Nonrandom Distribution of the One-Disulfide Intermediates in the Regeneration of Ribonuclease A[†]

Xiaobing Xu, David M. Rothwarf, and Harold A. Scheraga*

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301

Received January 16, 1996; Revised Manuscript Received March 21, 1996[⊗]

ABSTRACT: The one-disulfide intermediates formed during the oxidative refolding of ribonuclease A (RNase A) have been characterized. This information is important for understanding the folding pathways of RNase A. The one-disulfide intermediates were blocked with 2-aminoethyl methanethiosulfonate, fractionated using ion-exchange chromatography, and digested with trypsin and chymotrypsin. The resulting peptide fragments were fractionated using reversed phase high-performance liquid chromatography, and identified using mass spectrometry. The relative population of each one-disulfide intermediate was determined from its disulfide bond concentration using a postcolumn disulfide detection system. A total of 24 out of 28 possible one-disulfide intermediates were found to be populated (greater than 0.3%) in the one-disulfide mixture. The population of one-disulfide intermediates displays a nonrandom distribution. All four native disulfide pairings have populations greater than those predicted by loop entropy calculations, suggesting the presence of enthalpic contributions stabilizing these species. The one-disulfide intermediate [65, 72], containing the disulfide bond between cysteines 65 and 72, comprises 40% of the entire one-disulfide population. The interactions that stabilize this intermediate may play an important role in the regeneration pathways of RNase A.

The protein-folding problem, the question of how the amino acid sequence determines the three-dimensional structure of a protein, has fascinated researchers for many years. Although a wealth of information has been obtained in attempts to solve this problem, the answer is far from complete. Characterization of folding intermediates has proven to be a very powerful tool in the elucidation of protein-folding pathways (Kim & Baldwin, 1990; Matthews, 1993; Ptitsyn, 1995; Houry et al., 1995; Li et al., 1995). However, such investigations are hindered by the transient nature of folding intermediates. For proteins containing disulfide bonds, an alternative method, which utilizes the formation of disulfide bonds, is often used. In this method, the disulfide-bonded intermediates which form during the folding process are chemically trapped. These folding intermediates can then be isolated and structurally characterized. The folding pathways of several proteins have been studied in this manner. Among them, BPTI[1] (Creighton & Goldenberg, 1984; Weissman & Kim, 1991, 1992; Creighton, 1992; Darby et al., 1995) and RNase A[1] (Hantgan et al., 1974; Creighton, 1979; Scheraga et al., 1984; Rothwarf & Scheraga, 1993a−d; Li et al., 1995) have been the most extensively studied.

The study of protein folding through disulfide bond formation is normally carried out by characterizing the kinetics, thermodynamics, and structures of different disulfide intermediates. In order for this approach to provide meaningful information about protein folding, the conformational folding process that occurs during regeneration from the disulfide-reduced state to the native state must be coupled to the formation of disulfide bonds. Unfortunately, this is not true for the regeneration of BPTI. Fully reduced BPTI contains extensive nonrandom structure (Pan et al., 1995), and most disulfide-bonded intermediates of BPTI were found to be very native-like (Staley & Kim, 1992; van Mierlo et al., 1992, 1993; Weissman & Kim, 1995). Therefore, most of the regeneration pathway involves disulfide rearrangement among species with native-like structure rather than conformational folding. In contrast, RNase A provides a better system to study the folding problem through disulfide bond formation.

RNase A has served as a model protein for understanding the protein-folding problem ever since the landmark work of Anfinsen (1973). RNase A contains 124 residues, with four native disulfide bonds at positions 26−84, 40−95, 58−110, and 65−72. Previous studies of the oxidative refolding of RNase A (Konishi & Scheraga, 1980a,b; Rothwarf & Scheraga, 1993a,b) have shown that the one- and two-

[†] This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (Grant GM-24893). Support was also received from the National Foundation for Cancer Research. X. Xu is a Leukemia Society of America Special Fellow. A brief report of these results was presented at the Protein Society meeting in Boston, MA, in July 1995 [Xu, X., & Scheraga, H. A. (1995). The one-disulfide intermediates formed in the regeneration of ribonuclease A display a non-random distribution, *Protein Sci. 4* (Suppl. 2), 249-T].

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1996.

[1] Abbreviations: RNase A, bovine pancreatic ribonuclease A; BPTI, bovine pancreatic trypsin inhibitor; [65, 72], one-disulfide intermediate with a disulfide bond between the two cysteines denoted in the brackets; des-[40, 95]-RNase A, ribonuclease A lacking the disulfide bond between the cysteine residues denoted in the brackets; 1S, one-disulfide intermediates; 2S, two-disulfide intermediates; 3S, three-disulfide intermediates; AEMTS, 2-aminoethyl methanethiosulfonate [$(NH_2)C_2H_5$-SSO_2CH_3]; DDS, disulfide detection system; MS, mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization-time of flight; GSSG, oxidized glutathione; GSH, reduced glutathione; DTT^{red}, reduced dithiothreitol; DTT^{ox}, oxidized dithiothreitol; EDTA, ethylenediaminetetraacetic acid; GdnSCN, guanidine thiocyanate; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HPLC, high-performance liquid chromatography; Tris, tris(hydroxymethyl)aminomethane; TFA, trifluoroacetic acid; CFIS, chain-folding initiation site.

S0006-2960(96)00090-6 CCC: $12.00

disulfide intermediates are largely conformationally disordered, indicating that further disulfide bond formation is essential to the formation of significant native structure. In addition, it has been determined that the rate-limiting steps in the regeneration of RNase A correspond to the major conformational folding processes (Rothwarf et al., 1995). Therefore, it seems that folding to the native structure occurs along with the formation of disulfide bonds in RNase A. Characterization of the disulfide-bonded intermediates of RNase A should provide a detailed folding pathway for this protein.

However, since RNase A contains eight cysteines, the possible number of disulfide-bonded intermediates involved is very large [a total of 762 as opposed to 74 for the three-disulfide protein BPTI (excluding the native and reduced states)]. This makes it difficult to separate and identify different intermediates. In addition, early studies used GSSG/GSH as the redox couple which can form stable mixed disulfides with disulfide intermediates, making the problem even more complicated. Only in recent years has significant progress been made in such investigations by using DTT^{red}/DTT^{ox} as the redox pair and a novel thiol-blocking reagent, 2-aminoethyl methanethiosulfonate (AEMTS) (Rothwarf & Scheraga, 1993a−d; Li et al., 1995; Rothwarf et al., 1995). AEMTS blocks sulfhydryl groups at least 5 orders of magnitude faster than the conventionally used alkylhalides, e.g. iodoacetate, which ensures rapid quenching of the oxidative refolding process (Rothwarf & Scheraga, 1991). Blocking of a cysteine with AEMTS introduces an additional positive charge. This allows the disulfide intermediates to be separated according to their charges using ion-exchange chromatography. Using this technique, native RNase A was successfully regenerated with DTT^{red}/DTT^{ox} from the fully reduced and denatured form (Rothwarf & Scheraga, 1993a−d). The one-, two-, three-, and four-disulfide intermediates formed in the regeneration process were separated as groupings. The rates and equilibrium constants for the interconversion of these intermediates at 25 °C and pH 8.0 have been determined. It was found that the one-, two-, three-, and four (non-native)-disulfide intermediates achieve a steady-state distribution early in the regeneration process. Building upon these results, we have found that it is possible to characterize the regeneration pathways of RNase A further. Thus far, two three-disulfide intermediates, which form after the rate-determining step in the regeneration process, have been identified and characterized (Rothwarf & Scheraga, 1991; Talluri et al., 1994; Li et al., 1995; Rothwarf et al., 1995). It is of prime interest to determine the identities, importance, and even conformations/structures of different disulfide intermediates in order to obtain a complete understanding of the regeneration pathways of RNase A.

In this work, investigations have been carried out to characterize the one-disulfide intermediates (1S) involved in the regeneration of RNase A. Information about the distribution of the one-disulfide intermediates can provide important insight into the early stages of the regeneration process. Since RNase A contains 8 cysteines, there are 28 possible one-disulfide intermediates. Previous regeneration experiments suggest that a large number of one-disulfide intermediates are involved, since a number of distinct peaks were observed in the ion-exchange chromatogram for the one-disulfide grouping (Rothwarf & Scheraga, 1993a). The one-disulfide grouping is well-separated from other groupings, allowing for its separation as an ensemble from other disulfide intermediates. In the present work, we have isolated the blocked one-disulfide intermediates as a whole group using ion-exchange chromatography. The mixture of one-disulfide intermediates was then subjected to a tryptic−chymotryptic digestion. The digestion fragments were fractionated using reversed-phase HPLC and identified using MALDI-TOF mass spectrometry. The relative population of each one-disulfide intermediate was determined using a disulfide-specific postcolumn detection system. Using these techniques, we have determined the distribution of the one-disulfide intermediates involved in the regeneration process. We have also calculated the one-disulfide distribution from loop entropy considerations. These results and their significance for the regeneration process of RNase A are discussed.

MATERIALS AND METHODS

Materials. RNase A (type 1-A, Sigma Chemical Co.) was purified as described previously (Rothwarf & Scheraga, 1993a). Ultrapure DTT^{red} was obtained from Boehringer-Mannheim, and DTT^{ox} (Sigma) was purified by the method of Creighton (1977). AEMTS was synthesized as described by Bruice and Kenyon (1982). Trypsin (type III, from bovine pancreas) and α-chymotrypsin (type II, from bovine pancreas) from Sigma were used without further purification. All other reagents were of the highest grade commercially available and were used without further purification.

Reduction of RNase A. Fully reduced RNase A was obtained by adding 150 mg of the native protein to 10 mL of 100 mM DTT^{red} solution in a buffer containing 100 mM Tris, 2 mM EDTA, and 4 M GdnSCN at pH 8.0 and 25 °C and maintained in this buffer for 4 h. The mixture was then desalted using 100 mM acetic acid on a Sephadex G-25 column with continuous sparging of argon. The reduced protein was immediately frozen and lyophilized. The resulting protein was reconstituted into 100 mM acetic acid at a concentration of about 10 mg/mL and stored at −70 °C. The purity of the reduced protein was checked by blocking an aliquot of the reduced protein with a 50−100-fold molar excess of AEMTS at pH 8.0, followed by ion-exchange HPLC analysis using procedures described in the following section. No oxidized protein was detected even after 6 months of storage.

Regeneration and Isolation of One-Disulfide Intermediates. The regeneration of one-disulfide intermediates was initiated by mixing 1 volume of the reduced protein (∼10 mg/mL) with 10 volumes of DTT^{ox} in a buffer of 100 mM Tris and 2 mM EDTA at pH 8.0 that had been sparged with argon. The final concentrations were 73 μM RNase A and 20 mM DTT^{ox}. The reaction was typically quenched at 90 min with a 100-fold molar excess of AEMTS over protein in 100 mM Tris and 2 mM EDTA at pH 8.0. These conditions were chosen to obtain a good yield of one-disulfide intermediates while the formation of other disulfide species was suppressed (Rothwarf & Scheraga, 1993a,b). After 5−10 min, the pH of the solution was adjusted to 5.0 to facilitate subsequent chromatographic analysis and to ensure stability during storage (Rothwarf & Scheraga, 1993a). Protein which was not immediately injected onto the HPLC system was stored frozen at −70 °C until it was used. The regeneration process was also quenched at 50 min to see if the distribution of one-disulfide intermediates varied with regeneration time.

The one-disulfide intermediates were separated from the other intermediates and isolated as a group by using cation-exchange HPLC at room temperature.

The HPLC system was model SP8700 from Spectra Physics. The column was a Hydropore-5-SCX column (4.6 × 100 mm) from Rainin. The detector was an ISCO UA-5 with a type 9 optical unit and a 280 nm filter. A ternary gradient was used. Buffer A contained 25 mM HEPES and 1 mM EDTA (pH 7.0). Buffer B contained 25 mM HEPES, 1 mM EDTA, and 1 M NaCl (pH 7.0). Buffer C contained 50 mM acetic acid (pH 5.0). To separate the salts (e.g. AEMTS and DTT) from proteins, 100% C was run from 0 to 15 min, and then buffer C was changed to 0%, buffer A to 88%, and Buffer B to 12% within 2 min. The system was equilibrated under these conditions for 10 min. A NaCl gradient from 12 to 17% B over 40 min (at 0% C) was then run to elute different disulfide species. The fractions corresponding to the one-disulfide species from different runs were collected together. The mixtures of one-disulfide intermediates were dialyzed against 100 mM acetic acid for 24−36 h in a 4 °C cold room. The resulting protein solution was divided into several aliquots and stored in a −70 °C freezer. The one-disulfide intermediates appeared as several partially separated peaks on the ion-exchange chromatogram. They were collected as subfractions to facilitate the identification of these species. These subfractions were dialyzed and stored in the same manner as the whole one-disulfide (1S) grouping. The overall recovery of one-disulfide species is typically greater than 95%. However, since we are comparing the relative population of 1S species, overall recovery is not as important as avoiding the preferential loss of any particular species. Given that the intermediates are for the most part disordered, it is unlikely that there would be significant differences in the chromatographic recoveries of the various one-disulfide species.

Proteolytic Digestion. Both blocked one-disulfide intermediates and the fully reduced and blocked RNase A were lyophilized and reconstituted into 100 mM Tris buffer at pH 8.0 prior to digestion. The typical protein concentration was approximately 1 mg/mL. A combination of trypsin and chymotrypsin digestion was carried out to ensure cleavage between the eight cysteines. The enzyme/substrate ratio was about 1/75 to 1/100 for each enzyme. A typical digestion was carried out by adding trypsin for 1 h and then an equal amount of chymotrypsin for an addition 1 h at room temperature. The digestion was quenched by adjusting the solution pH to 2 using 5% TFA. The digested sample was stored at −70 °C. Both the one-disulfide mixture and the fully reduced and blocked RNase A were digested in parallel using the same enzyme solutions to minimize any difference in digestion conditions.

Peptide Mapping. The digestion fragments were analyzed using a reversed-phase HPLC system. This system consisted of an LKB 2149 pump system and a Gilson Model 116 variable wavelength detector. The detecting wavelength was set at 215 nm. A minibore column (ODS AQ 120A S3, 2 × 150 mm) from YMC was used to separate peptide fragments. The flow rate was 200 μL/min. A binary gradient was used. Buffer A contained 0.095% TFA, and buffer B contained 50% acetonitrile with 0.085% TFA. In the first 10 min, 6% B was run isocratically, and then an acetonitrile gradient was run from 6 to 65% B in 100 min. All peptide fragments eluted within this gradient range.

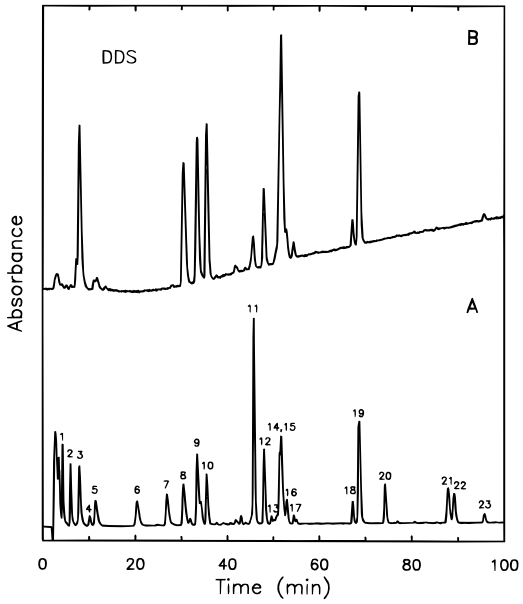

FIGURE 1: HPLC chromatogram of the tryptic−chymotryptic digestion of fully reduced and blocked RNase A detected at 215 nm (panel A). Labels for digestion peaks are the same as in Table 1, and their identities are given in Table 1. Panel B shows the chromatogram detected by the DDS.

The fractionated digestion fragments were collected and analyzed further using a MALDI-TOF mass spectrometer from Finniganmat (Lasermat model 2000). The masses obtained using this system have an error of ±0.1%. The masses of the digested fragments were then matched to the protein sequence to deduce their identities. Most fragments could be identified easily as shown in the Results. However, in some cases, more then one fragment had the same mass, and their identities could not be assigned unambiguously. In addition, species with masses less than 400 were obscured by the large matrix peaks and were generally not detected. In these cases, fragments were identified by amino acid analysis using the Cornell Biotechnology Analytical Facilities. A few short peptides containing only one or two amino acid residues coelute with salts at early times on the reversed-phase column. Since they did not contain cysteine residues, they were not identified.

Disulfide Bond Analysis. A disulfide-detecting system (DDS) was coupled directly to the HPLC system to identify and quantitate the disulfide bond-containing fractions. The details of the DDS were described by Thannhauser et al. (1985). The only difference here is that the effluent from the HPLC column was not split but passed through the UV detector and the DDS successively. The signals detected by both the DDS and UV absorbance were digitized and stored on a Sun IPC computer. Two chromatograms corresponding to UV detection and the DDS were obtained. From the DDS chromatogram, the disulfide-containing fractions in the UV detection chromatogram could easily be identified. An example is shown in Figure 1 where chromatograms monitored by both UV detection at 215 nm and the DDS are displayed for the digestion products of fully reduced and blocked RNase A. It is very easy to recognize which fractions in panel A contain disulfide bonds. In the fully reduced and blocked RNase A, each of the eight cysteines is blocked with AEMTS, resulting in formation of a new disulfide bond at each cysteine. Therefore, several

disulfide-containing fragments, corresponding to cysteine-containing fragments, were observed.

The DDS also provides a quantitative measure of the relative concentrations of disulfide-containing species, since the peak area depends linearly on the disulfide bond concentration. This allows for a direct comparison of the relative population of each one-disulfide intermediate in the whole 1S mixture. In the determination of the relative concentration of one-disulfide intermediates, the digestion procedure and chromatographic analysis were repeated three times. The standard deviation was found to be less than 5%. As an additional check, the areas of all of the peaks in the DDS chromatogram (corresponding to all the disulfide-containing species) of the digestion mixture of the one-disulfide intermediates were summed. The resulting sum was compared to the total area for the one-disulfide species (excluding the AEMTS-containing species) multiplied by 7, and both values were found to agree within 5% (in an AEMTS-blocked one-disulfide intermediate, the total concentration of all of the one-disulfide species should be one-seventh of the total disulfide concentration, since each protein molecule contains one intramolecular disulfide bond and six disulfide bonds from AEMTS-blocked cysteines).

Reducing and Reblocking of Disulfide-Containing Fragments. To verify the identities of the disulfide-containing fragments, each peptide fraction from the HPLC chromatogram was reduced with DTT^{red} and reblocked using AEMTS. These subfragments, obtained after reducing and reblocking, were then reanalyzed using HPLC and MS. The identification of these subfragments was made from their HPLC retention times and verified by MS. Therefore, the identities of any disulfide-containing fragments could be assigned unambiguously and confirmed. Typically, a peak was collected, lyophilized, and then reduced by addition of a 10 mM DTT^{red} solution in 100 mM Tris with 2 mM EDTA at pH 8.0. After about 30 min, an equal volume of 100 mM AEMTS solution in the same buffer was added to block the reduced species. The pH of the solution was adjusted to 5 using 0.1 N HCl after 5−10 min. The completion of both reduction and blocking processes was determined using reversed-phase HPLC. It must be noted that the concentration of AEMTS must be in a large excess compared to that of DTT^{red} to ensure completion of the reblocking process.

RESULTS

Formation of One-Disulfide Intermediates. The kinetics of formation of the one-, two-, three-, and four-disulfide intermediates as well as those of the reduced and native species have been studied in great detail in experiments on the regeneration of RNase A (Rothwarf & Scheraga, 1993a,b). Using these kinetics parameters and the DTT^{ox} concentration dependence, we were able to select a few conditions under which the formation of 1S is maximized while only a small amount of two- and three-disulfide intermediates is formed. A typical ion-exchange chromatogram is shown in Figure 2. The peak identities have been assigned previously (Rothwarf & Scheraga, 1993a) and are labeled on the figure. The yield is about 50% for 1S and 10% for 2S. The 1S species were isolated using chromatographic conditions under which 1S and 2S are well-separated. The peak profiles for 1S at the two different regeneration times that we used (90 and 50 min) are almost identical, suggesting that the one-disulfide distribution does not change with regeneration time.

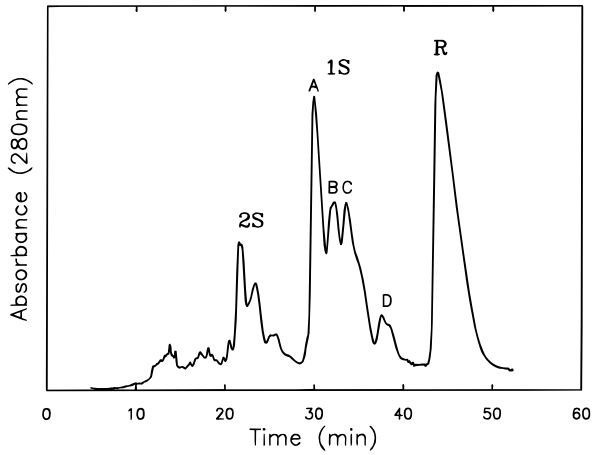

FIGURE 2: Ion-exchange chromatogram of the regeneration of RNase A (with 20 mM DTT^{ox} and 73 μM RNase A at pH 8.0 and 25 °C) quenched and blocked at 90 min. R, 1S, and 2S represent the AEMTS-blocked reduced, one-, and two-disulfide intermediates, respectively. The labels A−D represent subfractions of the one-disulfide intermediates.

In Figure 2, it is apparent that the 1S peak profile is not uniform. Four partially resolved peaks labeled A−D are clearly distinguishable. We collected these as subfractions to simplify identification. However, these subfractions do not necessarily represent pure species, because different one-disulfide intermediates could elute at the same retention time. From the subfractions of such chromatograms, we cannot determine the relative concentration of each one-disulfide species accurately since they are only partially resolved (but, see below for the proper method to identify the one-disulfide species using the *whole* 1S mixture).

Tryptic and Chymotryptic Digestion of Fully Reduced and Blocked RNase A. No tryptic−chymotryptic map has been reported previously for AEMTS-blocked RNase A. This map is necessary to simplify the assignment of the 1S digestion map, since the 1S digestion map contains all the fragments from the digestion of fully reduced and blocked protein in addition to the one-disulfide fragments. This map will therefore be used as a standard to help identify the one-disulfide fragments. We have carried out the digestion both for the fully reduced and blocked RNase A and for the blocked one-disulfide intermediates, under the same conditions, ensuring that the two maps can be compared directly.

Figure 1 shows a tryptic−chymotryptic map for fully reduced and blocked RNase A. The bottom curve (panel A) is a chromatogram obtained by UV detection, and the top curve (panel B) is the corresponding DDS chromatogram. In the fully reduced and blocked protein, only AEMTS-blocked cysteines contain disulfide bonds. Therefore, the peaks in the DDS chromatogram represent peptide fragments which contain cysteine residues. This information is important in fragment assignment. The identity of each fragment was determined from their masses, and the results are summarized in Table 1. The peak numbers in Table 1 correspond to the labels for the peaks in Figure 1. Almost all of the fragments could be assigned, and most of them resulted from specific cleavage at the expected trypsin and chymotrypsin cleavage sites. The cysteine-containing fragments were confirmed by the DDS and are marked with an * in Table 1. Only two significant atypical cleavage sites

Table 1: Tryptic−Chymotryptic Digestion Fragments of Fully Reduced and Blocked RNase A[a]

peak	observed mass[b]	fragment[c]	expected mass[d]
1	662	99−104	662
2	719	1−7	718
3	611	62−66*	610
4	ND[e]	121−124	318
5	475	34−37	475
6	397	74−76	397
7	451	8−10	451
8	830	26−31*	830
9	875	67−73*	875
10	770	80−85*	771
11	1573; 933	11−25; 92−98*	1573; 934
12	1525	86−98*	1524
13	867	1−8	866
14	884; 807	40−46*; 92−97*	884; 806
15	731	105−110*	731
16	1397	86−97*	1396
17	1090; 1156	77−85*; 38−46*	1089; 1156
18	1291	105−115*	1292
19	1692	47−61*	1690
20	971	116−124	970
21	1159	111−120	1159
22	1532	111−124	1531
23	1871	105−120*	1871

[a] The peak numbers are the same as the labels in Figure 1A. [b] Masses are determined from MALDI-TOF MS. [c] Assignments for fragments with the observed masses. Fragments with an * contain cysteine residues as also shown on the DDS chromatogram in Figure 1. [d] Calculated molecular masses for the corresponding peptide fragments. [e] This fragment was not detected by MS; its identity was determined from amino acid analysis.

were observed. One is between Met79 and Ser80, resulting in an 80−85 fragment (peak 10). This fragment was observed in previous tryptic−chymotryptic digestions of RNase A (Smyth et al,, 1963; McWherter et al., 1984) and was confirmed here again using amino acid analysis. The other atypical cleavage was observed after blocked cysteine 110 (peaks 15, 21, and 22). This is possibly due to the similarity between the side chains of lysine and AEMTS-blocked cysteine (2-aminoethylthiocysteine). Curiously, such behavior was not observed at a significant level for other blocked cysteines. Different cleavage rates have been reported at 2-aminoethylthiocysteinyl residues (Walsh, 1970) and in an AEMTS-blocked peptide fragment (50−79) of RNase A (Milburn & Scheraga, 1988).

The map in Figure 1 bears a large degree of similarity to the tryptic−chymotryptic map of carboxymethylated RNase A (McWherter et al., 1984), although different digestion times were required. We have found that the digestion of the AEMTS-blocked protein is much faster than that of the carboxymethylated protein. This could be due to differences in the net charges on the proteins. At pH 8, the native protein has a net positive charge of about +4. Reduction and blocking with a carboxylmethyl group results in a protein with a net charge of about −4, while blocking with AEMTS results in a net charge of about +12. It is possible that the additional net charge on the protein causes it to adopt a more extended conformation, making cleavage sites more accessible to proteolytic enzymes.

Tryptic and Chymotryptic Digestion of Blocked One-Disulfide Intermediates. The mixture of one-disulfide intermediates was subjected to trypsin and chymotrypsin digestion under the same conditions as for fully reduced and blocked RNase A. Ideally, in comparison with the standard

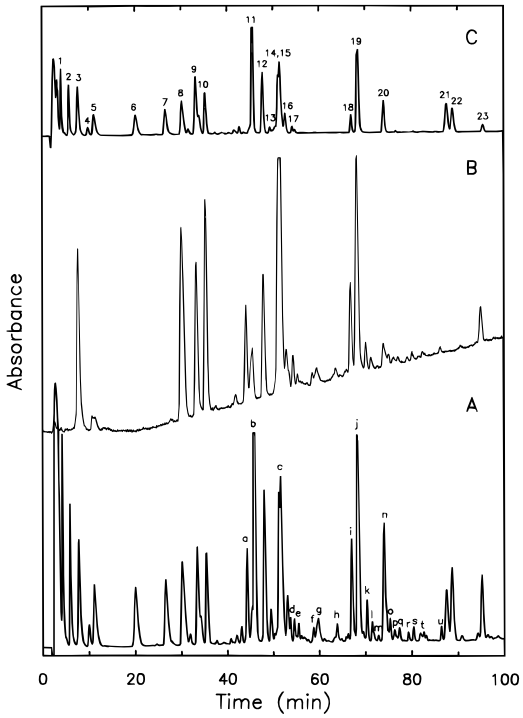

FIGURE 3: Chromatograms of the tryptic−chymotryptic digestion of the blocked one-disulfide intermediates in the regeneration of RNase A. Panel A shows the chromatogram detected at 215 nm. Peaks containing one-disulfide fragments are labeled alphabetically, and their identities are given in Table 2. Panel B is the corresponding chromatogram from the DDS. Panel C shows the chromatogram of the digestion of fully reduced and blocked RNase A detected at 215 nm (the same as in Figure 1A).

map in Figure 1A, if there is only a single one-disulfide intermediate present, the digestion of this species will result in the appearance of one new peak corresponding to the resulting two disulfide-bonded fragments, and the disappearance of the two peaks that correspond to the fragments containing the individual cysteines which form that disulfide. For a mixture of one-disulfide intermediates, a number of new fragments corresponding to disulfide-containing peptides will appear. Most peaks corresponding to cysteine-containing fragments will remain present, since the cysteine-containing fragments absent in the digestion of an individual one-disulfide species could exist in the digestion of another one-disulfide species. The intensities of these peaks corresponding to the cysteine-containing fragments will decrease. Therefore, by comparison of the digestion chromatogram of 1S to that of reduced and blocked RNase A, the fragments corresponding to the one-disulfide intermediates can be identified.

Figure 3 shows a tryptic−chymotryptic map of the one-disulfide mixture. The chromatogram obtained by UV detection is shown in panel A, and the corresponding DDS chromatogram is shown in panel B. A standard map for fully reduced and blocked RNase A (the same as in Figure 1A) is shown in panel C for comparison. Many new peaks are observed in the digestion of the 1S mixture, and they all contain disulfide bonds as indicated by the DDS. Most peaks, such as peaks d−h and o−u, are small in size. A control experiment was carried out using trypsin and chymotrypsin in the absence of RNase A to determine if any of the observed peaks arise from the digesting enzymes. No peaks were observed in the same gradient range used to elute

the RNase A fragments (data not shown). The observation of a large number of low-intensity disulfide-containing fragments indicates the presence of a large number of one-disulfide intermediates. The digestion fragments from both 1S and fully reduced and blocked RNase A were analyzed using MS. The fragments eluting before 40 min are the same for both blocked 1S and fully reduced and blocked RNase A. All new fragments elute after 44 min.

In Figure 3, peaks containing species not observed in the standard map of fully reduced and blocked RNase A (from the MS analysis) are labeled alphabetically, and their identities are given in Table 2. Peaks b, c, i, j, and n represent one-disulfide peaks which are overlapped with peaks 11, 15, 18, 19, and 20 in the standard map of Figure 3C. Peak a is the only nonoverlapping one-disulfide fragment with an intensity comparable to those of the standard peaks. One important feature in Table 2 is that many species coelute in the HPLC chromatogram as observed by multiple masses from a single peak. Most one-disulfide peaks not only overlap with the peaks in the standard map but also overlap with each other.

To identify all the one-disulfide species, each mass is matched to a table of expected masses of all possible one-disulfide combinations. In most cases, a match can be made and the species can thus be assigned with a disulfide pairing. However, in some cases, a species cannot be assigned unambiguously from its mass alone because several one-disulfide fragments with different disulfide pairings have similar masses or two species eluted at different times display similar masses. To solve this problem, a reducing and reblocking method, described in the Materials and Methods, is used to determine and, in some cases, confirm the assignments. Detailed results will be presented in the next section. The assignments shown in Table 2 are the confirmed results.

Table 2 also reveals that a single one-disulfide intermediate may give rise to more than one disulfide-containing fragment. For example, four different disulfide fragments, as seen in peaks a, b, d, and e, correspond to the disulfide pairing [65, 72]. This is due to incomplete digestion at certain cleavage sites in this intermediate. However, in most cases, the shortest one-disulfide fragment, which corresponds to complete digestion, has the largest population. For example, the one-disulfide fragment [65, 72] in peak a corresponds to the complete digestion product and comprises over 70% of the total population of digestion fragments from that one-disulfide intermediate.

Tryptic−chymotryptic digestion was also carried out on each subfraction of the one-disulfide intermediates (labeled A−D in Figure 2). The HPLC chromatograms of the digestion of each of these subfractions are shown in Figure 4 in panels 1SA, 1SB, 1SC, and 1SD, respectively. The digestion map for the whole 1S mixture (the same as in Figure 3A) is shown at the bottom for comparison. The masses of the digestion fragments of these subfractions were measured, and they are in excellent agreement with those from digestion of the whole mixture (data not shown). Peaks containing the same disulfide species, as in the digestion of the whole 1S mixture, are labeled with the same letters as in the 1S map of Figure 3A. Peaks in the individual chromatograms shown in Figure 4 are labeled only if they contain one-disulfide species as determined by MS. For example, although peaks are observed at the same retention time of peak b in every chromatogram, only in the 1S and 1SA chromatograms do those peaks contain one-disulfide species. In the 1SB, 1SC, and 1SD chromatograms, peak b corresponds to peak 11 of Figure 3C. It is interesting to note that the digestion fragments for each subfraction from

Table 2: Tryptic−Chymotryptic Digestion Fragments of One-Disulfide Intermediates[a]

peak	mass[b]	assignment[c]	expected mass[d]	disulfide[e]	fraction[f]
a	1332	(62−66, 67−73)	1332	[65, 72]	1.0
b	1573; 933	11−25; 92−98	1573; 933		
	1312	(62−73)	1313	[65, 72]	0.22
c	732	105−110	731		
	1448	(26−31, 80−85)	1448	[26, 84]	0.0078
d	1712	(62−66, 67−76)	1711	[65, 72]	0.17
	1983	(62−66, 86−98)	1981	[65, 95]	0.069
e	1691	(62−76)	1691	[65, 72]	0.52
	1491	(67−73, 80−85)	1493	[72, 84]	0.20
	2200	(26−31, 86−98)	2200	[26, 95]	0.18
	1341	(40−46, 62−66)	1341	[40, 65]	0.097
f	1552	(80−85, 92−98)	1552	[84, 95]	0.80
	1612	(38−46, 62−66)	1613	[40, 65]	0.20
g	2141	(80−85, 86−98)	2141	[84, 95]	0.64
	2246	(67−73, 86−98)	2245	[72, 95]	0.31
	1656	(67−73, 92−98)	1654	[72, 95]	0.054
h	2124	(47−66)	2127	[58, 65]	0.60
	2625	(67−76, 86−98)	2623	[72, 95]	0.20
	1424	(80−85, 92−97)	1423	[84, 95]	0.20
i	1292	105−115	1292		
	1561	(26−31, 40−46)	1560	[26, 40]	0.10
	2460	(77−85, 86−98)	2459	[84, 95]	0.028
	1832	(26−31, 38−46)	1832	[26, 40]	0.040
	2255	(40−46, 86−98)	2254	[40, 95]	0.040
j	1692	47−61	1690		
	1878	(38−46, 67−73)	1877	[40, 72]	0.010
	2190	(67−76, 77−85)	2189	[72, 84]	0.011
k	2147	(47−61, 62−66)	2146	[58, 65]	1.0
l	2368	(26−31, 47−61)	2366	[26, 58]	0.69
	1749	(62−66, 105−115)	1749	[65, 110]	0.15
m	1969	(26−31, 105−115)	1968	[26, 110]	1.0
n	970	116−124	970		
	2410	(47−61, 67−73)	2411	[58, 72]	0.49
	2307	(47−61, 80−85)	2307	[58, 84]	0.17
	2792	(47−61, 67−76)	2794	[58, 72]	0.069
	1908	(80−85, 105−115)	1908	[84, 110]	0.0085
o	2664	(86−98, 105−115)	2662	[95, 110]	0.26
p	2015	(67−73, 105−115)	2013	[72, 110]	0.27
	2474	(47−61, 92−98)	2470	[58, 95]	0.18
q	2073	(92−98, 105−115)	2071	[95, 110]	0.22
	2393	(67−76, 105−115)	2391	[72, 110]	0.33
r	2421	(40−46, 47−61)	2420	[40, 58]	0.45
s	2695	(38−46, 47−61)	2695	[40, 58]	0.41
	2624	(47−61, 77−85)	2625	[58, 84]	0.14
t	2297	(38−46, 105−115)	2294	[40, 110]	0.45
	2025	(40−46, 105−115)	2022	[40, 110]	0.45
	2230	(77−85, 105−115)	2227	[84, 110]	0.10
u	2831	(47−61, 105−115)	2828	[58, 110]	1.0

[a] Peak labels are the same as those in the chromatogram of the 1S digestion shown in Figure 3A. [b] The experimentally determined masses in each peak of the 1S digestion mixture. [c] Assignments for fragments with experimentally determined masses. One-disulfide species are represented in parentheses either by two subfragments which form that one-disulfide species with an inter-peptide disulfide bond or by one fragment with an intra-peptide disulfide bond. Species that are the same as observed in the digestion of fully reduced and blocked RNase A are indicated without parentheses. [d] Calculated masses for corresponding species. [e] Disulfide pairings of each one-disulfide species. [f] Fraction of each one-disulfide species within that peak of the 1S digestion chromatogram of Figure 3A obtained as described in the text. Some neighboring peaks are collected together to obtain an accurate fraction of each species, since they are partially overlapped. This was done for peaks d (with the neighboring peak corresponding to peak 16 in the standard map in Figure 3C), n and o, p and q, and r and s.

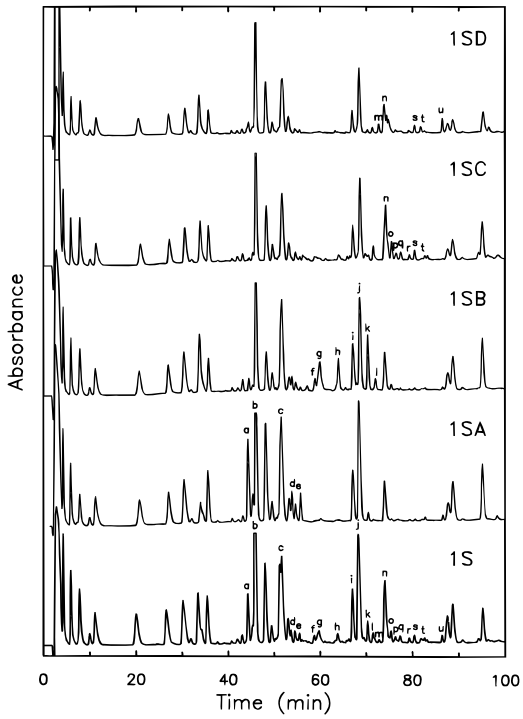

FIGURE 4: HPLC chromatograms with UV detection of the tryptic−chymotryptic digestion of the subfractions A−D of the one-disulfide intermediates in Figure 2. Panel 1S shows the digestion for the whole 1S mixture (the same as in Figure 3A). Panels 1SA, 1SB, 1SC, and 1SD are for subfractions A, B, C, and D, respectively. The alphabetic labels are the same as in Figure 3A, and their identities are given in Table 2.

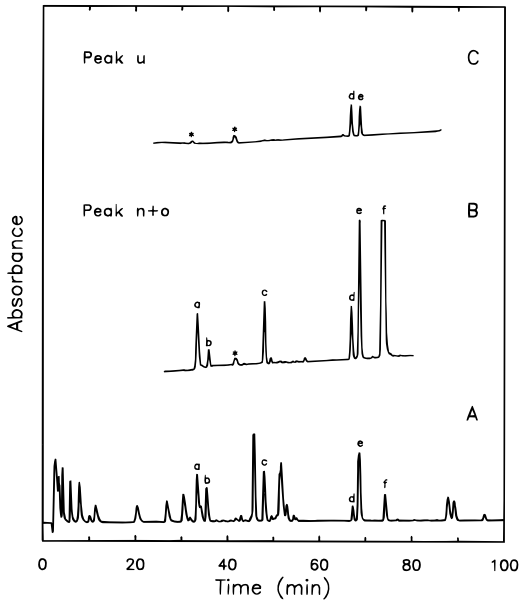

FIGURE 5: HPLC chromatograms of the reduced and reblocked peaks in 1S digestion detected at 215 nm. Panel A shows the standard tryptic−chymotryptic map (the same as in Figure 1A) for comparison. Panel B shows peaks n+o (in Figure 3A) after reduction and reblocking. Panel C shows peak u (in Figure 3A) after reduction and reblocking. Peaks labeled with an * in both panels B and C arise from the mixing of AEMTS and DTT^{red} used in the reduction and reblocking procedure and do not contain peptide fragments. The identities of the labeled peaks are as follows: a, 67−73; b, 80−85; c, 86−98; d, 105−115; e, 47−61; and f, 116−124. These letters label corresponding peaks in *all* three panels.

Figure 2 correspond to a specific set of fragments in the digestion map of the whole mixture. As seen in Figure 4, digestion of subfraction 1SA yields mainly one-disulfide fragments a−e in the 1S digestion map, 1SB gives fragments f−l, 1SC gives fragments n−t, and 1SD gives fragments m, n, and s−u. These data provide confirmation of the results from the digestion of the whole 1S mixture. One advantage of using subfractions from Figure 2 is that the peak intensities for the one-disulfide fragments are much larger than those from digestion of the whole mixture as clearly seen in the chromatograms of the subfraction digestion.

Confirmation of the Identities of One-Disulfide Intermediates. To resolve the ambiguities in the identification of the one-disulfide species from the mass determination, each disulfide-containing fragment obtained from tryptic−chymotryptic digestion was reduced with DTT^{red} and reblocked with AEMTS. There are three possible types of disulfide-containing species present after digestion of the one-disulfide mixture. The first one contains two peptide fragments with an inter-peptide disulfide bond. Most peaks belong to this type. The second one contains only a single peptide fragment with an intra-peptide disulfide bond between two neighboring cysteines. The third one contains a fragment with AEMTS-blocked cysteine. Reducing and reblocking of the first type of fragment will yield two subfragments, each containing an AEMTS-blocked cysteine. From the HPLC retention times of these species, their identities can readily be determined. For the second type of fragment, reduction will form only one species with retention time later than the original species and reblocking will form a new species with retention time earlier than the original one. For the third type, reduction will give a species eluting at a later time than the original species; however, reblocking will yield the original species. In this way, we can distinguish the AEMTS-blocked species from the true one-disulfide containing species. The identities of those fragments after reducing and reblocking were confirmed further by MS.

Figure 5 shows examples of how the reducing and reblocking method works. The chromatograms of peaks n+o and u (from the 1S digestion map in Figure 3A), after reducing and reblocking, are shown in panels B and C, respectively. A standard map of fully reduced and blocked RNase A (from Figure 1A) is aligned at the bottom for comparison. Peaks n and o were analyzed together and shown in panel B since they are partially overlapped in the 1S digestion chromatogram (Figure 3A).

Panel C in Figure 5 clearly shows that reducing and reblocking of peak u gives rise to two subfragments corresponding to peaks d and e in the standard map in the bottom panel. The identities of d and e are fragments 105−115 and 47−61, respectively, from the standard map and are confirmed by their masses. Therefore, peak u contains peptide subfragments 47−61 and 105−115 with an inter-peptide disulfide bond between Cys58 and Cys110. This assignment is also confirmed by the combination of the masses of the two subfragments. The sum of the masses of these two subfragments minus the mass for the blocking groups gives the expected masses in Table 2.

Panel B in Figure 5 shows the reducing and reblocking of peaks n and o. Six fragments corresponding to peaks a−f in the standard map are observed. Fragment f is a non-cysteine-containing fragment 116−124 and is consequently not involved in any one-disulfide fragments. The identities of the other five subfragments are given in the legend of Figure 5. To determine the identities and populations of the

1S pairings in peaks n+o, the masses of all possible pairings are summed (with subtraction of the masses of the blocking groups). These masses are then matched with the masses measured from peaks n and o from the 1S digestion prior to reducing and reblocking (shown in the second column of Table 2). Using this method, each species in peaks n and o was easily assigned as shown in Table 2. This technique provided unambiguous assignments for the species observed in the 1S digestion. The identities of the one-disulfide species in every peak of the 1S digestion chromatogram were assigned and confirmed using this method. A total of 24 different one-disulfide intermediates was identified, as shown in column 5 of Table 2.

The reducing and reblocking method also provides a quantitative measure of the relative concentration of each one-disulfide species in the 1S digestion peaks (Figure 3A) containing multiple species. Since reducing and reblocking will form subfragments containing AEMTS-blocked cysteines, their relative concentrations can be determined from the corresponding DDS chromatogram by simply comparing the integrated areas of the DDS peaks. For each peak in the 1S digestion chromatogram which contains more than one disulfide-containing species, the fraction of each one-disulfide species within that peak was calculated and listed in the last column of Table 2. Peaks b, c, i, and j all contain fragments with AEMTS-blocked cysteines as well as some one-disulfide species. Only the fractions of one-disulfide species are given in Table 2. For example, in peak b, both the one-disulfide species 62−73 and the AEMTS-blocked species 92−98 give rise to a signal in the DDS; the one-disulfide species 62−73 was found to give rise to 22% of the total disulfide signal within that chromatographic peak. The remaining 78% corresponds to the AEMTS-blocked species 92−98. The species 11−25 in peak b does not contribute to the disulfide signal in peak b and is not considered in the calculations. A unit fraction of 1.0 is used for peaks that contain only a single disulfide-containing species, such as peaks a, k, m, and u. In some cases in which the digestion peaks are partially overlapped, they were collected and analyzed together to obtain an accurate relative percentage for each species such as peaks n+o, p+q, and r+s shown in Table 2. Peak d is collected together with the adjacent peak (corresponding to peak 16 in the standard map in Figure 3C) containing AEMTS-blocked cysteine; therefore, the sum of the fractions in peak d of Table 2 is less than 1.0. These data are indispensable for determining the relative population of each one-disulfide intermediate in the one-disulfide mixture.

Distribution of One-Disulfide Intermediates. The relative population of each one-disulfide intermediate was calculated from the DDS chromatogram, since it is directly proportional to the integrated peak area. Therefore, the relative population of a one-disulfide intermediate is obtained by using the disulfide peak areas of that intermediate divided by the total areas for all the one-disulfide intermediates. For any given one-disulfide intermediate, several digested one-disulfide species exist that correspond to the same disulfide pairing. The peak areas for any given one-disulfide intermediate are, therefore, obtained by summing all the areas for that same disulfide species. These values are easily calculated using the fractions for each species in Table 2 multiplied by the corresponding peak areas. The total areas of all the one-disulfide intermediates are the sum of all the one-disulfide areas. The relative populations of each of the one-disulfide intermediates are calculated in this manner, and the results are listed in Table 3. One notable feature in this table is that the one-disulfide intermediate [65, 72] is significantly more populated than any other species, about 40% of the total population. The populations of all other species are less than 10%, with most of them except [58, 65], [58, 72], and [84, 95] being less than 5%.

Table 3: Distribution of One-Disulfide Intermediates in the Regeneration of RNase A

disulfide pairing[a]	percentage[b]	random percentage[c]	ratio[d]
[26, 40]	4.0	5.9	0.7
[26, 58]	2.6	1.7	1.5
[*26*, *65*]	ND[e]	1.3	–
[*26*, *72*]	ND	1.0	–
[**26**, **84**]	1.3	0.7	1.8
[26, 95]	1.1	0.5	2.2
[26, 110]	1.0	0.4	2.5
[40, 58]	2.6	4.1	0.6
[40, 65]	1.4	2.5	0.6
[40, 72]	0.8	1.7	0.5
[*40*, *84*]	ND	1.1	–
[**40**, **95**]	1.2	0.8	1.5
[40, 110]	1.8	0.5	3.6
[58, 65]	9.4	16.6	0.6
[58, 72]	6.6	5.9	1.1
[58, 84]	2.4	2.3	1.0
[58, 95]	0.6	1.4	0.4
[**58**, **110**]	1.5	0.8	1.9
[**65**, **72**]	39.5	16.6	2.4
[*65*, *84*]	ND	3.7	–
[65, 95]	1.2	1.9	0.6
[65, 110]	0.6	1.0	0.6
[72, 84]	2.2	7.4	0.3
[72, 95]	3.0	2.8	1.1
[72, 110]	2.0	1.3	1.5
[84, 95]	8.6	8.5	1.0
[84, 110]	0.3	2.3	0.1
[95, 110]	3.8	5.3	0.7

[a] The disulfide pairing of each one-disulfide intermediate. The one-disulfides in boldface type are the ones with native disulfide pairings. The italicized one-disulfides are those that were not detected experimentally. [b] The experimentally determined relative population of one-disulfide intermediates. [c] Random distribution of one-disulfide intermediates calculated from chain entropy effects as described in the text. [d] The ratio is obtained as the experimentally determined population (column 2) divided by the random values (column 3). [e] ND indicates that the populations for these species were not determined, since they are not detected experimentally. No ratio is therefore available in column 4 for these species.

In order to assess the contribution of specific conformational effects to the stability of the various one-disulfide species, it is useful to compare the experimentally observed distribution of one-disulfide species to the expected random distribution. However, the random distribution is not a uniform one. It depends on the size of the loops created by formation of the disulfide bonds. We have taken this into account and determined the random distribution by calculating the loop entropy involved in formation of all possible one-disulfide species using the following equation (Lin et al., 1984):

$$\Delta S = -R[3.47 + 1.5 \ln(N)] \quad (1)$$

where ΔS is the entropy decrease due to formation of a disulfide bond, R is the gas constant, and N is the number of residues between the C^{α} atoms of the two cysteines. The equilibrium population of any one-disulfide based purely on

entropy considerations can thus be calculated from the following formula:

$$c_i = \frac{c_0}{e^{3.47} N^{1.5}} \quad (2)$$

where c_i is the equilibrium concentration of the one-disulfide intermediate i and c_0 is the concentration of the fully reduced species and is a constant, independent of the one-disulfide pairings. The relative population for each one-disulfide intermediate is thus obtained using the concentration c_i divided by the sum of c_i ($=\sum c_i$) for all the intermediates. The results are listed in the third column of Table 3. We also calculated the random distribution using a similar formula from Harrison and Sternberg (1994):

$$\Delta S = -R[1.73 + 1.5 \ln(N)] \quad (3)$$

which differs from eq 1 in the tolerance volume for ring closure. The results differ from those presented here by less than 1% for all the one-disulfide species.

From Table 3, the one disulfides [58, 65] and [65, 72] display the largest expected random relative population of 16.6%, since the distances between the C^{α} atoms of the cysteines are shortest ($N = 7$). One interesting feature is that, although both [65, 72] and [58, 65] have the same expected population entropically due to equal inter-cysteine residue distances, the actual population of [65, 72] is about 4 times greater than that of [58, 65]. This result suggests the presence of interactions that favor the formation of [65, 72]. The details of these interactions will be addressed in the Discussion.

In order to distinguish the effects of conformational preferences on the equilibrium population of one-disulfide intermediates from the entropic contribution arising from loop entropy, the ratio of the population determined experimentally for each one-disulfide species over the random population is calculated and listed in the last column of Table 3. A ratio greater than 1.0 indicates that the one-disulfide has an above-random population. Likewise, a ratio less than 1.0 indicates that this one-disulfide has a less than random population. The existence of ratios much larger or smaller than 1.0 suggests that interactions exist which favor or disfavor the formation of certain one-disulfide intermediates. Among the 24 identified species, 9 one-disulfide intermediates have ratios significantly greater than 1, 4 of them have ratios close to 1 (values of 1.0 and 1.1), and 11 of them have ratios less than 1. The deviations of the populations from the random values suggest that enthalpic contributions are involved in the formation of different intermediates. This will be discussed in detail below.

DISCUSSION

Effects of Disulfide Rearrangements

One concern about the experimental design is that disulfide rearrangement may occur during the digestion period since it was carried out at pH 8.0. The occurrence of disulfide rearrangement will result in deviation of the measured one-disulfide distribution from the true one. Several experimental results suggest that this is not the case.

First, the digestion was typically carried out for a relatively short time range (trypsin for 1 h and chymotrypsin for an additional 1 h), minimizing the extent of disulfide rearrangement. The digestion was also carried out at pH 6.5, where the rate of disulfide rearrangement is expected to be about 30 times slower than at pH 8.0, solely on the basis of the lower concentration of thiolate ion at pH 6.5. The peptide maps obtained at pH 6.5 and 8.0 are nearly identical (data not shown). MS measurements show that the same fragments were formed at both pH's. These results suggest that no significant disulfide rearrangement occurred during digestion at pH 8.0. We also carried out digestion in the presence of a 100-fold molar excess of AEMTS. Since AEMTS can block free thiols very rapidly, any thiols formed transiently prior to disulfide rearrangement would be blocked by AEMTS. If disulfide rearrangement were important, the digestion maps in the presence and absence of AEMTS would be very different. In fact, both peptide maps were very similar, indicating again that no significant disulfide rearrangement occurred.

An additional piece of supporting evidence is that digestion of each subfraction of the one-disulfide intermediates (peaks A−D in Figure 2) gives some of the peaks observed in the digestion chromatogram of the whole one-disulfide mixture (Figure 4). If significant disulfide rearrangement occurred during digestion, digestion of each subfraction would give a map similar to that of the entire 1S mixture. This was not observed. As seen in Figure 4, only certain sets of peaks in the digestion of the 1S mixture were observed in the digestion of each subfraction. Therefore, it appears that disulfide rearrangement is not significant under the experimental conditions used, and the one-disulfide distribution determined here represents the true distribution.

Interactions Promoting Formation of [65, 72]

It is clearly seen in Table 3 that the formation of the one-disulfide species [65, 72] is strongly favored over that of other one-disulfide species. The large population of [65, 72] is the result of both enthalpic and entropic contributions. Since the [65−72] disulfide loop contains the least number of residues, the entropic contribution to the stability of this one-disulfide intermediate is one of the largest (the other being [58, 65]), as shown in column 3 of Table 3. The measured relative population of [65, 72] is 39.5%, which is more than twice as large as the value expected purely on the basis of loop entropy considerations (16.6%), suggesting that enthalpic interactions contribute to the stability of this disulfide bond. The contributions of enthalpy to the stability of [65, 72] can also be evaluated by comparing it with [58, 65]. As mentioned in the Results, both intermediates are expected to have the same random populations as a result of equal loop size. However, the relative population of [65, 72] is about 4 times larger than that of [58, 65] (see Table 3). In previous studies on a short peptide containing residues 58−72 of RNase A (Altmann & Scheraga, 1990), it was found that the formation of the *native* disulfide bond between 65 and 72 is also about 4 times more favored than that of the *non-native* one between 58 and 65. The similarity between the enhancement for formation of [65, 72] in the short peptide and in the entire protein suggests that the interactions promoting the formation of [65, 72] in both the peptide and protein systems are similar. NMR studies of the short peptide containing the disulfide loop between 65 and 72 indicate that a type II β turn is formed from residues 66 to 69 which stabilizes the disulfide bond between 65 and

72 (Talluri et al., 1993). In native RNase A, a type III β turn from residues 65 to 68 was observed (Borkakoti et al., 1982). It seems likely that, in the one-disulfide [65, 72], a β turn exists which stabilizes the disulfide bond between 65 and 72.

The pK_a's of cysteines 58 and 72 may affect the relative populations of the one-disulfide species [58, 65] and [65, 72]. The details of how the pK_a's of cysteines influence the equilibrium between these two species in a short peptide containing residues 58−72 of RNase A have been discussed (Altmann & Scheraga, 1990). Briefly, at a certain pH, the pK_a of a cysteine will determine the degree of ionization of the thiol group, which in turn contributes to the ability of that thiol to form a disulfide bond. Therefore, if the difference in the relative population of species [58, 65] and [65, 72] is due to the difference in pK_a's of cysteines 58 and 72, a change in pH from 8 to 9 will result in a marked change in their relative populations. This experiment has been carried out for the short peptide 58−72 of RNase A (Altmann & Scheraga, 1990). Those results revealed that the relative populations of these two species did not change with pH. Therefore, it seems that the large difference in the relative populations of the one-disulfide species [58, 65] and [65, 72] does not result from a difference in the pK_a's of the cysteines but rather from specific stabilizing interactions within the 65−72 region as discussed above.

Enthalpic Contributions to the Distribution of One-Disulfide Intermediates

In Table 3, nine one-disulfides were detected with populations significantly greater than the random values (from chain entropy calculations) as indicated by ratios greater than 1. Four are roughly equal to the random values (values of 1.0 and 1.1), and the remaining eleven have populations less than random values, plus four more which were not detected. These data can provide us with important information about which one-disulfide intermediates have favorable interactions that stabilize each disulfide bond. A ratio greater than 1.0 indicates that there is a favorable enthalpic contribution involved in that intermediate. A ratio less than 1.0 suggests that there are some unfavorable interactions that decrease the population of that one-disulfide intermediate.

However, caution must be exercised in interpreting these distribution data. Since the total concentration of one-disulfide intermediates is a fixed number, an increase in the equilibrium populations of some one-disulfide intermediates will necessarily result in a decrease in the equilibrium populations of other one-disulfide species. As a result, the latter species will have below-random populations. We can estimate the random population for these species by applying an average scaling factor obtained from the experimental data. From column 2 of Table 3, the sum of the relative populations for the nine species with above-random population is 52%, while the sum of the expected random values (from column 3 of Table 3) is only 23.3%. The greater than random population of these nine one-disulfide species will reduce the expected random populations of the other one-disulfide species by a factor of 0.6 [$=(1 - 0.52)/(1 - 0.233)$]. Application of this factor brings the random values of those species with below-random populations close to the experimentally observed values. For example, the rescaled random population for [58, 65] is 10% ($=16.6\% \times 0.6$), which is very close to the observed relative population of 9.4%. The purpose of these estimates is to demonstrate that the majority of those species with below-random populations are indeed following the distribution controlled by loop entropy with very little enthalpic contributions. The use of the rescaling factor also results in the four species with close-to-random populations having values that are above random. Therefore, of the 24 observed one-disulfide species, nine have significantly above-random populations, four have above-random populations, and the rest have populations close to random.

It seems that the four species that we did not observe experimentally have relative populations much lower than random, since the rescaled random populations for these species would still be in the detectable range. The absence of these species could also be due to the limitation of the mass spectrometer in detecting very low concentration species. Nevertheless, even if they are present, they should not have large populations and therefore will not markedly affect the observed relative distribution.

Implications for the Folding of RNase A

(a) Refolding Pathway of RNase A. Knowledge of the steady-state distribution of one-disulfide intermediates formed during the regeneration of RNase A provides important information about early folding events, since the one-disulfide intermediates form first during the refolding process. It is now known that two three-disulfide intermediates are formed after the rate-limiting step which leads to the rapid formation of native RNase A (Rothwarf et al., 1995). Each contains three native disulfides missing a disulfide bond between either 40 and 95 or 65 and 72, designated as des-[40, 95]-RNase A and des-[65, 72]-RNase A, respectively (Talluri et al., 1994; Li et al., 1995). The observation of 24 one-disulfide intermediates in the one-disulfide ensemble indicates that the folding from one-disulfide to three-disulfide species must involve the rearrangement of a large number of disulfide bonds. Since no information is yet available about the distribution of the two- and three-disulfide intermediates formed prior to the rate-determining step, we cannot at present provide a clear picture of the refolding pathway.

Studies of the kinetics of the regeneration of RNase A reveal that the one-disulfide intermediates display a steady-state distribution (Rothwarf & Scheraga, 1993a,b), indicating that the distribution of one-disulfide intermediates is thermodynamically controlled. Such a conclusion is consistent with the results presented here; i.e. the distributions of one-disulfide intermediates obtained by quenching the regeneration at 50 and 90 min are similar. Therefore, the one disulfide species [65, 72], which occupies approximately 40% of the total one-disulfide population, is thermodynamically more stable than other one-disulfide species. The stability of [65, 72] is a result of both enthalpic and entropic contributions as discussed above.

From the equilibrium relative population of [65, 72], we estimate that the enthalpic stabilization of this species relative to [58, 65] (the random distribution) is RT ln(39.5/9.4) or approximately 1 kcal/mol. The stabilization of the disulfide bond between 65 and 72 appears to arise from interactions within the disulfide loop since, as discussed above, similar enthalpic stabilization is observed in small peptides contain-

ing the same local sequence. Given that, in the kinetic studies of the regeneration of RNase A, the distributions of the 2S and most of the 3S species (Rothwarf & Scheraga, 1993a,b; Rothwarf et al., 1995) are also thermodynamically determined, it appears likely that the 65−72 disulfide bond will be highly populated in the 2S and 3S intermediates. This should favor the regeneration pathways in which the 65−72 disulfide bond is already formed. While further experiments are needed to confirm such speculation, this is consistent with the experimental observation that about 80% of the regeneration occurs through a pathway containing the three-disulfide intermediate des-[40, 95]-RNase A, which contains the 65−72 disulfide bond (Rothwarf et al., 1995).

(b) Evidence for the Presence of a CFIS. It has been proposed that, in the early stages of folding, there are some "short-range interactions" in a polypeptide chain which form transient local structures. These local structures may play important roles in the early folding process (Tanaka & Scheraga, 1977; Matheson & Scheraga, 1978; Wright et al., 1988; Montelione & Scheraga, 1989). They are termed chain-folding initiation sites (CFISs) (Montelione & Scheraga, 1989). The presence of a CFIS restricts the available conformational space, influencing the subsequent folding events. The results presented here, and NMR studies on the short peptide containing the disulfide bond between 65 and 72 (Talluri et al., 1993), suggest the presence of a turn structure which stabilizes this disulfide bond. An ordered structure, possibly turnlike, is also observed in this same region in an NMR structural determination of the three-disulfide mutant des-[65, 72]-RNase A which is missing the disulfide bond between 65 and 72 (Shimotakahara et al., 1995). The observation of a significant population of [65, 72] is a result of the presence of this CFIS which not only promotes the formation of this disulfide bond but also influences the subsequent folding processes. From the crystal structure of RNase A (Borkakoti et al., 1982; Wlodawer et al., 1988), this reverse turn connects two β strands (residues 61−64 and 71−75) of a four-stranded antiparallel β sheet. Therefore, formation of this CFIS may stabilize the two β strands and subsequently the β sheet region of the native structure.

(c) Roles of Native and Non-Native Disulfide Bonds. Although only nine one-disulfide intermediates were observed to have relative populations significantly greater than the random values, all four one-disulfide intermediates that contain the native disulfide bonds are among them. This indicates the presence of favorable interactions that promote the formation of the native disulfide bonds. This also suggests that the formation of native disulfide bonds is important in the refolding process. The importance of native disulfide bonds in the refolding of proteins has been demonstrated in the study of BPTI (Weissman & Kim, 1991; Dadlez & Kim, 1995).

A total of 20 out of 24 possible one-disulfide intermediates with non-native disulfide bonds were also identified to be involved in the regeneration of RNase A. In particular, two one-disulfide intermediates [26, 110] and [40, 110] have significantly greater than random populations, possibly due to some local interactions. However, the absolute equilibrium populations of these two species are very small (1.0 and 1.8%). Consequently, the significance of these two species to the folding of RNase A is not clear at this point. While the existence of 20 intermediates indicates that non-native disulfide bonds are involved in the refolding process, this result does not necessarily imply that non-native disulfide bonds play any role in directing subsequent folding events.

Comparison with Previous Work

Determination of the distribution of one-disulfide intermediates in the regeneration of RNase A was attempted previously using GSSG/GSH as the redox pair and was deduced to follow a random pattern (Creighton, 1979). However, in that study, only the relative amount of each cysteine involved in a disulfide bond, not the specific disulfide bond pairings, was determined. Therefore, the distribution of one-disulfide intermediates was inferred. In addition, a different blocking reagent, iodoacetate, was used in that study. With slow iodoacetate blocking, the one-disulfide mixture may have had enough time to reshuffle and rearrange to form a close to random distribution. In studies of BPTI, it has been demonstrated that significant disulfide rearrangement can occur when iodoacetate is used as a blocking reagent (Weissman & Kim, 1991). The blocking reagent AEMTS, which we used in the studies presented here, blocks at least 5 orders of magnitude faster than iodoacetate (Rothwarf & Scheraga, 1991). Therefore, the distribution that we have determined for the one-disulfide intermediates more accurately represents the actual distribution.

It should be pointed out that the procedures presented here to determine the distribution of one-disulfide intermediates of RNase A can be applied to any disulfide-containing protein or group of intermediates. In particular, these procedures should be useful in the determination of the disulfide pairings present in the two-, three-, and four-disulfide intermediates formed during the regeneration of RNase A.

SUMMARY

The identities and relative populations of one-disulfide intermediates formed during the regeneration of RNase A have been determined using peptide mapping, disulfide-specific detection, and mass spectrometry. A total of 24 out of 28 possible one-disulfide intermediates have been detected, and the population of these one-disulfide intermediates does not follow the random distribution expected on the basis of loop entropy. The one-disulfide intermediate with a disulfide bond between 65 and 72 is the most populated species, occupying 40% of the entire one-disulfide ensemble. The population of each of the other individual one-disulfide species is less than 10%. The relative population of [65, 72] was also found to be about a factor of 4 larger than that of [58, 65], which has the same random population from loop entropy. The stability of [65, 72] is attributed to the formation of a β turn structure which may also serve as a chain-folding initiation site (CFIS). These results provide important insight into the folding pathways of RNase A and suggest that the interactions that stabilize the disulfide bond between 65 and 72 play an important role in determining the regeneration pathways of RNase A.

ACKNOWLEDGMENT

We thank Drs. Yue-Jin Li and Robert W. Dodge for their valuable suggestions and experimental assistance.

REFERENCES

Altmann, K.-H., & Scheraga, H. A. (1990) *J. Am. Chem. Soc. 112*, 4926.

Anfinsen, C. B. (1973) *Science 181*, 223.
Borkakoti, N., Moss, D. S., & Palmer, R. A. (1982) *Acta Crystallogr. 38B*, 2210.
Bruice, T. W., & Kenyon, G. L. (1982) *J. Protein Chem. 1*, 47.
Creighton, T. E. (1977) *J. Mol. Biol. 113*, 329.
Creighton, T. E. (1979) *J. Mol. Biol. 129*, 411.
Creighton, T. E. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 301−351, W. H. Freeman, New York.
Creighton, T. E., & Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497.
Dadlez, M., & Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 674.
Darby, N. J., Morin, P. E., Talbo, G., & Creighton, T. E. (1995) *J. Mol. Biol. 249*, 463.
Hantgan, R. R., Hammes, G. G., & Scheraga, H. A. (1974) *Biochemistry 13*, 3421.
Harrison, P. M., & Sternberg, M. J. E. (1994) *J. Mol. Biol. 244*, 448.
Houry, W., Rothwarf, D. M., & Scheraga, H. A. (1995) *Nat. Struct. Biol. 2*, 495.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631.
Konishi, Y., & Scheraga, H. A. (1980a) *Biochemistry 19*, 1308.
Konishi, Y., & Scheraga, H. A. (1980b) *Biochemistry 19*, 1316.
Li, Y. -J., Rothwarf, D. M., & Scheraga, H. A. (1995) *Nat. Struct. Biol. 2*, 489.
Lin, S. H., Konishi, Y., Denton, M. E., & Scheraga, H. A. (1984) *Biochemistry 23*, 5504.
Matheson, R. R., Jr., & Scheraga, H. A. (1978) *Macromolecules 11*, 819.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653.
McWherter, C. A., Thannhauser, T. W., Fredrickson, R. A., Zagotta, M. T., & Scheraga, H. A. (1984) *Anal. Biochem. 141*, 523.
Milburn, P. J., & Scheraga, H. A. (1988) *J. Protein Chem. 7*, 377.
Montelione, G. T., & Scheraga, H. A. (1989) *Acc. Chem. Res. 22*, 70.
Pan, H., Barbar, E., Barany, G., & Woodward, C. (1995) *Biochemistry 34*, 13974.
Ptitsyn, O. B. (1995) *Curr. Opin. Struct. Biol. 5*, 74.
Rothwarf, D. M., & Scheraga, H. A. (1991) *J. Am. Chem. Soc. 113*, 6293.
Rothwarf, D. M., & Scheraga, H. A. (1993a) *Biochemistry 32*, 2671.
Rothwarf, D. M., & Scheraga, H. A. (1993b) *Biochemistry 32*, 2680.
Rothwarf, D. M., & Scheraga, H. A. (1993c) *Biochemistry 32*, 2690.
Rothwarf, D. M., & Scheraga, H. A. (1993d) *Biochemistry 32*, 2698.
Rothwarf, D. M., Li, Y.-J., & Scheraga, H. A. (1995) *Protein Sci. 4* (Suppl. 2), 237-S.
Scheraga, H. A., Konishi, Y., & Ooi, T. (1984) *Adv. Biophys. 18*, 21.
Shimotakahara, S., Rios, C. B., Laity, J. H., Montelione, G. T., & Scheraga, H. A. (1995) *Protein Sci. 4* (Suppl. 2), 242-S.
Smyth, D. G., Stein, W. H., & Moore, S. (1963) *J. Biol. Chem. 238*, 227.
Staley, J. P., & Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1519.
Talluri, S., Falcomer, C. M., & Scheraga, H. A. (1993) *J. Am. Chem. Soc. 115*, 3041.
Talluri, S., Rothwarf, D. M., & Scheraga, H. A. (1994) *Biochemistry 33*, 10437.
Tanaka, S., & Scheraga, H. A. (1977) *Macromolecules 10*, 291.
Thannhauser, T. W., McWherter, C. A., & Scheraga, H. A. (1985) *Anal. Biochem. 149*, 322.
van Mierlo, C. P. M., Darby, N. J., & Creighton, T. E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6775.
van Mierlo, C. P. M., Darby, N. J., Keeler, J., Neuhaus, D., & Creighton, T. E. (1993) *J. Mol. Biol. 229*, 1125.
Walsh, K. A. (1970) *Methods Enzymol. 19*, 41.
Weissman, J. S., & Kim, P. S. (1991) *Science 253*, 1386.
Weissman, J. S., & Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9900.
Weissman, J. S., & Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 1123.
Wlodawer, A., Svensson, L. A., Sjölin, L., & Gilliland, G. L. (1988) *Biochemistry 27*, 2705.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167.

BI960090D