Centre Regional de Sismologia para America del Sur, Lima.

NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A

Jayant B. Udgaonkar & Robert L. Baldwin

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA

The presence of an early intermediate on the folding pathway of ribonuclease A has been demonstrated by a study of the exchange reaction between the backbone amide protons in the folding protein and solvent protons using rapid mixing techniques. A structural analysis of the intermediate by two-dimensional ^{1}H-NMR is consistent with the framework model of protein folding in which stable secondary structure first forms the framework necessary for the subsequent formation of the complete tertiary structure.

To obtain detailed structural information about the folding pathway of a protein has been a challenge of long standing. The single-domain protein ribonuclease A (RNase A) has traditionally served as an archetypal model for small proteins in protein folding and its folding pathway has been extensively studied. Multiple unfolded forms, which differ in their refolding rates, have been identified[1]. There is strong evidence for the presence of both an early hydrogen-bonded intermediate[2,3] and a late native-like intermediate[4,5] on the folding pathway of the main unfolded species, but little is known about the structure or stability of these kinetic intermediates. Such information should provide the most useful evidence for establishing the folding pathway of RNase A.

Here we report the use of the exchange reaction of backbone amide protons with solvent protons to obtain detailed structural information about early events during the folding of RNase A. This experimental approach is also being used with another small protein, cytochrome *c* (ref. 6). In our experiments, exchange is initiated by rapid (millisecond) mixing of a solution of RNase A in D_2O with an H_2O buffer at different times after initiation of refolding of the unfolded protein. The accessibility of individual backbone amide protons to exchange with solvent protons at each timepoint is measured by two-dimensional ^{1}H-NMR after folding is complete; it is this measurement that allows a structural characterization of the folding process. Our results reveal that a stable structural intermediate is formed within the first 1.5 seconds of folding. The rapid-mixing technique makes it possible to monitor the kinetics of formation of this intermediate, in which a large number of backbone amide protons are protected against exchange. NMR techniques have allowed assignment of several of these amide protons to specific residues found in the β-sheet of native RNase A. Thus, the β-sheet may be formed in the first major folding reaction of RNase A. The preliminary results reported here support the

framework model of protein folding[7-9], in which the formation of stable secondary structure is a prerequisite to the formation of the final tertiary structure. The potential of the technique introduced here to provide more detailed structural information on protein-folding pathways is established.

Experimental strategy

Two-dimensional ^{1}H-NMR techniques[10] allow the resolution of nearly all the backbone amide proton resonances in RNase A; for example, in a COSY (ref. 11) spectrum of RNase A in water, nearly all of the 120 or so expected C_αH-NH crosspeaks are observed in the fingerprint region (data not shown). The NMR techniques also make it possible to assign each such crosspeak to a specific backbone amide proton in the amino-acid sequence[10]. Twenty such assignments were reported earlier for RNase A(ref. 12), and additional assignments have been made in this laboratory (A. D. Robertson, unpublished results). The fact that the fingerprint region of the COSY spectrum of RNase A in H_2O is fully resolved implies that it will be possible to obtain a complete sequential assignment of all the crosspeaks in this part of the spectrum without too much difficulty. Recently, essentially complete sequential ^{1}H-NMR assignments have been reported for several proteins that are comparable in size to RNase A, including cytochrome *c* (ref. 13), lysozyme[14] and thioredoxin[15].

Elegant studies of the folding pathway of disulphide-reduced bovine pancreas trypsin inhibitor (refs 16 and 17) have shown that it is possible to trap covalently and then purify intermediates containing different disulphide bonds, and to study these intermediates using NMR techniques[18]. In the case of protein-folding reactions in which the disulphide bonds are left intact or otherwise are absent, it is not possible to trap folding intermediates covalently. Thus, although two-dimensional ^{1}H-NMR techniques can yield detailed structural information about a native protein the size of RNase A (ref. 10), there is a basic problem in obtaining similar information about the folding pathway of the disulphide-intact protein: the NMR techniques are slow (several hours are required to record a spectrum), whereas folding intermediates are typically populated only for short times (a few seconds or less). To use the capabilities of the NMR techniques, it is necessary first to label the structure of the folding intermediate in such a way that it can be studied in the native protein after folding is complete[19-22]. The exchange reaction of the backbone amide protons with solvent protons can be exploited for this purpose if the folding intermediate contains stable secondary structure (such as α-helices and β-sheets). The exchange rates of some of the backbone amide protons in a fully folded protein are slower than those in the unfolded protein by factors of as much as 10^9, because these protons participate in hydrogen-bonded secondary structure in the fully folded protein[23,24]. Hydrogen-bonded structure directly inhibits exchange because hydrogen bonds must be broken for exchange to occur[25]. Accessibility to solvent is also required for exchange, and most highly protected amide protons in native proteins are both hydrogen-bonded and inaccessible to solvent. Thus at any time during the folding process, an amide proton may become resistant to exchange with solvent protons, either as a result of the involvement of that proton in secondary structure, or as a result of the seclusion of that proton from solvent, or both.

In the case of native RNase A, most of the 27 backbone amide protons that were protected against exchange with solvent deuterons over the one-year time period required for a neutron diffraction study[26,27] are involved in hydrogen-bonding. These 27 amide protons are also distributed throughout the structure of the protein molecule[26,27]. Under the conditions required to record the COSY NMR spectrum shown in Fig. 1, about 40 backbone amide protons are seen to be stable to exchange with solvent deuterons. Of these, 13 have been sequentially assigned

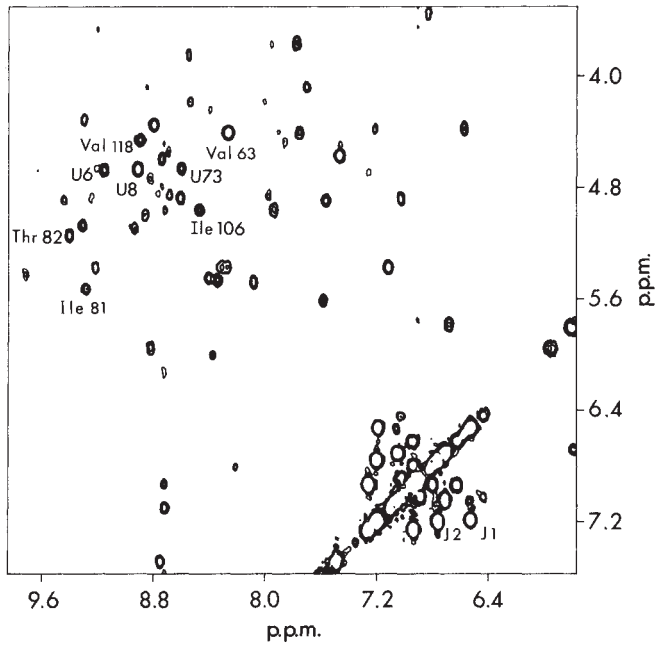

Fig. 1 Reference COSY spectrum of RNase A in D_2O, *p*H 4, 22 °C. Only the region showing C_α H-NH crosspeaks and aromatic ring-proton crosspeaks is depicted. The latter are shown at the lower right corner next to the diagonal of the spectrum. Only ~40 C_αH-NH crosspeaks are seen; the amide protons contributing to them are stable to exchange with solvent deuterons. Some of the sequential assignments for the C_αH-NH crosspeaks are shown. U6, U8 and U73 are C_αH-NH crosspeaks that have not yet been assigned to specific residues in the protein. The crosspeaks labelled J1 and J2 arise from completely non-exchangeable aromatic ring protons, and were chosen as internal intensity reference peaks (see Fig. 2 legend).

Methods. The preparation of the RNase A sample for the reference spectrum was as follows. RNase A (40 mg) were dissolved in 1.6 ml D_2O and 8.4 ml H_2O at *p*H 3 and the solution heated to 60 °C for 20 min; after lyophilization, the procedure was repeated to obtain protein that was 84% protonated at all amide hydrogen sites. The protein was then redissolved in 16% deuterated water at *p*H 3. The final NMR sample was prepared by ultrafiltration of the protein solution at *p*H 3, 5 °C; the solvent was completely deuterated by repeated concentration and then dilution of the sample in D_2O. The *p*H of the concentrated NMR sample (4 mM) in D_2O was raised to 4, and the sample was then kept at 5 °C for 20 h to allow exchange out of all protons labile to exchange. The two-dimensional homonuclear J-correlated (COSY) spectrum[11] was recorded at 500 MHz on a General Electric GN-500 spectrometer. The spectrum was obtained from 450 values of t_1 between 0 and 75.6 ms; each free induction decay consisted of 1,024 complex data points, and 64 scans were collected. Total data acquisition time was 10.5 h. The FTNMR program of Dennis Hare was used for data processing on a VAX 8550 computer. Before Fourier transformation, the time-domain data were multiplied by phase-shifted sinebell squared window functions, using phase shifts of 5° and 10° in the t_1 and t_2 dimensions respectively, and the t_1 domain data were extended to 1,024 points with zero-filling. An absolute magnitude calculation then yielded the final spectrum.

by NMR (ref. 12). All 13 are hydrogen-bonded in the crystal structure, according to the neutron diffraction data[26,27].

Two-dimensional ^{1}H-NMR techniques not only monitor the individual backbone amide protons, but also discriminate between protons and deuterons. If an amide proton site is initially deuterated when the protein is unfolded in D_2O, then the resistance (or accessibility) of the amide deuteron to exchange with solvent protons at a specific time during the folding process can be determined from the extent to which the site is protonated when H_2O is added at that time. If the amide proton is also stable to exchange in the native protein, the extent of proton-labelling can be determined after folding of the protein

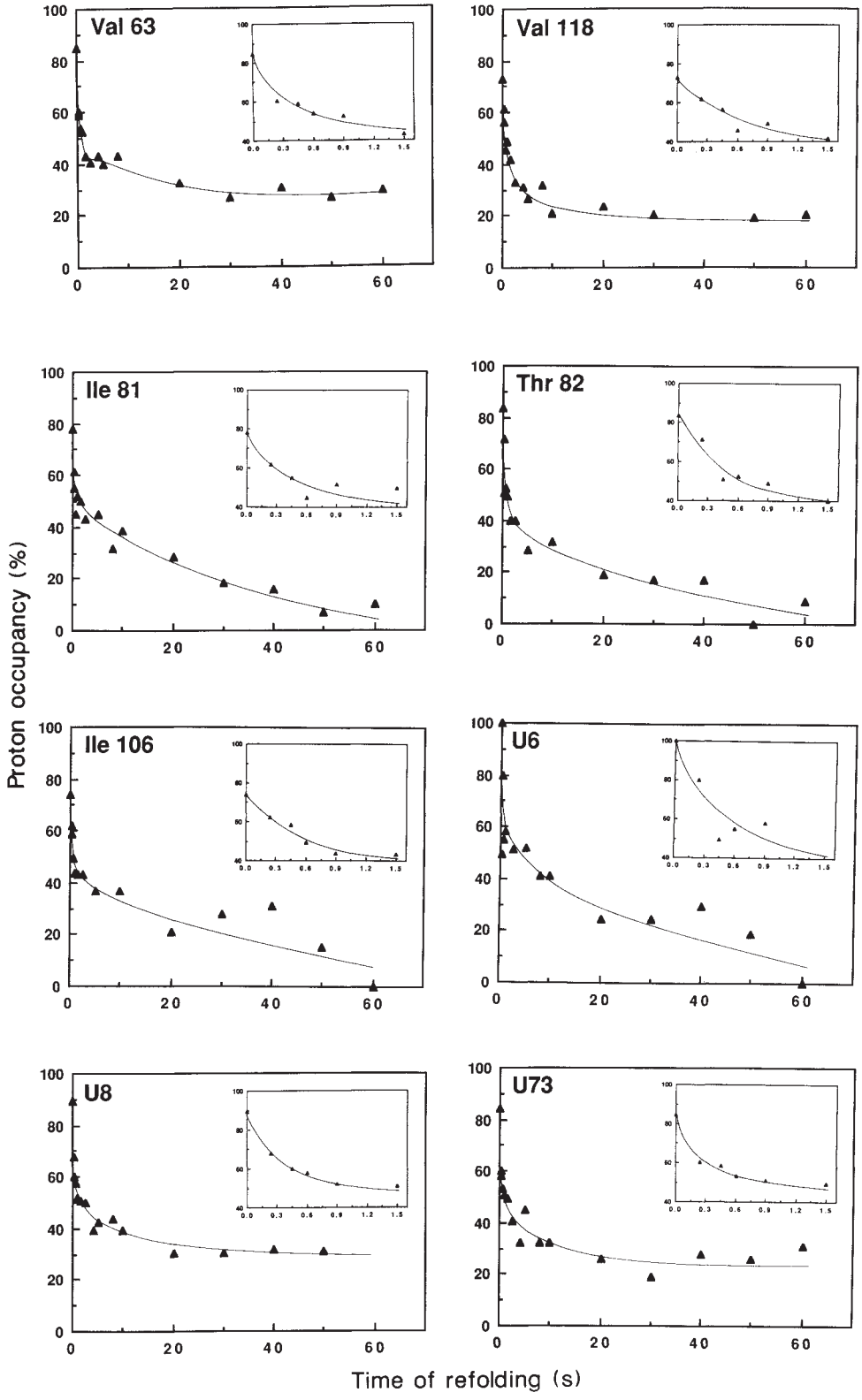

Fig. 2 Kinetics of protection of backbone amide deuterons from exchange with solvent protons. The accessibility to exchange (proton occupancy) of 8 backbone amide proton sites is plotted versus time after initiation of refolding. The C_αH-NH crosspeaks arising from these 8 amide protons are among the most intense of such crosspeaks in the COSY spectrum recorded at 22 °C (Fig. 1). The lines through the data points were drawn by inspection only. The amide protons of Val 63, Val 118, U8 and U73 are partially accessible to exchange even 100 seconds after initiation of refolding.

Methods. Bovine pancreatic ribonuclease A (Sigma, grade XIIA) was purified chromatographically[28] and deuterated by dissolving it in D_2O at *p*H 3, then heating the solution to 60 °C for 20 min; the procedure was repeated after lyophilization. The kinetic experiments were all performed at 5 °C, using a rapid (millisecond) mixing pulsed quench-flow machine that has been described in detail[29]. The deuterated RNase A was unfolded in an unfolding buffer (2.62 M guanidine hydrochloride, 40 mM glycine, in D_2O, final *p*H 2); this unfolded RNase A solution (80 mg ml^{-1}) was allowed to equilibrate for 2 h at 5 °C before use. To initiate refolding, the unfolded RNase A solution was diluted 10.5-fold into a refolding buffer (0.442 M ammonium sulphate, 0.11 M sodium borate in D_2O, *p*H 9.05). At different times after initiation of refolding, the deuterated and partly folded protein solution was diluted 6.25-fold into an exchange buffer (0.4 M ammonium sulphate, 0.25 M guanidine hydrochloride, 0.1 M sodium borate in H_2O, *p*H 9). The mixing dead-time was 10 ms for both dilution steps. After 10 s, amide hydrogen exchange in the 16% deuterated buffer at *p*H 9 was quenched by rapidly mixing with a quench buffer (50 mM sodium formate, *p*H 1.25), so that the final *p*H was 2.9. The refolding reaction was then allowed to go to completion (10 min) at this *p*H. For the zero-time point, the refolding buffer (see above) contained H_2O, not D_2O; thus, refolding and exchange were initiated simultaneously. As for the other time points, exchange was quenched after 10 s, and refolding was allowed to go to completion. The NMR sample of the fully folded RNase A was then prepared as described for the reference spectrum in the Fig. 1 legend. A two-dimensional COSY spectrum of the sample from each time point was recorded at 22 °C. Data recording conditions and data processing were as described for the reference spectrum in Fig. 1 legend. The intensities of the C_αH-NH crosspeaks (the proton occupancies) in each spectrum were determined by calculating the volume integrals of the crosspeaks, after first setting the baseline of the spectrum to zero. The average error in the determination of the volume integrals was less than ±25%. For each spectrum, the average volume of two crosspeaks arising from non-exchangeable protons (the two tyrosine ring-proton crosspeaks, J1 and J2 in Fig. 1) was used as the internal intensity standard. For each C_αH-NH crosspeak, the intensity of the corresponding crosspeak in the reference spectrum (Fig. 1) was chosen to represent 100% proton occupancy.

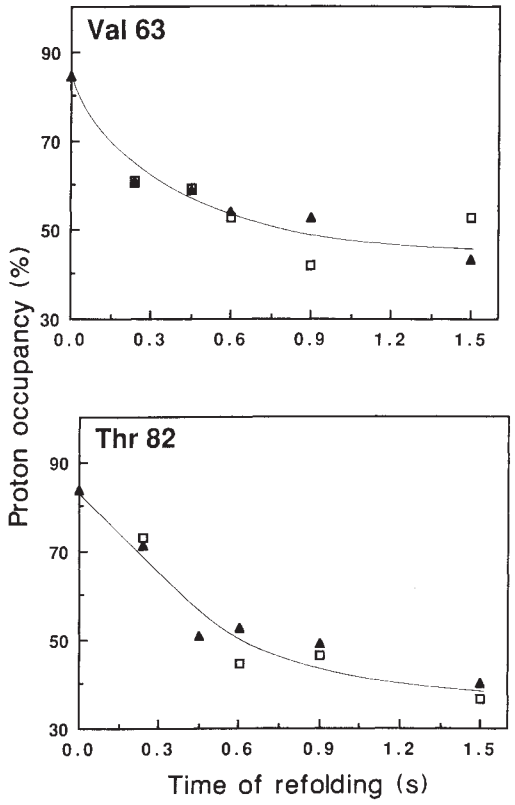

Fig. 3 *p*H dependence of the formation of the early intermediate. Closed symbols, *p*H 9; open symbols, *p*H 10. The experiments at *p*H 10 were done using the same procedures as those used for the experiments at *p*H 9 that are described in legends to Figs 1 and 2. Except for the difference in *p*H, the only other difference was that sodium sulphate was used instead of ammonium sulphate in the buffers at *p*H 10.

is complete, using ^{1}H-NMR techniques. It is necessary to obtain ^{1}H resonance assignments only for the native protein.

Results

In Fig. 2, the accessibility of 8 different backbone amide protons to exchange with solvent protons is plotted against time after initiation of refolding. Deuterated RNase A was allowed to refold in a D_2O buffer for a variable duration of time before exposure to a 10-second proton-labelling pulse at *p*H 9. Proton-labelling was terminated at the end of the pulse by rapidly dropping the *p*H of the solution to 2.9. At this *p*H, amide proton exchange is quenched but folding can proceed to completion. The accessibility to exchange (proton occupancy) of each backbone amide proton site at each time point was determined from the intensity of the corresponding C_αH-NH crosspeak in the COSY spectrum of the fully folded protein sample. A control experiment in which native deuterated RNase A was exposed to a 10-second proton-labelling pulse at *p*H 9, indicated that none of these amide proton sites showed any accessibility to exchange during the exchange pulse (the sites showed zero proton occupancy when assayed by the COSY spectrum). Only 3 out of 40 amide protons showed any degree of proton occupancy in the control experiment, thereby being disqualified from serving as probes for structure formation.

The amide protons shown in Fig. 2, as well as others (data not shown), become protected against exchange according to kinetic curves that are multiphasic and distinct from one another. From the kinetic curves in Fig. 2, it is evident that a kinetic intermediate is formed within the first 1.5 s of refolding. In experiments in which the duration of the proton-labelling pulse, applied 1.5 s after initiation of refolding, was varied between 2

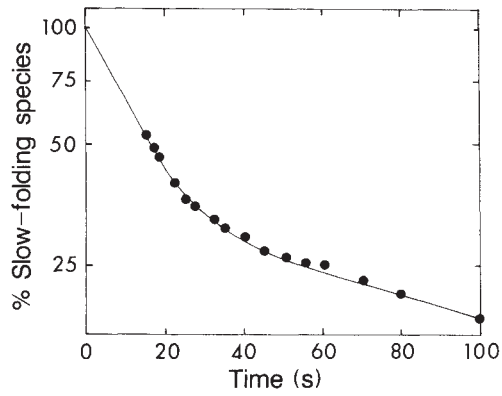

Fig. 4 Refolding kinetics of the slow-folding forms of RNase A. Refolding was monitored by measuring the increase in tyrosine absorbance at 287 nm, which accompanied refolding in a 0.4 M ammonium sulphate, 0.25 M guanidine hydrochloride, 0.1 M sodium borate buffer at *p*H 9 and 5 °C. The solid line is a double exponential fit through the data; the faster phase has a time constant of 6 s, and the slower phase has a time constant of 128 s. Deuterated RNase A was dissolved in an unfolding buffer at a concentration of 20 mg ml^{-1}; refolding was initiated by a manual 10.5-fold dilution into a refolding buffer. Both unfolding and refolding buffers and consequently the conditions of refolding were the same as those used in the experiments described in Fig. 2.

and 20 s, the extent of labelling at several amide proton sites was unchanged (data not shown). This indicates that the extent of labelling at this time after initiation of refolding does not depend on the exchange behaviour of protected protons in the early folding intermediate. Instead, at least for some amide protons, the extent of labelling depends solely on the fraction of RNase A molecules that have formed the early folding intermediate.

To test this conclusion, the 10-s pulse experiment was repeated at *p*H 10. Amide proton exchange in intact proteins is base-catalysed above *p*H 3 in most conditions[19,20]; thus the driving force for exchange (the intrinsic exchange rate) is most conveniently increased 10-fold by increasing the *p*H of the proton-labelling pulse from 9 to 10. As seen in Fig. 3, however, the extent of labelling for the Val 63 and Ile 106 amide protons during the first folding reaction detected is not increased at *p*H 10; several other amide protons also behave similarly (data not shown). This confirms that it is the formation of the early intermediate that is being observed in the pulse-labelling experiments. Under the conditions of the proton-labelling pulse (*p*H 10, 5 °C), the intrinsic amide proton exchange process in unfolded RNase A has a time constant of 0.6 milliseconds[19,20]. As the duration of the pulse was 10 s, the exchange rate of each amide proton shown in Fig. 3 must be slowed by a large factor in the intermediate to avoid complete labelling during the pulse (see Discussion). When the pulse is applied 1.5 s after initiation of refolding at *p*H 9 or 10, the amide proton sites in Figs 2 and 3 are all labelled to an extent of only ~40%, which corresponds to the fraction of protein molecules that have not formed the intermediate during the first 1.5 s of folding. It has been previously shown in the case of the S-protein portion of ribonuclease S (ref. 30), which is a compact globular protein having both secondary and tertiary structure[31], that the exchange rate constants are reduced only ~1,000-fold (ref. 32). The large reduction in exchange rate constants for some of the amide protons in the folding intermediate is therefore quite remarkable.

The refolding of RNase A can also be monitored by observing the change in tyrosine absorbance (Fig. 4) that accompanies the folding process. Unfolded RNase A consists of a mixture of a fast-folding species, U_F, and at least two slow-folding species, U_S^I and U_S^{II} (refs 1 and 4), which probably arise from proline

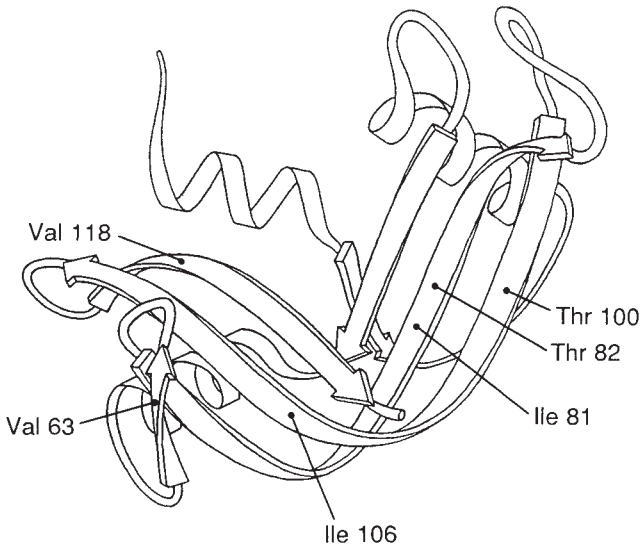

Fig. 5 Tertiary structure of RNase A. The 3 α-helices and the β-sheet are clearly depicted in this ribbon diagram (courtesy of Jane Richardson). The locations of some of the amide protons that have been used as examples in Figs 2 and 3 are shown.

cis-trans isomerization after unfolding[5]. The two kinetic phases observed in the refolding reaction (Fig. 4) reflect the folding of U_S species; the faster phase reflects the folding of the major U_S^{II} species and the slower phase reflects the folding of the minor U_S^{I} species[4]. The refolding of U_F is complete within the dead time (10 s) of the manual mixing; in fact, stopped-flow experiments have indicated that the time constant for refolding of U_F is less than 50 ms both at 25 °C (ref. 33) and at 10 °C (ref. 34). It should be noted that the design of the pulse-labelling experiments in Figs 2 and 3 is also such that only the refolding of the U_S species is being monitored; even at the time of application of the earliest pulse, which was 250 ms after initiation of refolding (except for the zero time points), U_F is completely folded[34] and therefore will not be labelled.

Tyrosine absorbance monitors the formation of the hydrophobic core of the protein, as solvent is excluded from the vicinity of 3 tyrosine residues that are buried in native RNase A. Like many spectroscopic probes, tyrosine absorbance monitors formation of tertiary structure during folding, whereas the amide proton probes monitor formation of secondary structure, although they may also be sensitive to later formation of tertiary structure. During the folding of both the U_S^{I} and U_S^{II} species of RNase A, the tyrosine absorbance probe monitors the formation of a native-like intermediate, I_N (refs 4 and 5), which possesses tertiary structure and subsequently forms native RNase A, probably as the result of one or more proline isomerization reactions[4]. A comparison of Figs 2 and 4 shows that the kinetics of folding of the slow-folding species of RNase A, when measured by the tyrosine absorbance method, are considerably slower than when measured using individual amide protons as probes for folding. This suggests that formation of stable secondary structure precedes the formation of the final tertiary structure during the folding of RNase A, in agreement with earlier low-resolution studies[2,3]. This result with RNase A is different from the result obtained earlier with bovine pancreas trypsin inhibitor (ref. 35), where the amide proton probes and the tyrosine absorbance probe all showed similar folding kinetics.

A competition method[2] was used in that study of bovine pancreas trypsin inhibitor[35]: the competition between folding and amide proton exchange was measured as a function of *p*H because exchange was base-catalysed, whereas the folding kinetics were nearly independent of *p*H in the range studied. The experiments reported here demonstrate the power of the pulse-labelling method. In competition experiments, a folding intermediate can be detected only if it is formed rapidly compared with the rate of amide proton exchange from the unfolded form, which is not the case here. Moreover, if the folding kinetics are complex, as they are for RNase A, the competition experiment does not give information on the kinetics of formation of the intermediate under conditions where it is detected.

Discussion

The kinetics of formation of the early intermediate on the folding pathway of RNase A have been confirmed by [^{3}H]H_2O pulse-labelling experiments (data not shown) using the same rapid (millisecond) mixing method. Earlier manual mixing experiments had shown that an early intermediate, I_1, with many exchange-resistant amide protons[2,3], is formed before the formation of the final tertiary structure. In these low-resolution studies, two mechanisms were considered[2] for the formation of I_1 from the unfolded protein U. In one mechanism, the equilibrium stability model, I_1 is formed very rapidly from U and a rapid pre-equilibrium is established between I_1 and U; later, I_1 undergoes the further folding reactions that lead to formation of the final tertiary structure. The mechanism of amide proton exchange from I_1 is transient unfolding of I_1 to U, with exchange into U. I_1 is found to be only moderately stable compared with U according to this mechanism[2]. In a second mechanism, the kinetic model, I_1 is very stable compared with U; consequently, the formation of I_1 from U is essentially irreversible, and occurs at a rate that is slower than or competitive with exchange into U. The earlier studies could not distinguish between the two mechanisms because neither the stability of I_1 nor the kinetics of formation of I_1 could be directly determined by the methods used, which had both low temporal and structural resolution. The results depicted in Figs 2 and 3 attest to both the high stability of I_1 and a rate of formation of I_1 that, although rapid, is still competitive with the rate of amide proton exchange into U at the *p*H (~7) which was used in the low-resolution studies[2,3]. When the rate of amide proton exchange into U is much faster than the rate of formation of I_1, which is true under the conditions of exchange (*p*H 9 or 10) used in this study, the second mechanism dictates that the extent of proton-labelling of I_1 should be independent of *p*H, as we observe (Fig. 3). It thus appears that the second mechanism, the kinetic model, is more appropriate for describing both the results of the earlier low-resolution studies[2,3] and those reported here.

There are two possible mechanisms[23,24] to explain the strong and *p*H-independent resistance (under our experimental conditions) to exchange of the early intermediate I_1. The first is the EX_2 mechanism, in which a fast pre-equilibrium determines the fraction of I_1 available for exchange. The normal test of this

mechanism is that exchange should be base-catalysed above *p*H 3. As we do not observe base-catalysed exchange into I_1 (Fig. 3), either exchange does not follow the EX_2 mechanism, or otherwise I_1 is completely resistant to exchange under our conditions. The second mechanism is the EX_1 mechanism, in which the observed exchange rate equals the rate of opening of I_1 to exchange. In this case, the observed exchange rate is not base-catalysed and may be *p*H-independent. If exchange into I_1 is occurring in our experiments and takes place by an EX_1 mechanism, then the extent of exchange must increase with the duration of the pulse. This will be tested in future experiments. Usually, exchange into intact proteins occurs by the EX_2 mechanism[23,24].

In a molecular dynamics simulation of the folding of crambin[36] using nuclear Overhauser-effect constraints from NMR data, it has been observed that secondary structure forms before tertiary structure during the folding process. Kinetic circular dichroism studies have indicated that secondary structure is formed before the final tertiary structure during the folding of RNase S (ref. 37), α-lactalbumin[38,39], lysozyme [39], carbonic anhydrase[40-42], cytochrome *c* (ref. 43) and β-lactoglobulin[43]. Such studies, however, do not indicate whether the units of secondary structure are the same as in the native protein, and little is known about the structural stability of folding intermediates. The low *p*H form of α-lactalbumin has properties resembling those of the kinetic folding intermediate[39,44], and tritium-exchange studies suggest that it has only marginal stability[45]. In contrast the early intermediate, I_1, on the folding pathway of RNase A has been shown here to possess very stable secondary structure. It has yet to be determined whether or not the simultaneous protection of several β-sheet amide protons in I_1 means that a β-sheet has formed in the intermediate, and if so, whether it has a structure similar to the β-sheet in the native protein. To do this, it will be necessary to obtain kinetic curves of proton occupancy versus time of folding for a much larger number of sequentially assigned amide protons.

Two fundamentally different models for protein folding have been proposed. In the so-called jigsaw puzzle model[46], which can be traced back to the early two-state model for protein folding, there is neither a defined starting point for folding nor are any unique structural intermediates formed during the folding process. It is proposed that multiple folding pathways exist which are equally accessible to the unfolded protein. Tertiary structure can form without the need for first assembling the backbone; instead, it is assembled in much the same way as a jigsaw puzzle is fitted together. In support of this model, there are numerous possible ways to assemble the native structures of both myoglobin and RNase A by association of compact structural units which are contiguous in the sequences of the proteins[47]. In the framework model[7-9], folding is directed along one or a few well defined sequential pathways with unique structural intermediates. It is proposed that a framework of stable secondary structure is first formed. This secondary structure is presumably stabilized by a small number of specific tertiary interactions; for example, the stabilization of α-helices during the folding of myoglobin is expected to result from as few as 6 principal pairwise helix-helix contacts[7,8]. The subsequent development of additional tertiary interactions leads to the formation of the native tertiary structure of the protein. Kinetic intermediates on the folding pathway are expected to have at least some stable units of secondary structure. The preliminary results reported here show that the technique used in our studies can distinguish between these two distinct models, and the results support the framework model.

At least 5 of the 8 backbone amide protons that are used as examples in Fig. 2 are located in the β-sheet of native RNase A (Fig. 5). All are protected early on during the folding process. Our preliminary results, although they do not exclude the concurrent formation of other secondary structural units, suggest that the β-sheet may be formed in an early folding reaction of RNase A. It is not possible, at this initial stage of investigation, to interpret the complex kinetics evident at later stages of the folding process (Fig. 2). In part, the complexity must reflect the kinetic heterogeneity of unfolded RNase A (ref. 1); the minor slow-folding species U_S^I is known to refold last. The existence of more than one folding pathway for any one of the 3 species of unfolded RNase A cannot as yet be ruled out.

In conclusion, our investigations have shown that a kinetic intermediate with many protected amide protons is formed in an early folding reaction of RNase A. The work reported is consistent with the framework model of protein folding, in which formation of stable secondary structure occurs before the assembly of the final tertiary structure. Recently, considerable progress has been made towards a complete sequential NMR assignment of RNase A (A. D. Robertson, E. O. Purisma, & M. A. Eastman, unpublished results). Once this is complete, the use of nearly 40 backbone amide protons as probes for folding will yield a more detailed characterization of structural intermediates formed on the folding pathway.

We thank Kevin Shoemaker for preliminary experiments, Dennis Hare for a copy of the FTNMR program, Virginia Robbins for purifying the protein, and John Osterhout, Andrew Robertson, Heinrich Roder, Franz Schmid, Kevin Shoemaker and David Wemmer for helpful discussions. J.B.U. is a Fellow of The Jane Coffin Childs Memorial Fund for Medical Research. This investigation has been aided by a grant from the Jane Coffin Childs Memorial Fund for Medical Research and by the National Institutes of Health.

Received 13 June; accepted 31 August 1988.

1. Garel, J.-R. & Baldwin, R. L. *Proc. natn. Acad. Sci. U.S.A.* **70,** 3347-3351 (1973).
2. Schmid, F. X. & Baldwin, R. L. *J. molec. Biol.* **135,** 199-215 (1979).
3. Kim, P. S. & Baldwin, R. L. *Biochemistry* **19,** 6124-6129 (1980).
4. Schmid, F. X. *Biochemistry* **22,** 4690-4696 (1983).
5. Cook, K. H., Schmid, F. X. & Baldwin, R. L. *Proc. natn. Acad. U.S.A.* **76,** 6157-6161 (1979).
6. Roder, H., Elove, G. A. & Englander, S. W. *Nature* **335,** 700-704 (1988).
7. Ptitsyn, O. B. & Rashin, A. A. *Biophys. Chem.* **3,** 1-20 (1975).
8. Richmond, T. J. & Richards, F. M. *J. molec. Biol.* **119,** 537-555 (1978).
9. Kim, P. S. & Baldwin, R. L. *A. Rev. Biochem.* **51,** 459-489 (1982).
10. Wüthrich, K. *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
11. Aue, W. P., Bartoldi, E. & Ernst, R. R. *J. chem. Phys.* **64,** 2229-2246 (1976).
12. Hahn, U. & Rüterjans, H. *Eur. J. Biochem.* **152,** 481-491 (1985).
13. Wand, A. J., Di Stefano, D. S., Feng, Y., Roder, H. & Englander, S. W. *Biochemistry* (in the press).
14. Redfield, C. & Dobson, C. M. *Biochemistry* **27,** 122-136 (1988).
15. LeMaster, D. M. & Richards, F. M. *Biochemistry* **27,** 142-150 (1988).
16. Creighton, T. E. *Prog. Biophys. molec. Biol.* **33,** 231-297 (1978).
17. Creighton, T. E. *Proc. natn. Acad. Sci. U.S.A.* **85,** 5082-5086 (1988).
18. States, D. J., Creighton, T. E., Dobson, C. M. & Karplus, M. *J. molec. Biol.* **195,** 731-739 (1987).
19. Kuwajima, K., Kim, P. S. & Baldwin, R. L. *Biopolymers* **22,** 59-67 (1983).
20. Kim, P. S. *Meth. Enzym.* **131,** 136-156 (1986).
21. Brems, D. N. & Baldwin, R. L. *Biochemistry* **24,** 1689-1693 (1985).
22. Roder, H. *Meth. Enzym.* (in the press).
23. Englander, S. W. & Kallenbach, N. R. *Q. Rev. Biophys.* **16,** 521-655 (1984).
24. Wagner, G. *Q. Rev. Biophys.* **16,** 1-57 (1983).
25. Englander, S. W., Donner, N. N. & Teitelbaum, H. *A. Rev. Biochem.* **41,** 903-924 (1972).
26. Wlodawer, A. & Sjölin, L. *Proc. natn. Acad. Sci. U.S.A.* **79,** 1418-1422 (1982).
27. Wlodawer, A. & Sjölin, L. *Biochemistry* **22,** 2720-2728 (1983).
28. Garel, J.-R. *Eur. J. Biochem.* **70,** 179-189 (1976).
29. Cash, D. J. & Hess, G. P. *Analyt. Biochem.* **112,** 39-51 (1981).
30. Richards, F. M. & Vithayathil, P. J. *J. biol. Chem.* **234,** 1459-1465 (1959).
31. Richards, F. M. & Wyckoff, H. W. in *The Enzymes 3rd edn* Vol. 4 (ed Boyer, P. D.) 647-806 (Academic, New York 1971).
32. Rosa, J. J. & Richards, F. M. *J. molec. Biol.* **145,** 835-851 (1979).
33. Garel, J.-R., Nall, B. T. & Baldwin, R. L. *Proc. natn. Acad. Sci. U.S.A.* **73,** 1853-1857 (1976).
34. Lin, L.-N. & Brandts, J. F. *Biochemistry* **22,** 564-573 (1983).
35. Roder, H. & Wüthrich, K. *Proteins* **1,** 34-42 (1986).
36. Brünger, A. T., Clore, G. M., Groneborn, A. M. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **83,** 3801-3805 (1986).
37. Labhardt, A. M. *Proc. natn. Acad. Sci. U.S.A.* **81,** 7674-7678 (1984).
38. Gilmanshin, R. I. & Ptitsyn, O. B. *FEBS Lett.* **223,** 327-329 (1987).
39. Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. *Biochemistry* **24,** 874-881 (1985).
40. McCoy, L. F., Rowe, E. S. & Wong, K. P. *Biochemistry* **19,** 4738-4743 (1980).
41. Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A. & Ptitsyn, O. B. *FEBS Lett.* **165,** 88-92 (1984).
42. Semisotnov, G. V. *et al. FEBS Lett.* **224,** 9-13 (1987).
43. Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. & Nagamura, T. *FEBS Lett.* **221,** 115-118 (1987).
44. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. *Biochemistry* **25,** 6965-6972 (1986).
45. Dolgikh, D. A. *et al. Eur. biophys. J.* **13,** 109-121 (1985).
46. Harrison, S. C. & Durbin, R. *Proc. natn. Acad. Sci. U.S.A.* **82,** 4028-4030 (1985).
47. Lesk, A. M. & Rose, G. D. *Proc. natn. Acad. Sci. U.S.A.* **78,** 4304-4308 (1981).