Characterizing the Unstructured Intermediates in Oxidative Folding †

ARTICLE *in* BIOCHEMISTRY · JULY 2003

Impact Factor: 3.02 · DOI: 10.1021/bi030054w · Source: PubMed

3 AUTHORS, INCLUDING:

Ervin Welker
Hungarian Academy of Sciences
47 PUBLICATIONS **2,544** CITATIONS

SEE PROFILE

Characterizing the Unstructured Intermediates in Oxidative Folding†

Mahesh Narayan, Ervin Welker, and Harold A. Scheraga*

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

Received March 3, 2003; Revised Manuscript Received April 15, 2003

ABSTRACT: A recently developed method is used here to characterize some of the folding intermediates, and the oxidative folding processes, of RNase A. This method is based on the ability of *trans*-$[Pt(en)_2Cl_2]^{2+}$ to oxidize cysteine residues to form disulfide bonds faster than the disulfide bonds can be rearranged by reshuffling or reduction. Variations of this method have enabled us to address three issues. (i) How the nature of the residual structure and/or conformational order that is present, or develops, during the initial stages of folding can be elucidated. It is shown here that there is a 10-fold increase in the propensity of the unfolded reduced forms of RNase A to form the native set of disulfides directly, compared to the propensity under strongly denaturing conditions (4−6 M GdnHCl). Thus, the unfolded reduced forms of RNase A are not statistical coils with a more condensed form than in the GdnHCl-denatured state; rather, it is suggested that reduced RNase A has a little bias toward a native topology. (ii) The structural characterization of oxidative folding intermediates in terms of disulfide pairing is demonstrated; specifically, a lower-limit estimate is made of the percentage of native disulfide-containing molecules in the two-disulfide ensemble of RNase A. (iii) The critical role of structured intermediate species in determining the oxidative folding pathways of proteins was shown previously. Here, we demonstrate that the presence of a structured intermediate in the oxidative folding of proteins can be revealed by this method.

Oxidative folding is the concomitant formation of the native disulfide bonds and the native tertiary structure from the reduced and unfolded polypeptide (*1*). Much interest has been centered on deciphering the key characteristics of the initial folding stages of unfolded polypeptides and their role in guiding the subsequent stages of the folding process (*2*−*8*).

Fully reduced forms of disulfide-containing proteins serve as analogues for their denatured disulfide-intact counterparts since they can remain unfolded even under "folding conditions" (because of the absence of their disulfides), thereby permitting the measurements that are used to characterize them. Of particular interest is the extent to which *residual structure* or *conformational order* is present in the fully and partially reduced forms of proteins under folding conditions, compared to strongly denaturing conditions. As defined previously, *conformational order* signifies a nonrandom distribution of conformational variables without requiring that the spatial arrangement of the atoms be similar in all conformations of the ensemble (*9*). Thus, a hydrophobically collapsed homopolymer may be conformationally ordered without being structured; its equilibrium ensemble of conformations is characterized by a short radius of gyration (the nonrandom conformational variable) but may include many dissimilar conformations, with large variations in corresponding interatomic distances (*9*).

Bovine pancreatic ribonuclease A (RNase A),[1] a four-disulfide-containing protein, and its mutants, have served as a model for understanding the oxidative folding mechanism of proteins in general (*1*, *9*−*11*). Recent studies of fully reduced RNase A, using double-point mutations of the protein in conjunction with FRET, have revealed a native-like trace of the C-terminal portion of the chain but with a large separation between the N- and C-terminal subdomains in the reduced protein when it is placed under folding conditions (*6*, *8*). Additional studies using NMR (*12*) and small-angle X-ray scattering (*13*) have suggested that, under folding conditions, the reduced form, i.e., R, is not a statistical coil but is partially collapsed with a significant fraction of tertiary topology. Attempts have also been made to detect some degree of native-like character in R using an antigenic response (*14*).

While FRET (*6*, *8*, *15*), in combination with mutagenesis, is very informative in contributing both distance information and a folding hierarchy in almost any region of the protein, whereas the other aforementioned techniques are unable to provide comparable levels of detail, it still remains a laborious task requiring sophisticated instrumentation, and is not applicable over certain distance ranges. Moreover,

† 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.

* To whom correspondence should be addressed. Telephone: (607) 255-4034. Fax: (607) 254-4700. E-mail: has5@cornell.edu.

[1] Abbreviations: RNase A, bovine pancreatic ribonuclease A; *n*S, ensemble of disulfide-containing intermediates each having *n* disulfide bonds; N, native RNase A; R, reduced RNase A; N_R, N_{1S}, N_{R+1S}, and N_{2S}, native protein formed by the direct oxidation of the R, 1S, R + 1S, and 2S, ensembles, respectively; $N_{R+1S+2S+3S}$, definition similar to that for the C[40,95]A mutant; des [*x*, *y*], intermediate of RNase A having all native disulfide bonds but lacking the [*x*, *y*] disulfide bond; BPTI, bovine pancreatic trypsin inhibitor; AEMTS, aminoethylmethylthiosulfonate; DTT^{red}, reduced dithiothreitol; en, ethylenediamine {in *trans*-$[Pt(en)_2Cl_2]^{2+}$}; FRET, fluorescence resonance energy transfer.

10.1021/bi030054w CCC: $25.00
Published on Web 05/16/2003

introduction of point mutations throughout the polypeptide, followed by selective fluorescence labeling, is a cumbersome process in itself. Furthermore, this technique can successfully characterize the unfolded state, but unless mutants are introduced everywhere, it leaves open the question of the ability of a particular conformation to align cysteines correctly to form the native disulfide bonds.

A new procedure (*16*) was recently developed to assess the propensity of the reduced forms of RNase A to form the full set of native disulfide bonds in the absence of reshuffling and reduction reactions (i.e., its native tendency). It should be noted that the native tendency does not necessarily indicate the presence of a population with an overall native topology; however, it provides a quantitative measure of how the amino acid sequence favors the direct formation of the native disulfide bonds over the non-native ones. Here, reduced forms refer to the fully reduced species, and native disulfide-containing intermediates in which more than one disulfide is reduced. The relevant species in RNase A (fully reduced, one- and two-native-disulfide-bonded intermediates) are unstructured (*1*). It should be noted that, in this procedure, the des [*x*, *y*] species containing one pair of half-cystines cannot be oxidized to the scrambled 4S. Thus, their ability to align their cysteines would not affect the amount of native protein that is formed in this experiment. Results from the previous study (*16*) revealed a close to random probability for the reduced species to form the native structure, N. In the study presented here, we use the same procedure to address three questions. The first question pertains to the assessment of the native tendency of the reduced forms of RNase A [which is represented by the ratio of the formed native protein to the fully oxidized non-native (scrambled) species (*16*)] under a variety of conditions to obtain information about the dependence of the native tendency under those conditions. Since the native tendency is influenced by the conformational order in the polypeptide chain, we gain information about the nature of the conformational order present in these forms of RNase A under the given solution conditions.

Therefore, we have used *trans*-$[Pt(en)_2Cl_2]^{2+}$ [which has previously been shown to oxidize cysteines in polypeptides rapidly and selectively to form disulfide bonds (*17*, *18*)] to oxidize the thiols of reduced forms of ribonuclease A and form the terminally oxidized molecules. As demonstrated previously (*16*), by separating the non-native disulfide-containing isomers (4S) from the native molecule (containing all four native disulfide bonds) by HPLC, it is possible to evaluate the ratio of native to scrambled, four-disulfide-containing molecules and, hence, to assess the conformational order present in the aforementioned forms of the polypeptide.

The second question that is addressed in this study concerns the amount of native disulfide pairing in intermediates that are populated on the oxidative folding pathway of RNase A. During the course of the oxidative folding process ($R \rightarrow N$) of any multi-disulfide-containing protein, ensembles of unstructured intermediates containing increasing numbers of disulfide bonds may be formed sequentially (for example, $R \rightarrow R + 1S \rightarrow R + 1S + 2S \rightarrow R + 1S + 2S + ... + nS$) (*1*, *19*, *20*). Within each ensemble of unstructured intermediates, intramolecular thiol−disulfide exchange reactions take place between the protein thiols and existing disulfide(s), leading to an equilibrium distribution of disulfides that is independent of the redox reagent (dithiothreitol). In the absence of stable structure, this equilibrium distribution of disulfide bonds is influenced by both entropic and enthalpic factors (*4*, *21*).

Mapping of these disulfide bonds in combination with a calculated theoretical entropic-driven distribution has proven to be a useful experimental tool for distinguishing between these entropic and enthalpic factors and for gaining insight into the percentage of native disulfide bonds that are present in any ensemble; for example, mapping studies have shown that 40% of the disulfide bonds in the 1S ensemble of RNase A is the native [65−72] disulfide bond (*4*). Mapping, however, does not report on the pairing of disulfide bonds in a particular species of the ensemble. For example, we know that, in the 2S ensemble of RNase A, ∼54% of the disulfides form between cysteines 65 and 72, and 2.2% between cysteines 26 and 84 (*21*). However, we do not know the percentage of molecules in the 2S ensemble that contain both the [65−72] and [26−84] disulfide bonds. Knowledge of such natively paired disulfides is critical for an understanding of the factors involved in the oxidative folding process and is indicative of the native tendency of each relevant unstructured ensemble.

In this study, we show that it is possible to sample each *intramolecularly equilibrated* unstructured ensemble of RNase A sequentially, using *trans*-$[Pt(en)_2Cl_2]^{2+}$, for the propensity to form the native protein (native tendency). The 2S ensemble is the only relevant one in RNase A for estimating native pairing. This is simply because the 1S ensemble contains only one disulfide bond per molecule, and when the 3S ensemble is equilibrated, it reshuffles to form two structured native-like des [*x*, *y*] intermediates (*22*, *23*).

Let us now consider in some detail the propensity of the 2S ensemble of RNase A to form the native molecule (native tendency). Each of the species in the 2S ensemble has two disulfide bonds and four cysteines. Since species that have non-native disulfide bond(s) cannot be oxidized to N under our conditions (no reshuffling), N can form only from those 2S species that contain only native disulfide bonds. However, certain combinations of the remaining four cysteines in these native disulfide bond-containing 2S species result in non-native disulfide bonds and, thus, scrambled molecules. Consequently, the 2S ensemble may contain more species with only native disulfide bonds than is calculated from the amount of N formed. Thus, the propensity to form the native molecule from the 2S ensemble provides a lower-limit estimate for the percentage of molecules within that ensemble that contain only native disulfide bonds (native pairing). Here, we show that it is at least $0.80 \pm 0.15\%$.

Finally, we address a third question, the demonstration of a procedure which helps to elucidate the character of the oxidative folding pathways of disulfide-containing proteins. It was suggested that the pivotal event in oxidative folding is the formation of stable tertiary structure, and hence, the pathways were classified according to the existence and characteristics of structured intermediates (*1*, *24*). In certain proteins, the macromolecule folds conformationally before the formation of the last disulfide bond. When such events occur, the pathway is denoted as a des N pathway, referring to the existence of the structured des [*x*, *y*] species. Examples are pathways involving des [65−72] and des [40−95] species in RNase A (*22*, *23*) and the des [14−38] species of BPTI

(*25*). Apart from having productive (disulfide-secure) intermediates, the des N pathway can also have structured species that are kinetic traps or species that must be unfolded locally before they can be oxidized to form N (*1*). The other pathway, which does not contain any structured intermediates, i.e., the des [*x*, *y*] species is unfolded, is denoted as a des U pathway [with examples being Hirudin (*20*) and stable mutants of RNase A containing three disulfide bonds (*26*, *27*)].

To decipher the nature of the oxidative folding pathway, kinetic fitting of the rate of regeneration as a function of both the redox reagent and protein concentration must be accomplished (*19*). Given the complicated nature of such a study, it is not surprising that there have been no reports in such detail of the oxidative folding pathways of proteins containing more than four disulfide bonds.

Here we demonstrate the efficacy of a simple method of delivering oxidative pulses at varied time intervals and show that it is possible to diagnose whether the regeneration takes place through unstructured intermediates (des U pathway) or through native-like structured intermediates (des N type pathway) and, if so, the stage at which structured intermediates are populated.

When applied to wild-type RNase A and its three-disulfide stable mutant, C[40,95]A, the method successfully reveals these key features, in agreement with previous results obtained from the more laborious kinetic fitting (*22*, *23*, *27*).

EXPERIMENTAL PROCEDURES

Materials. RNase A (type 1-A, Sigma) was purified by cation-exchange HPLC (SP-Sepharose, Amersham) according to the procedure of Rothwarf and Scheraga (*19*). Reduced RNase A (R) was prepared from the purified native protein as described previously (*19*). The folded C[40,95]A mutant of RNase A was obtained from a previous study (*27*) and was reduced as described elsewhere (*27*). DTT^{red} (ultrapure) was obtained from Sigma and used without further purification. The thiol-blocking reagent AEMTS was synthesized by the procedure described by Bruice and Kenyon (*28*). *trans*-$[Pt(en)_2Cl_2]^{2+}$ was prepared by oxidation of *trans*-$[Pt(en)_2Cl_2]$ (*17*) and used directly in solution without further purification. The final concentration of *trans*-$[Pt(en)_2Cl_2]^{2+}$ was estimated from the absorbance at 332 nm (*17*) and, as a check, by reacting it with known concentrations of DTT^{red} at pH 5.

Conformational Ordering in the Reduced Forms of RNase A. Fully reduced RNase A was dissolved (final concentration of 150 μM) in a pH 2, 10 mM glycine buffer containing GdnHCl (0−6 M) and allowed to equilibrate for 10 min under these conditions. The protein was then diluted 10-fold with a solution containing a 10-fold molar excess of the platinum reagent over the free thiols (pH 7, 100 mM HEPES, 5 mM EDTA buffer) and an appropriate amount of GdnHCl (0−6 M in a pH 7, 100 mM HEPES, 5 mM EDTA buffer) to bring the final pH to 7. The concentration of GdnHCl does not vary upon dilution (eq 1a); at the end of this reaction, all of the protein is in either the N or 4S form. After 5 min, the pH was brought to 8 (by the addition of a pH 8 solution containing 1 M Tris-HCl), and a "5-min reduction pulse" was applied to the sample by adding DTT^{red} (final concentration of 10 mM). Under these conditions, only the unstructured 4S species are fully reduced to R (*22*), and the native protein remains unaffected (eq 1b). After an additional 5 min, excess AEMTS was added to block all free thiolates (eq 1c). All experiments were carried out at room temperature, and the samples were desalted on a Sephadex G25 column before injection onto a cation-exchange column (Rainin Hydropore SCX) for separation using HPLC. During the HPLC runs, the salt gradient was increased linearly from 50 to 150 mM over a period of 130 min. AEMTS-blocking introduces a positive charge for each blocked thiol and facilitates the separation of the *n*S intermediates of RNase A (*19*).

$$\text{R (pH 2, } X \text{ M GdnHCl)} \xrightarrow[\text{pH 7}]{\text{Pt Rgt} + X \text{ M GdnHCl}} \text{N} + 4\text{S} \quad (1a)$$

$$\text{N} + 4\text{S} \xrightarrow[\text{5 min, pH 8}]{\text{10 mM DTT}^{red}} \text{N} + \text{R} \quad (1b)$$

$$\text{N} + \text{R} \xrightarrow[\text{cation-exchange HPLC}]{\text{AEMTS}} \underset{(\%\text{R} = \%4\text{S})}{\%\text{N}, \%\text{R}} \quad (1c)$$

Native Tendencies in Intramolecularly Equilibrated 1S and 2S Ensembles of RNase A, and Estimation of Native Pairing in the 2S Ensemble. The 3S ensemble of wild-type RNase A is not considered here because it contains structured intermediates. The species R (890 μM) that had been dissolved in a pH 2 buffer (10 mM glycine), with no GdnHCl present, was diluted into a pH 8 buffer (100 mM Tris-HCl and 1 mM EDTA; final R concentration of 10 μM) containing *trans*-$[Pt(en)_2Cl_2]^{2+}$ such that the concentration of the Pt reagent was much lower than that of the free -SH groups. With variation of the concentration of *trans*-$[Pt(en)_2Cl_2]^{2+}$, it is possible to oxidize R partially in a controlled manner so that all the disulfide-containing ensembles of the protein are sequentially populated (see, for example, eq 2a); i.e., in separate experiments (i−iii), the procedure results in the following independent mixtures at pH 8: (i) R {no *trans*-$[Pt(en)_2Cl_2]^{2+}$ needs to be added}; (ii) R + 1S, and (iii) R + 1S + 2S (at this stage, the types of species present in each experiment and their concentrations are easily checked by blocking an aliquot of each mixture with AEMTS and separating the ensembles of intermediates using HPLC).

$$\text{R} + \textit{trans}\text{-}[Pt(en)_2Cl_2]^{2+} \rightarrow \text{R} + 1\text{S} \quad (2a)$$

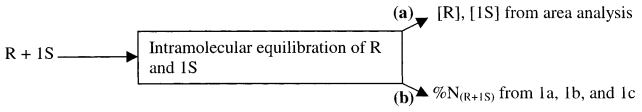
The individual steps required to calculate the native tendency for each ensemble are described below.

(I) First, the native tendency of R, i.e., $\%N_R$, is determined by rapid oxidation of R using *trans*-$[Pt(en)_2Cl_2]^{2+}$ (see eqs 1a−1c; 0 M Gdn and pH 8).

(II) Next, the native tendency in the intramolecularly equilibrated 1S ensemble (i.e., $\%N_{1S}$) is determined as follows. After the mixture containing "R + 1S" (i.e., mixture ii) had been allowed to equilibrate for 10 min at pH 8 (by intramolecular reshuffling), the sample was split equally into two parts, a and b. Free thiols in sample a were immediately blocked using excess AEMTS; the sample was desalted, and separation was carried out using HPLC. The fractional concentrations of both R ([R]) and 1S ([1S]) are determined by area analysis (eq 2b).

The propensity to form N from "R + 1S" (i.e., $\%N_{R+1S}$) is obtained by complete and rapid oxidation of the second

half of the mixture (i.e., part b) to form [N_{R+1S} + 4S], using excess *trans*-[$Pt(en)_2Cl_2$]$^{2+}$ (20× total protein) followed by a reduction pulse with DTT^{red}, blocking of the free thiols with AEMTS, and area analysis using HPLC (eq 2b).



(2b)

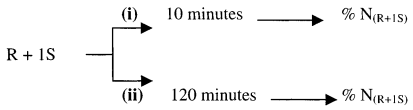
From

$$[R][\%N_R] + [1S][\%N_{1S}] = \%N_{R+1S} \qquad (2c)$$

we obtain the native tendency for the 1S ensemble [$\%N_{1S}$] since all other terms are known, [R] being the fractional concentration of fully reduced RNase A, [1S] being the fractional concentration of the ensemble containing one disulfide bonds, and [$\%N_R$] being the native tendency in R (determined in the previous experiment).

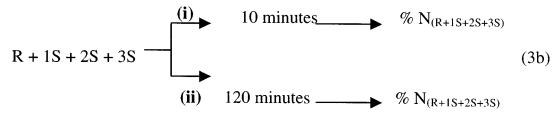
(III) The procedure followed in parts I and II is then repeated for the R + 1S + 2S mixture to obtain $\%N_{2S}$ (i.e., the native tendency in the 2S ensemble).

Procedure for Diagnosing the Oxidative Folding Pathway of Wild-Type RNase A and Its Three-Disulfide Mutant, C[40,95]A. As described above, the following mixtures are generated from reduced wild-type RNase A (10 μM) at pH 8: (a) R, (b) R + 1S, and (c) R + 1S + 2S. Each mixture was then divided into two parts, i and ii, and allowed to equilibrate (intramolecularly; an exception is R which can only achieve conformational equilibrium). To part i was added *trans*-[$Pt(en)_2Cl_2$]$^{2+}$ in a 20-fold excess (over protein concentration) after 10 min, and it was added to part ii after 2 h. After 5 min of each such addition, a reduction pulse was applied to each mixture, followed by blocking, and the samples were desalted and analyzed as described above (for an example of the applications of this method to mixture b, see eq 3a).



(3a)

For the C[40,95]A mutant, a slightly different procedure was followed. Fully reduced C[40,95]A was partially oxidized with *trans*-[$Pt(en)_2Cl_2$]$^{2+}$ (at pH 8) to generate the R + 1S + 2S + 3S mixture (verified by blocking an aliquot of the sample with AEMTS; there is no 4S, since this mutant contains only three disulfide bonds). The sample was then split into two parts, and [$Pt(en)_2Cl_2$]$^{2+}$ was added in a 20-fold excess (over protein concentration) to one part after 10 min and to the other part after 2 h (eq 3a). A reduction pulse was applied to both samples, which were then analyzed as described above (eq 3b).



(3b)

RESULTS

Assessing Conformational Ordering in the Reduced, Unfolded Forms of RNase A. trans-[$Pt(en)_2Cl_2$]$^{2+}$ is able to oxidize cysteines rapidly and specifically to form disulfide bonds. Figure 1 shows the treatment of reduced RNase A (Figure 1A) with the reagent, resulting in the terminal oxidation of the protein to form 4S + N (Figure 1B). The absence of other *n*S ensembles (such as 1S−3S) in the chromatogram indicates that the oxidation of the protein is complete. By integration of the area under the N and R peaks (in Figure 1C), the R peak representing the concentration of 4S, it is possible to calculate %N that is formed [i.e., %N = ($area_N$)/($area_R$ + $area_N$) × 100].

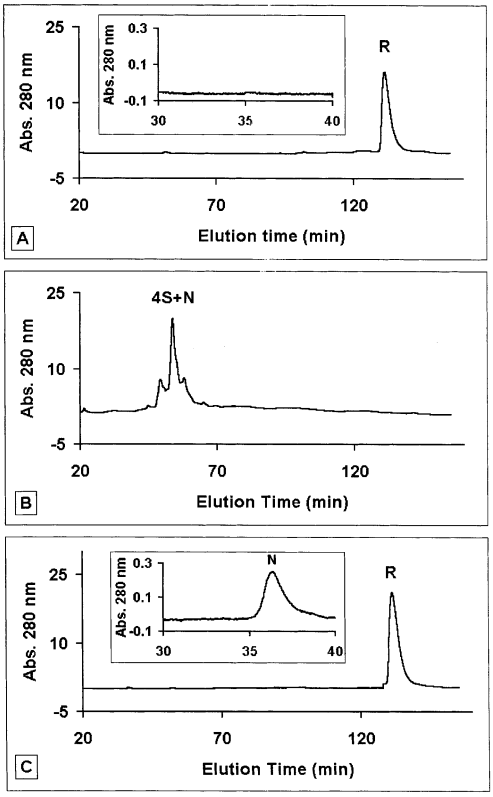


FIGURE 1: Sequential steps involving oxidation of fully reduced RNase A (150 μM) at pH 7 (100 mM HEPES and 5 mM EDTA). (A) Fully reduced RNase A blocked with AEMTS. (B) Addition of excess *trans*-[$Pt(en)_2Cl_2$]$^{2+}$ to the mixture used for panel A followed by blocking of free thiols (if any) with AEMTS. (C) The mixture used for panel B subjected to a 5 min reduction pulse followed by blocking with AEMTS. The inset of panel A shows the absence of N in the mixture containing R. The inset of panel C shows the presence of native RNase A (N) after a reduction pulse had been applied to the mixture used for panel B.

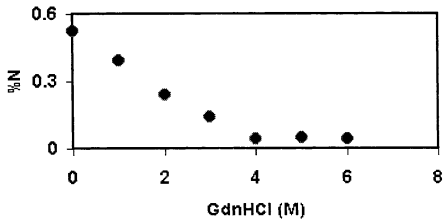


FIGURE 2: Native tendency in R plotted as a function of GdnHCl concentration.

Figure 2 is a plot of %N formed by the oxidation of R by a 10-fold excess of *trans*-[$Pt(en)_2Cl_2$]$^{2+}$ over protein thiols in the presence of increasing amounts of GdnHCl. The data indicate a continuous decrease in %N with increasing concentrations of the denaturant up to 4 M. Between 4 and

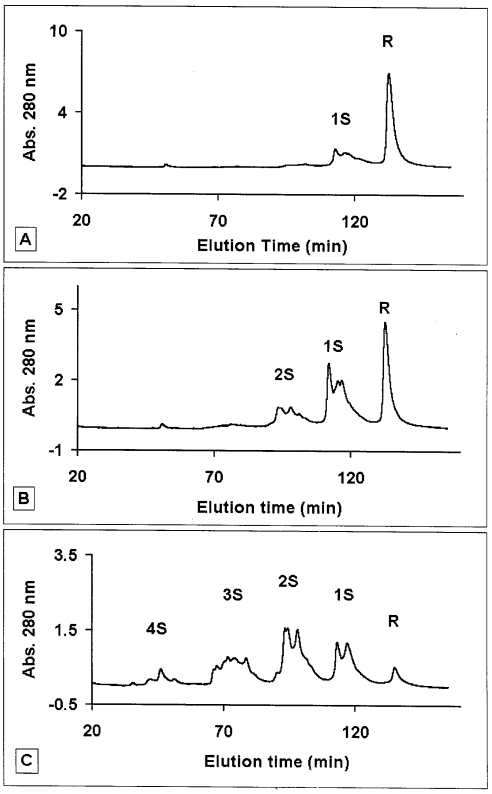


FIGURE 3: Sequential oxidation of R using *trans*-$[Pt(en)_2Cl_2]^{2+}$ to populate each of the *n*S ensembles of RNase A from R (final concentration of 10 μM). In each sample, any free thiols were blocked using AEMTS. (A) R + 1S were generated by adding 2 μM *trans*-$[Pt(en)_2Cl_2]^{2+}$ to the solution containing R. (B) R + 1S + 2S were obtained by adding 4 μM *trans*-$[Pt(en)_2Cl_2]^{2+}$ to the solution containing R. (C) R + 1S + 2S + 3S + 4S were obtained by adding 20 μM *trans*-$[Pt(en)_2Cl_2]^{2+}$ to the solution containing R.

6 M, however, the %N does not change within the error of the experiment. An ~10-fold increase (from ~0.05 to 0.58%) in the propensity to form N when starting from the fully reduced protein is observed under folding conditions (pH 7, 0 M GdnHCl), as compared to more strongly denaturing conditions (pH 7, 4−6 M GdnHCl).

Native Tendencies in the 1S and 2S Ensembles of RNase A. In the previous section, the evaluation of the effect of folding conditions on the native tendency of the reduced forms of RNase A was carried out under conditions where oxidation is much faster than reshuffling. During the course of the oxidation of R (to N + 4S), no intramolecular equilibration of the ensuing *n*S intermediates is achieved.

In this section, we have sequentially evaluated the native tendency (as defined in the introductory section) of the 1S and 2S ensembles of RNase A that have first been allowed to *equilibrate intramolecularly*. It should be noted that the final oxidation from each intramolecularly equilibrated ensemble to 4S (and N) is carried out without allowing for intramolecular equilibration of the subsequent intermediates (i.e., oxidation is much faster than reshuffling).

Figure 3 includes HPLC chromatograms showing the partial oxidation of R using *trans*-$[Pt(en)_2Cl_2]^{2+}$. By carefully adjusting the concentration of this oxidizing agent, we were able to populate the various disulfide-containing ensembles of intermediates in RNase A selectively and sequentially.

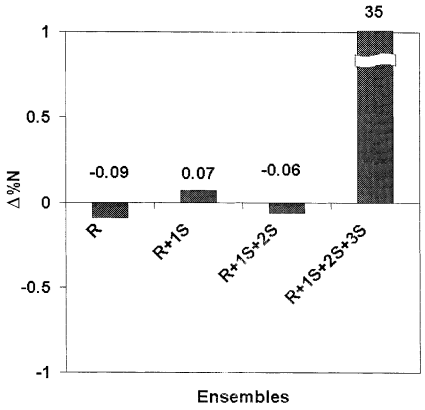


FIGURE 4: Difference between the native tendencies of the ensembles (R, R+1S, R+1S+2S, and R+1S+2S+3S) of RNase A (at 120 min minus 10 min) after they are formed from R using *trans*-$[Pt(en)_2Cl_2]^{2+}$.

Table 1: Propensity of Relevant Unstructured Intermediates of RNase A To Form the Full Set of Native Disulfide Bonds at pH 8 and 25 °C upon Oxidation with *trans*-$[Pt(en)_2Cl_2]^{2+}$

unstructured species of RNase A	$\%N_{nS}$
R	0.62 ± 0.20
1S	0.75 ± 0.30
2S	0.80 ± 0.15

Table 1 shows the native tendency of intramolecularly equilibrated 1S and 2S ensembles of RNase A. [As noted in the introductory section and Experimental Procedures, 3S is not considered because it includes two thermodynamically stable and structured intermediates, the des [x, y] species, to which the remaining isomers in the 3S ensemble gradually reshuffle (*22*, *23*). This would have an overbearing influence on the native propensity of the ensemble.] The native tendency of R at pH 8 is also shown and does not differ from the results obtained at pH 7.

Oxidative Folding Pathway of RNase A Using trans-[Pt(en)$_2$Cl$_2$]$^{2+}$. The oxidative folding pathway of RNase A contains structured disulfide-secure intermediates that improve the efficiency of regeneration of the native molecule by burying the native disulfide bonds that are already formed in a stable structure, thereby preventing them from being reshuffled or reduced (*1*).

Figure 4 shows the difference in %N generated when the intramolecularly equilibrated sets of unstructured ensembles of RNase A obtained using *trans*-$[Pt(en)_2Cl_2]^{2+}$ are sampled successively for the propensity to form the full set of native disulfide bonds over a time interval of 110 min. The graph shows a large difference in %N regenerated when the 3S ensemble is present in the sampling mixture, whereas sampling of other combinations of intermediates at two time intervals results in the same amount of N formed ($\Delta\%N \sim 0$).

Elucidation of the Oxidative Folding Pathway of C[40,95]-A, a Three-Disulfide-Containing Mutant of RNase A. The same method that is described above was applied to a stable three-disulfide mutant of RNase A whose folding pathway has previously been well-characterized in our laboratory using traditional kinetic fitting (*27*, *29*).

Table 2 shows the amount of $N_{(R+1S+2S+3S)}$ obtained by terminal oxidation of these mixtures after 10 min and after 120 min and also the (negligible) difference between them.

Table 2: Propensity of an Intramolecularly Equilibrated Mixture of Intermediates of C[40,95]A To Form N at pH 8 and 25 °C

species	$\%N_{R+1S+2S+3S}$ (10 min)	$\%N_{R+1S+2S+3S}$ (120 min)	$\Delta\%N_{R+1S+2S+3S}$
R+1S+2S+3S	4.2 ± 0.6	5.1 ± 1.5	0.9

DISCUSSION

A Quick Diagnostic for Conformational Ordering in Unstructured, Reduced Forms of Polypeptides. A variety of methods can be used to assess conformational ordering in unstructured polypeptides. For example, in the case of RNase A, NMR (*12*), antigenic response (*14*), FRET (*6*, *8*, *15*), small-angle X-ray scattering (*13*), and dynamic light scattering (*30*) have been applied to assess residual structure in R as well as in 1S−4S ensembles.

Recently, a new procedure for evaluating the propensity of the reduced forms of a polypeptide to form the native structure (the so-called native tendency) was developed (*16*). Here, we used this procedure to evaluate the conformational ordering present in these forms of RNase A under a variety of conditions (0−6 M GdnHCl). As reported previously (*16*), we have used the ability of *trans*-$[Pt(en)_2Cl_2]^{2+}$ to oxidize the free thiols of the reduced forms of RNase A rapidly and efficiently to form an ensemble of terminally oxidized species (4S) on a time scale during which no significant reshuffling or reduction could occur. Among these 4S species (105 being theoretically possible for a four-disulfide-containing protein), there is only one isomer, N, having the full set of native disulfide bonds. By separating N from the 4S isomers [by converting the latter to R, blocking its free thiols using the charged molecule AEMTS, and then separating N and blocked R on a cation-exchange column (*16*)] and calculating the ratio of N to (N + 4S), we were able to determine the extent of N formed from the aforementioned reduced forms under given conditions.

Next, we varied the conditions under which the reduced protein is placed, and reassessed the propensity to form N under the new set of conditions using the same method. The %N that is formed from the reduced and one- and two-native-disulfide-containing reduced forms when subjected to various conditions of interest enables us to compare the "conformational ordering" that is present in the reduced forms under those conditions. *This is simply because the technique exploits the ability of the "conformational order" that is present in the reduced forms of the polypeptide to manifest itself in the ability of the macromolecule to align the relevant pairs of its cysteines correctly and, as a result, form the full set of native disulfide bonds upon oxidation.* For example, as seen here in the case of the reduced forms of RNase A, a change in its conformational ordering, when transferred from folding to fully denaturing conditions (Figure 2), manifests itself as a change in the propensity to form N under those conditions. We then monitor this change, if any, to comment on the conformational ordering of the reduced forms under a variety of conditions. For RNase A, at concentrations of >4 M GdnHCl, there is no further decrease in the ability to form N, indicating that 4 M GdnHCl may be sufficient to "fully denature" the reduced protein. Under our experimental conditions, a 10-fold increase in N is observed when R is transferred from "fully-denaturing" to folding conditions, reflecting a small bias in R toward the native topology, in agreement with previous studies (*6*, *8*, *16*).

The method presented here is relatively simple and appears to be applicable to any unstructured reduced form(s) of disulfide-containing polypeptides. It can also be used to assess conformational ordering under a variety of conditions as in the presence of stabilizing salts such as sulfate and phosphate and other agents of interest such as trifluoroethanol.

Another widely used technique, FRET, has revealed intramolecular details and folding scenarios in reduced proteins, by means of distance measurements, to a far greater extent (*6*, *8*). However, it is a much more labor-intensive and time-consuming technique and is less able to assess the effects of conformational ordering in the context of oxidative folding.

Native Tendencies in Other Intramolecularly Equilibrated Ensembles of RNase A, and the Quantitative Estimation of Species in its 2S Ensemble Containing Two Native Disulfide Bonds. To characterize unstructured intermediates during the oxidative folding process, mapping of the disulfides has been used successfully (*4*, *21*, *29*, *31*−*33*). Previous mapping studies of the disulfide bond distribution of the 1S and 2S ensembles of RNase A have revealed the predominance of the [65−72] disulfide bond over that favored by entropic factors alone (*4*, *21*, *29*). In the absence of protective structure, this bias has been attributed to enthalpic effects. What is not known from mapping studies is the precise pairing of disulfides in unstructured species with more than one disulfide bond, and especially the population of those 2S species containing only native disulfide bonds. These native bond-containing unstructured species can significantly affect the rates of regeneration in certain cases and thus are important players in the oxidative folding process.

By sequentially sampling the intramolecularly equilibrated unstructured ensembles of RNase A, we have been able to assess the native tendency in relevant disulfide-containing ensembles (1S and 2S) of RNase A. It should be noted that the quantitative estimation of the native tendency in R ($[\%N_R]$) includes not only the conformational order present in R but also the conformational order in the populated one- and two-native disulfide species during the oxidation of R to form N. This is simply because R can be oxidized to N **only** sequentially (Figure 5); i.e., the fully reduced protein is first oxidized to form the 1S ensemble, a fraction of which has native disulfide bonds. In the next step, a certain percentage of the native disulfide bond-containing 1S molecules will be oxidized to form native disulfide bond-containing 2S intermediates, with all other molecules forming the remainder of the 2S ensemble (having at least one non-native bond). A fraction of the native disulfide bond-containing 2S molecules will be oxidized to native disulfide bond-containing 3S intermediates. All native disulfide bond-containing 3S intermediates are oxidized to N under conditions where reshuffling is slower than oxidation. Therefore, in our current experimental setup, any conformational ordering present in the native disulfide bond-containing 1S or 2S molecules will count as part of the native tendency in R ($[\%N_R]$). Similarly, the native tendency in 1S ($[\%N_{1S}]$) is influenced by the conformational order present in both the populated one- and two-native-disulfide-containing species. The difference between the native tendency measured in R

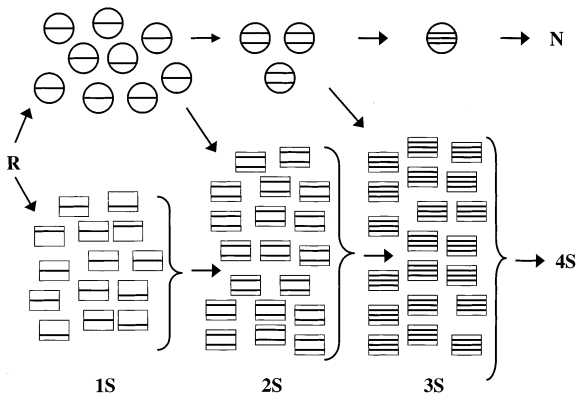


FIGURE 5: Cartoon representation of the sequential oxidation of R to form N and 4S using *trans*-$[Pt(en)_2Cl_2]^{2+}$. Rectangles represent intermediates that contain at least one non-native disulfide bond. Open circles represent species containing *only* native disulfide bonds. The number of bonds in a species is indicated by the horizontal line within the rectangle or circle. Under the given experimental setup, there is no intramolecular equilibration (reshuffling reactions) of molecules within any ensemble or any reduction. Therefore, any 1S species containing a non-native bond is relegated to forming a 4S (non-native) species (sequentially through the rectangles). On the other hand, native disulfide-containing molecules can be oxidized to either non-native bond-containing species or molecules containing only native disulfide bonds (an exception is the native bond-containing 3S population which can form only N).

versus that in 1S (or 1S vs 2S) reflects the difference between the disulfide distribution (i.e., native disulfide content) of the 1S (or 2S) ensemble that was produced by oxidation of *trans*-$[Pt(en)_2Cl_2]^{2+}$ versus that which was produced by equilibrating the ensemble intramolecularly. The native tendency in the 1S species of RNase A, estimated from these studies, is $0.75 \pm 0.30\%$ [as compared to a $\%N_R$ of $0.62 \pm 0.20\%$ (Table 1)]. This indicates no significant difference between the *equilibrium* disulfide distribution of the 1S ensemble and *that distribution* which was transiently populated during the rapid oxidation of R to N (and 4S) by excess *trans*-$[Pt(en)_2Cl_2]^{2+}$. Thus, this result confirmed that the native tendency measured with *trans*-$[Pt(en)_2Cl_2]^{2+}$ is overwhelmingly influenced by the conformational order present in the species and not by the different reactivity of the reagent toward certain cysteines.

The native content of the equilibrated 2S ensemble is especially interesting. As discussed in the introductory section, the native content provides a lower-limit estimate for the number of molecules in this ensemble containing native disulfide bonds. This value of $0.80 \pm 0.15\%$ is not very different from the corresponding values for the native tendencies of R and 1S, indicating no additional bias toward native disulfide pairings in the 2S ensemble of RNase A.

We now discuss a potential source of error in this procedure. For example, consider the measurement involving the native content of the 2S ensemble of RNase A (i.e., $\%N_{2S}$). It is crucial that the 3S species not be present in the mixture containing R, 1S, and 2S. This can be achieved only under conditions in which the concentration of 2S is lower than those of R and 1S. Measurement of the area under the smaller chromatogram peak represented by 2S is then more prone to errors and will result in a discrepancy in the amount of N_{2S} calculated. One way to reduce this effect is to use more protein [but at a low concentration (*40*)] so that the error in estimating areas is reduced. This can be accomplished by using large volumes of dilute solutions.

Deciphering the Pathway of Oxidative Folding. The oxidative folding pathways of a few disulfide-containing proteins have been worked out in detail (*20*, *22*−*24*, *34*−*39*). These proteins regenerate either through unstructured intermediates [a des U pathway (*1*)] or through structured, native-like precursors that are populated during the course of the regeneration [a des N pathway (*1*)].

Studies of the oxidative folding pathways are normally carried out by kinetic fitting of the regeneration rates under varying concentrations of the reduced polypeptide and the redox couple (dithiothreitol or glutathione) (*19*, *34*, *38*, *41*). While such a study is relatively simple for two- or three-disulfide-containing proteins, for a four-disulfide protein such as RNase A, the number of kinetic parameters and fitting becomes quite complicated. This is further compounded if the number of structured des [*x*, *y*] intermediates increases [from two at 25 °C to four at 15 °C in the case of RNase A (*24*)]. To our knowledge, a detailed characterization of the oxidative folding pathway of proteins containing five or more disulfides has not been published. Needless to say, the experimental protocol and modeling only become more difficult.

In the method presented here, we periodically monitor the amount of native protein produced from unstructured ensembles by essentially allowing for the "shielding" of formed native disulfide bonds from intramolecular thiol−disulfide reshuffling reactions by the formation of stable structure. When such shielding is present, it signifies the presence of a structured intermediate(s). Oxidation of these structured intermediates gives rise to a large difference in the amount of N formed at two different times, indicating the presence of a des N pathway (the interval between these two different times should be large enough so that a structured intermediate can form by reshuffling from its unstructured isomers, e.g., 3S to the des [*x*, *y*] species in RNase A). On the contrary, in a des U pathway, there is no strong structural influence to protect native disulfide bonds and, accordingly, the difference between the instantaneous rates of regeneration of N at any two different times is zero ($\Delta\%N \sim 0$).

An important condition that must be met before sampling at any time is that the equilibrium distribution of disulfide bonds within each ensemble must be established. This is necessary to avoid artifacts arising from the generation of the ensemble by an oxidizing agent that may preferentially attack certain thiols in the polypeptide chain over the others.

trans-$[Pt(en)_2Cl_2]^{2+}$ is especially useful in being able to oxidize cysteines rapidly and specifically (*17*, *18*). The reduced form of this compound is inactive, and hence, there is no conversion between ensembles, provided that the protein concentration is sufficiently low (*40*). Moreover, by controlling the oxidizing conditions, one can obtain any desired distribution of oxidized intermediates. While this may not be an equilibrium distribution among ensembles, given enough time for reshuffling, an equilibrium distribution within each ensemble of intermediates can be achieved.

To test this method, the oxidative folding pathway of RNase A, a four-disulfide protein, was re-examined. Using dithiothreitol as the redox reagent, this protein leads to N through the oxidation of two structured three-disulfide intermediates which are formed by reshuffling reactions (formation of the des [*x*, *y*] species from 3S) from their unstructured 3S isomers (a des N pathway) (*19*, *22*, *23*).

By first using less than stoichiometric equivalents of *trans*-$[Pt(en)_2Cl_2]^{2+}$ (compared to protein thiols), we successively populated each ensemble and then applied the above test under strongly oxidizing conditions {using excess *trans*-$[Pt(en)_2Cl_2]^{2+}$}. A significant difference in N between two different times is observed only when the 3S species are present (Figure 4). This indicates that a native-like structured species is formed in this time interval from which N is able to regenerate (des N pathway). The absence of such a difference in the 1S- and 2S-containing samples indicates that there are no structured species in these intermediates. Therefore, we are able to conclude that RNase A regenerates through a des N pathway with at least one structured intermediate populated at the 3S stage.

For further testing of the method, it was applied to the oxidative folding of a previously studied three-disulfide stable mutant of RNase A, C[40,95]A. Kinetic studies (*27*) have indicated that this mutant regenerates through a des U pathway by oxidation of unstructured two-disulfide intermediates to form the native protein (2S → N). In our experiments using *trans*-$[Pt(en)_2Cl_2]^{2+}$, all the unstructured intermediates were populated simultaneously, divided in two and equilibrated for different time periods, and oxidized with *trans*-$[Pt(en)_2Cl_2]^{2+}$. No difference in %N at two sampling times was found (Table 2), indicating the absence of any structured species in any ensemble. Thus, the mutant regenerates through a des U pathway.

In practice, the following steps can be applied to diagnose the nature of an oxidative folding pathway.

Populate all the unstructured intermediates. After intramolecular equilibration has been established within each ensemble, split the mixture into two parts. Rapidly oxidize one part using *trans*-$[Pt(en)_2Cl_2]^{2+}$, and evaluate %N. Oxidize the second part after a certain time interval (say 2 h). Calculate the difference in %N. If this difference is close to zero, then it signifies a des U pathway. If there is a significant difference (i.e., $\Delta\%N \gg 0$), then it is a des N pathway. In such a des N pathway, the ensemble from which the structured species arises can then be found by sequentially evaluating each ensemble by applying the same test.

We now focus attention on certain finer aspects of this method which can potentially cause confusion while analyzing results.

(1) "Significant difference ($\Delta\%N \gg 0$)" is somewhat of a subjective term since we cannot comment precisely on the amount of N formed in a des N pathway of a particular protein. The reason is that we have no idea about the time required to form the structured species from its unstructured isomers. This problem can be overcome by simply sampling the mixture at a longer time interval, thereby allowing the structured species to accumulate by shielding its formed native bonds.

(2) A second potential source for misinterpretation of the results in a des N type of pathway can arise when the *n*S → *n*S* step is very fast (where the asterisk implies a structured isomer). Evaluation of %N {again using excess *trans*-$[Pt(en)_2Cl_2]^{2+}$} at the two times (10 and 120 min) will result in a similar large amount of N since the structured intermediates would have accumulated fully within a few minutes and there are no more unstructured isomers left from which the structured isomer can form. $\Delta\%N$ is then ~0, and the oxidative folding pathway may be incorrectly interpreted as a des U pathway. By sampling the mixture at frequent time intervals, one can resolve this problem.

In conclusion, we have introduced a simple method for evaluating changes in the residual structure in the initial folding stages of multi-disulfide-containing proteins as a function of the folding conditions. Variations of this method have enabled us to provide a lower estimate of the native disulfide pairing in the 2S ensemble of RNase A. Nevertheless, the methods described appear to be applicable to any disulfide-containing protein. Finally, we demonstrate that modifications of our method can be used to detect the presence of structured intermediates in oxidative folding (verified by comparison with previous results from oxidative folding studies of RNase A and its three-disulfide-containing mutant obtained by fitting of kinetic data). Therefore, the major use of this method lies in its application to identify the presence of structured intermediates that accumulate during the regeneration process under conditions in which kinetic fitting becomes too complicated, and an enormous amount of effort is required to collect the data that must be accumulated for such fitting.

REFERENCES

1. Narayan, M., Welker, E., Wedemyer, W. J., and Scheraga, H. A. (2000) *Acc. Chem. Res. 33*, 805−812.
2. Matheson, R. R., Jr., and Scheraga, H. A. (1978) *Macromolecules 11*, 819−829.
3. Kemmink, J., and Creighton, T. E. (1993) *J. Mol. Biol. 234*, 861−878.
4. Xu, X., Rothwarf, D. M., and Scheraga, H. A. (1996) *Biochemistry 35*, 6406−6417.
5. Redfield, C., Schulman, B. A., Milhollen, M. A., Kim, P. S., and Dobson, C. M. (1999) *Nat. Struct. Biol. 6*, 948−952.
6. Navon, A., Ittah, V., Landsman, P., Scheraga, H. A., and Haas, E. (2001) *Biochemistry 40*, 105−118.
7. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) *Science 295*, 1719−1722.
8. Navon, A., Ittah, V., Scheraga, H. A., and Haas, E. (2002) *Biochemistry 41*, 14225−14231.
9. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) *Biochemistry 39*, 4207−4216.
10. Neira, J. L., and Rico, M. (1997) *Folding Des. 2*, R1−R11.
11. Raines, R. T. (1998) *Chem. Rev. 98*, 1045−1065.
12. Denisov, V. P., Jonsson, B.-H., and Halle, B. (1999) *Nat. Struct. Biol. 6*, 253−260.
13. Zhou, J.-M., Fan, Y.-X., Kihara, H., Kimura, K., and Amemiya, Y. (1998) *FEBS Lett. 430*, 275−277.
14. Chavez, L. G., Jr., and Scheraga, H. A. (1980) *Biochemistry 19*, 996−1004.
15. Haas, E., McWherter, C. A., and Scheraga, H. A. (1988) *Biopolymers 27*, 1−21.
16. Narayan, M., Welker, E., and Scheraga, H. A. (2003) *J. Am. Chem. Soc. 125*, 2036−2037.
17. Shi, T., and Rabenstein, D. L. (2000) *J. Am. Chem. Soc. 122*, 6809−6815.
18. Shi, T., and Rabenstein, D. L. (2001) *Tetrahedron Lett. 42*, 7203−7206.
19. Rothwarf, D. M., and Scheraga, H. A. (1993) *Biochemistry 32*, 2671−2679.
20. Thannhauser, T. W., Rothwarf, D. M., and Scheraga, H. A. (1997) *Biochemistry 36*, 2154−2165.
21. Volles, M. J., Xu, X., and Scheraga, H. A. (1999) *Biochemistry 38*, 7284−7293.
22. Rothwarf, D. M., Li, Y.-J., and Scheraga, H. A. (1998) *Biochemistry 37*, 3760−3766.
23. Rothwarf, D. M., Li, Y.-J., and Scheraga, H. A. (1998) *Biochemistry 37*, 3767−3776.
24. Welker, E., Narayan, M., Wedemeyer, W. J., and Scheraga, H. A. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 2312−2316.
25. Oas, T. G., and Kim. P. S. (1988) *Nature 336*, 42−48.

26. Iwaoka, M., Juminaga, D., and Scheraga, H. A. (1998) *Biochemistry 37*, 4490−4501.
27. Xu, X., and Scheraga, H. A. (1998) *Biochemistry 37*, 7561−7571.
28. Bruice, T. W., and Kenyon, G. L. (1982) *J. Protein Chem. 1*, 47−58.
29. Wedemeyer, W. J., Xu, X., Welker, E., and Scheraga, H. A. (2002) *Biochemistry 41*, 1483−1491.
30. Noppert, A., Gast, K., Muller-Frohne, M., Zirwer, D., and Damaschun, G. (1996) *FEBS Lett. 380*, 179−182.
31. Thannhauser, T. W., Konishi, Y., and Scheraga, H. A. (1987) *Methods Enzymol. 143*, 115−119.
32. Ruoppolo, M., Vinci, F., Klink, T. A., Raines, R. T., and Marino, G. (2000) *Biochemistry 39*, 12033−12042.
33. Chang, J.-Y., Li, L., and Lai, P.-H. (2001) *J. Biol. Chem. 276*, 4845−4852.
34. Pace, C. N., and Creighton, T. E. (1986) *J. Mol. Biol. 188*, 477−486.
35. Creighton, T. E. (1992) *Science 256*, 111−112.
36. Weissman, J. S., and Kim, P. S. (1992) *Science 256*, 112−114.
37. Ewbank, J. J., and Creighton, T. E. (1993) *Biochemistry 32*, 3677−3693.
38. Roux, P., Ruoppolo, M., Chaffotte, A.-F., and Goldberg, M. E. (1999) *Protein Sci. 8*, 2751−2760.
39. Jarrett, N. M., Djavadi-Ohaniance, L., Willson, R. C., Tachibana, H., and Goldberg, M. E. (2002) *Protein Sci. 11*, 2584−2595.
40. Song, M.-C., and Scheraga, H. A. (2000) *FEBS Lett. 471*, 177−181.
41. Konishi, Y., Ooi, T., and Scheraga, H. A. (1982) *Biochemistry 21*, 4734−4740.

BI030054W