Identification of a New Site of Conformational Heterogeneity in Unfolded Ribonuclease A

Marc Adler[1,2] and Harold A. Scheraga[1,3]

The results presented here indicate that there are two slowly exchanging conformational isomers in unfolded bovine pancreatic ribonuclease A (RNase A) in the vicinity of Lys-41. The conformational heterogeneity is not observed in the fully folded protein. Therefore, one of the isomers may correspond to one of the slow-folding forms of the protein observed when refolding is initiated. These results were obtained from a chemically modified form of the protein, CL(7–41) RNase A, that has a dinitrophenyl cross-link between the ε-amino groups of Lys-7 and Lys-41. Extensive physical studies have shown that the cross-link does not significantly perturb the structure or the folding pathways of the protein. Therefore, the results obtained from this modified form of the protein are relevant to intact RNase A. The one-dimensional (1D) NMR spectrum of heat-unfolded CL(7–41) RNase A reveals that the singlet resonance for the C_3H ring proton of the dinitrophenyl cross-link has been split into two unequal peaks in a $3:1$ ratio, indicating that there are two distinct environments for the dinitrophenyl group. Variations in temperature, and the addition of urea, do not affect the relative peak intensities. The two peaks collapse into one after the protein is refolded. The observed splitting must originate from a slow reversible isomerization (>100 msec) in a neighboring bond. The two most likely candidates are either the *cis/trans* isomerization of the Lys-41–Pro-42 peptide bond or hindered rotation about the disulfide bond between Cys-40 and Cys-95.

1. INTRODUCTION[4]

Reversibly denatured bovine pancreatic ribonuclease A (RNase A) refolds in three phases, U_f, U_s^I, and

---

[1] Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301.

[2] Present address: Biophysics Research Division, Institute of Science and Technology, University of Michigan, Ann Arbor, Michigan 48109.

[3] To whom all correspondence should be addressed.

[4] Abbreviations: RNase A, bovine pancreatic ribonuclease A; CL(7–41) RNase A, RNase A modified by a dinitrophenyl cross-link between the ε-amino groups of Lys-7 and Lys-41; 8SO$_3^-$ CL(7–41) RNase A, CL(7–41) RNase A in which the four internal disulfide bonds have been reduced and then blocked by SO$_3^-$; CL and CL*, the resonance of the C_3H ring proton of the dinitrophenyl cross-link in CL(7–41) RNase A in the folded and unfolded protein, respectively; U_f, U_s^I and U_s^{II}, the single fast and two slow phases found when RNase A refolds; NMR, nuclear magnetic resonance; ppm, parts per million; DSS, 2, 2-dimethyl 2-silapentane sulfonate.

U_s^{II}. The fast phase, U_f, refolds on the millisecond time scale. The refolding of the two other phases, U_s^I and U_s^{II}, is retarded by two to three orders of magnitude (Tsong *et al.*, 1972; Garel and Baldwin, 1973; Garel *et al.*, 1976; Hagerman *et al.*, 1979; Cook *et al.*, 1979; Schmid and Blaschek, 1981; Lin and Brandts, 1983a). It is believed that the slow phases arise from nonnative conformational isomers of key covalent bonds. The *cis/trans* isomerizations of the four proline peptide bonds (Schmid and Baldwin, 1978; Lin and Brandts, 1983b, 1984, 1987a, b; Schmid, 1986; Schmid *et al.*, 1986; Adler and Scheraga, 1990) have been identified as possible causes for the slow-folding species in RNase A. There are also four intramolecular disulfides in RNase A. It has been hypothesized that the nonnative isomers of these disulfide bonds might also give rise to slow-folding species (Mui *et al.*, 1985). The C^β–S–S–C^β bond is stable in two conformations with the C^β–S bonds projecting

out at right angles to each other. Thus, both left- and right-handed forms are stable. The barrier to rotation about the sulfur-sulfur bond is estimated to be as high as $15.5 \, \text{kcal} \, \text{mol}^{-1}$, based on studies with model compounds (Kessler and Rundel, 1968; Fraser et al., 1971), and might be higher inside a globular protein.

Unfortunately, little progress has been made in identifying which isomers are responsible for the slow-folding species. Progress has been slow because it is difficult to measure the percentage of native and nonnative isomers of specific bonds. Recently, some progress has been made. Previous work from this laboratory (Adler and Scheraga, 1990) has shown that the percentage of trans isomers of Pro-93 and Pro-114 can be determined from the 1D NMR spectra of heat-unfolded RNase A. The results indicate that useful data can be extracted from the 1D NMR spectra of unfolded RNase A in spite of the problems of overlapping resonances.

In this report, the 1D NMR studies have been extended to a chemically modified form of the protein, CL(7-41) RNase A (Lin et al., 1984). CL(7-41) RNase A has a dinitrophenyl cross-link between the ε-amino groups of Lys-7 and Lys-41. Extensive physical studies (Lin et al., 1984), including X-ray crystallography (Weber et al., 1985), have shown that the cross-link does not perturb the structure of the protein. Furthermore, CL(7-41) RNase A refolds in three phases, whose rates of refolding and relative proportions mirror those of native ribonuclease A (Lin et al., 1985).

There are a few important differences between the native and modified forms of the protein. The cross-link is attached to active-site lysines. Therefore, CL(7-41) RNase A is inactive. The cross-link also increases the melting temperature of RNase A by 20°C, by reducing the conformational entropy of the unfolded state (Lin et al., 1984). Finally, since native RNase A does not contain any dinitrophenyl groups, the C_3H ring proton of the cross-link is resolved from all other resonances in the 1D NMR spectra of both the native and unfolded forms of the protein.

In the unfolded protein, the C_3H ring proton has two distinct chemical shifts. This indicates that the proton senses two conformational isomers of a neighboring group, either the cis/trans isomerization of Pro-42 or possibly the rotational isomers of the Cys-40-Cys-95 disulfide bond. This paper discusses these results in the light of recently acquired data.

2. MATERIALS AND METHODS

CL(7-41) RNase A was prepared and isolated by the method of Lin et al. (1984), and was a generous gift from S. H. Lin and D. M Rothwarf. $8SO_3^-$ CL(7-41) RNase A was prepared as described by Lin et al. (1984), and was also a generous gift from D. M. Rothwarf. Deuterated urea was purchased from Bio-Rad (Richmond, California), and was used without any further purification. The exchangeable protons of both CL(7-41) and $8SO_3^-$ CL(7-41) were replaced by deuterium by heating the protein to 60°C in 99.6% D_2O (Aldrich, Milwaukee, Wisconsin) and then removing the solvent by lyophilization. The process was repeated three times. The pH was adjusted to 2.3 for CL(7-41) RNase A and to 7 for $8SO_3^-$ CL(7-41) RNase A with DCl and NaOD. pH 7 was used for $8SO_3^-$ CL(7-41) RNase A because of its low solubility at acidic pH. The spectra were referenced to an internal standard of DSS.

NMR spectra were measured on a Varian VXL-400 spectrometer using a 20 μsec 90° pulse and a total cycle time of 4 sec. 512 free-induction decays were measured for each spectrum. The data presented in Figs. 1-3 were smoothed with 1 Hz, 1 Hz, and 0.5 Hz exponential line-broadening, respectively (Ferrige and Lindon, 1978).

For integration purposes, baseline drift was corrected using a first-order polynomial. The data quoted in this paper were derived by setting the upper and lower bounds for integration very close to the sharp rise in intensity of the peak. This gave the most reproducible results (25% and 75% intensity for the minor and major peaks, $1\sigma = 2\%$). If the limits of integration were broadened, the ratio of the two peaks changed to 23% and 77%, and the fluctuation between measurements doubled.

Saturation transfer experiments (Fox et al., 1986; Evans et al., 1987, 1989) were carried out by pre-irradiating the native resonance of the cross-link for 1 sec. The total cycle time between 90° pulses was 3 sec. A total of 1600 acquisitions was made at each decoupler setting. Control spectra were obtained by pre-irradiating the neighboring native $C^\varepsilon H$ His peak as well as irradiating an area with no resonances.

After the measurements were made, both CL(7-41) RNase A and $8SO_3^-$ CL(7-41) RNase A were checked for chemical degradation by HPLC on a Mono S high-performance cation exchange column at pH 7 (Pharmacia, Upsala, Sweden). There were no observable breakdown products for CL(7-41) RNase A (<2%). By contrast, 25% of the $8SO_3^-$ CL(7-

41) RNase A had been modified, presumably by deamidation at Asn-67, which is also seen in $8SO_3^-$ RNase A under similar conditions (Thannhauser and Scheraga, 1985). However, this chemical heterogeneity did not affect the resonances examined in this paper to any measurable extent.

1D NMR spectra of CL(7-41) RNase A did reveal that 20% of the protein did not refold after prolonged heating even though there was no apparent breakdown as seen by HPLC. This has also been observed for intact RNase A (Adler and Scheraga, 1990). However, the ratio of the major and minor peaks of the C_3H ring proton of the cross-link did not correlate with the amount of nonfolding forms of the protein.

3. RESULTS

The downfield region of the 1D proton NMR spectrum of CL(7-41) RNase A at the midpoint of the thermal transition appears in Fig. 1. The chemical shift of the four $C^\varepsilon H$ histidine resonances, labeled H1-H4, correspond closely to those seen in unmodified RNase A under similar conditions (Markley, 1975). If we assume that the cross-link did not perturb the chemical shifts, then the resonances can be assigned as follows: H1 is His-105, H2 is His-119, H3 is His-12, and H4 is His-48. However, the exact assignments of the histidines are not important for this study, and no attempt was made to verify them.

The NMR spectrum of CL(7-41) at the midpoint of the thermal transition at pH 2.3 contains two resonances, labeled as CL and CL*. These resonances would not be present in unmodified RNase A under similar conditions. Saturation transfer experiments

(results not shown) demonstrate that these resonances belong to the same proton in the folded and unfolded states, respectively. CL* has a similar chemical shift as the C_3H proton of the phenyl ring of N^ε, $N^{\varepsilon'}$-(2, 4-dinitrophenylene-1, 5)-bis(N^α-acetyl-L-lysine methylamide), which is a model compound for the dinitrophenyl cross-link (Lin et al., 1984). From this evidence, we conclude that CL and CL* belong to the C_3H proton of the phenyl ring of the cross-link in the folded and unfolded states, respectively.

The integrated intensities of the peaks of CL and CL* were monitored as a function of temperature throughout the transition region, and they appear to follow the cooperative melting process at the same rate as the four histidines (Fig. 1 and additional data not shown).

In Fig. 2, the peak corresponding to the C_3H proton of the cross-link in the unfolded protein, CL*,

is shown in greater detail. It actually consists of two resonances, referred to as the major and minor peaks that exist in a 75% to 25% ± 2% ratio. The relative intensity of the two peaks was not affected by changes in temperature from 51–63° ($T_m = 53°C$, Fig. 2a and b, and additional data not shown). Similar experiments were carried out in the presence of 5 M urea, pH 2.3. Once again, the relative intensities of the minor and major peaks were not affected throughout the transition region and beyond (from 22–56°C, $T_m = 25°C$; Fig. 2c and d, and additional data not shown).

Furthermore, these two peaks remained distinct at higher temperatures (Fig. 2b and d). This indicates that the rate of interconversion of the two isomeric states is slow (>100 msec) on the NMR time scale. If the rate of exchange were faster, the two peaks would have merged into one.

The resonance of the C_3H proton in the native form, CL, appears as a well-resolved singlet under all the conditions used in this study (Fig. 1 and additional data not shown). Therefore, one of the isomeric states is unstable in the fully folded protein and reverts back to the native form.

One possibility is that the splitting of CL* reflects the formation of compact structure in a region of the protein. However, changes in temperature as well as the addition of 5 M urea should disrupt the formation of any local ordered structure and, in turn, alter the ratio of the two peaks. Such effects have been seen for the folding of the N-terminal α-helix (residues 3–13) in partially folded forms of RNase A (Kim and Baldwin, 1984; Swadesh et al., 1984; Rico et al., 1986; Adler and Scheraga, 1988). The helix itself unfolds at higher temperatures and its folding could not account for the observed effects. It is much more likely that the splitting of CL* reflects the isomerization of a single covalent bond into native and nonnative forms.

An examination of the amino acid sequence of ribonuclease A reveals that there are no covalent bonds flanking Lys-7 that might have two stable isomers in the unfolded protein. However, Lys-41 is flanked by Pro-42 and a disulfide bond between Cys-40 and Cys-95. The X-Pro peptide bond can isomerize between the cis and trans forms in the unfolded state (Lin and Brandts, 1983b, 1987a, b; Schmid, 1986; Schmid et al., 1986; Adler and Scheraga, 1990). Studies on model compounds indicate that the rotation around disulfide bonds, such as the one between Cys-40 and 95, is hindered and two conformations are stable (Kessler and Rundel, 1968; Fraser et al., 1971). Longer-range interactions with groups such as
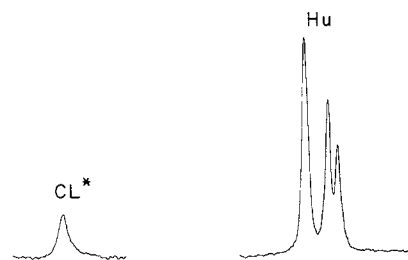

Fig. 3. The downfield region of $8SO_3^-$ CL(7–41) RNase A at pH 7, 14°C. CL* corresponds to the C_3H proton of the dinitrophenyl cross-link of the unfolded protein. The peaks corresponding to the four $C^\varepsilon H$ protons of His, labeled Hu, are included for comparison. Due to local conformation effects in the unfolded form (Swadesh et al., 1984), the four $C^\varepsilon H$ protons do not collapse into a single peak.

Pro-93 can probably be excluded since neither high temperature nor 5 M urea perturbed the NMR spectrum of the two isomers (Fig. 2).

If the splitting seen in CL* arises from interactions with the two rotational isomers of the disulfide, then it should be eliminated when the four disulfide bonds of RNase A are reduced and blocked with a symmetrical ligand such as $SO_3^=$. (The reduction of the four disulfide bonds also causes the protein to unfold.) Indeed, the NMR spectrum of $8SO_3^-$ CL(7–41) RNase A shows a single resolved resonance for the C_3H proton of the dinitrophenyl cross-link (Fig. 3). Unfortunately, this negative result is inconclusive. The reduction of the disulfide bonds may have disrupted the interactions between the cross-link and the proline. Therefore, more experiments need to be performed to determine which covalent bond is responsible for the splitting of the CL* resonance.

4. DISCUSSION

The experiments presented here demonstrate that there is a slowly exchanging isomer of a covalent bond in unfolded CL(7–41) RNase A, and presumably in unfolded RNase A, in the vicinity of either Lys-7 or Lys-41. The ratio of the two isomers is 25% to 75% ± 2%. Possible sites for such an isomerization include the peptide bond of Lys-41–Pro-42 and the disulfide bond between Cys-40 and Cys-95.

Preliminary experiments were carried out with a reduced form of the cross-linked protein, $8SO_3^-$ CL(7–41) RNase A, in order to identify the site of conformational heterogeneity sensed by CL*. The experiments indicated that the disulfide bond between Cys-40 and Cys-95 was responsible for the observed isomers (see Section 3). Unfortunately, the results were inconclusive. A more definitive answer could be obtained by

repeating the NMR measurements on a modified form of the cross-linked protein, where glycine is substituted for Pro-41. If the two peaks of CL* collapsed into one, then this would be strong indication that the *cis/trans* isomerism of Pro-42 generated the two isomers. Genetic substitution of glycine for proline has been used to investigate the conformational isomers of staphylococcal nuclease (Evans *et al.*, 1987) and calbindin D_{9k} (Chazin *et al.*, 1989). Work is in progress in our laboratory to try to produce the modified RNase A in sufficient amounts for NMR experiments.

The results presented here can be combined with earlier measurements of the percentage of *cis* and *trans* isomers of Pro-93 and Pro-114 in heat-denatured RNase A. Earlier work by Adler and Scheraga (1990) showed that only 15% of the unfolded RNase A retained the native *cis* isomer of both of these prolines. Assuming that there is no interaction between the three sites, estimates can be made of the remaining unfolded species that retain the native isomer at all three sites. If the major peak of CL* is the native isomer, then only 11% of the denatured protein retains the native conformation around all the relevant bonds. If the minor peak represents the native isomers, then no more than 4% of the unfolded protein retains all the native isomers. However, 22% (instead of 11%) of RNase A refolds as a single homogeneous phase on the millisecond time scale (Tsong *et al.*, 1972; Garel and Baldwin, 1973; Garel *et al.*, 1976). These data indicate that the fast phase is a heterogenous mixture that includes nonnative isomers of covalent bonds, such as prolines and/or disulfide bonds.

This paper, when combined with the previous work (Adler and Scheraga, 1990), indicates that certain covalent bonds in the unfolded protein are free to interconvert between native and nonnative isomers. Therefore, only a small amount of the denatured protein corresponds to the native isomer. When the protein is returned to folding conditions, a large percentage of the protein refolds on the millisecond time scale in spite of the incorrect isomer(s). Therefore, some of these nonnative isomers do not inhibit refolding. One possibility is that the rate-limiting step for refolding occurs early in the pathway, and it involves only a small number of residues. Alternatively, there may be a degree of conformational flexibility in the early stages of refolding that accommodates many different isomers. Therefore, the structure of the folding intermediates would be much looser than the native protein.

For completeness, one more theory should be discussed. It is possible that the U_f species is a mixture of many nonnative isomers, all of which refold on roughly the same time scale (i.e., msec). The totally native isomer of the unfolded protein may be present in extremely low concentration and has escaped detection in conventional stop-flow experiments. U_f refolds to the native form on the msec time scale in spite of several nonnative isomers or, alternatively, the rate of isomerization may be accelerated by the locally ordered structures during refolding. It is interesting to note that the enthalpy of refolding of U_f to native is 10 kcal mol^{-1} (Lin and Brants, 1983a), which could reflect the rate of isomerization of a covalent bond(s).

ACKNOWLEDGMENTS

REFERENCES

Adler, M., and Scheraga, H. A. (1988). *Biochem.* **27**, 2471–2480.

Adler, M., and Scheraga, H. A. (1990). *Biochem.* **29**, 8211–8216.

Chazin, W. J., Kordel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundstrom, T., and Forsen, S. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 2195–2198.

Cook, K. H., Schmid, F. X., and Baldwin, R. L. (1979). *Proc. Natl. Acad. Sci. USA* **76**, 6157–6161.

Evans, P. A., Dobson, C. M., Kautz, R. A., Hatfull, G., and Fox, R. O. (1987). *Nature* **329**, 266–268.

Evans, P. A., Kautz, R. A., Fox, R. O., and Dobson, C. M. (1989). *Biochem.* **28**, 362–370.

Ferrige, A. G., and Lindon, J. C. (1978). *J. Mag. Reson.* **31**, 337–340.

Fox, R. O., Evans, P. A., and Dobson, C. M. (1986). *Nature* **320**, 192–194.

Fraser, R. R., Boussard, G., Saunders, J. K., Lambert, J. B., and Mixan, C. E. (1971). *J. Am. Chem. Soc.* **93**, 3822–3823.

Garel, J.-R., and Baldwin, R. L. (1973). *Proc. Natl. Acad. Sci. USA* **70**, 3347–3351.

Garel, J.-R., Nall, B. T., and Baldwin, R. L. (1976). *Proc. Natl. Acad. Sci. USA* **73**, 1853–1857.

Hagerman, P. J., Schmid, F. X., and Baldwin, R. L. (1979). *Biochem.* **18**, 293–297.

Kessler, H., and Rundel, W. (1968). *Chem. Ber.* **101**, 3350–3357.

Kim, P. S., and Baldwin, R. L. (1984). *Nature* **307**, 329–334.

Lin, L.-N., and Brandts, J. F. (1983a). *Biochem.* **22**, 564–573.

Lin, L.-N., and Brandts, J. F. (1983b). *Biochem.* **22**, 559–563.

Lin, L.-N., and Brandts, J. F. (1984). *Biochem.* **23**, 5713–5723.

Lin, L.-N., and Brandts, J. F. (1987a). *Biochem.* **26**, 1826–1830.

Lin, L.-N., and Brandts, J. F. (1987b). *Biochem.* **26**, 3537-3543.

Lin, S. H., Konishi, Y., Denton, M. E., and Scheraga, H. A. (1984). *Biochem.* **23**, 5504-5512.

Lin, S. H., Konishi, Y., Nall, B. T., and Scheraga, H. A. (1985). *Biochem.* **24**, 2680-2686.

Markley, J. L. (1975). *Biochem.* **14**, 3546-3554.

Mui, P. W., Konishi, Y., and Scheraga, H. A. (1985). *Biochem.* **24**, 4481-4489.

Rico, M., Santoro, J., Bermejo, F. J., Herranz, J., Nieto, J. L., Gallego, E., and Jiménez, M. A. (1986). *Biopolymers* **25**, 1031-1053.

Schmid, F. X. (1986). *FEBS Lett.* **198**, 217-220.

Schmid, F. X., and Baldwin, R. L. (1978). *Proc. Natl. Acad. Sci. USA* **75**, 4764-4768.

Schmid, F. X., and Blaschek, H. (1981). *Eur. J. Biochem.* **114**, 111-117.

Schmid, F. X., Grafl, R., Wrba, A., and Beintema, J. J. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 872-876.

Swadesh, J. K., Montelione, G. T., Thannhauser, T. W., and Scheraga, H. A. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 4606-4610.

Thannhauser, T. W., and Scheraga, H. A. (1985). *Biochem.* **24**, 7681-7688.

Tsong, T. Y., Baldwin, R. L., and Elson, E. L. (1972). *Proc. Natl. Acad. Sci. USA* **69**, 1809-1812.

Weber, P. C., Sheriff, S., Ohlendorf, D. H., Finzel, B. C., and Salemme, F. R. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 8473-8477.