Oxidative Folding of Proteins

MAHESH NARAYAN, ERVIN WELKER, WILLIAM J. WEDEMEYER, AND HAROLD A. SCHERAGA*

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

Received April 18, 2000

ABSTRACT

The oxidative folding of proteins is reviewed and illustrated with bovine pancreatic ribonuclease A (RNase A). The mutual effects of conformational folding and disulfide bond regeneration are emphasized, particularly the "locking in" of native disulfide bonds by stable tertiary structure in disulfide intermediates. Two types of structured metastable disulfide species are discerned, depending on the relative protection of their disulfide bonds and thiol groups. Four generic pathways for oxidative folding are identified and characterized.

Introduction

Disulfide bonds between cysteine residues are a key component of many proteins.[1] Such bonds can stabilize the native structure significantly, both by lowering the entropy of the unfolded state[2−4] and by forming stabilizing interactions in the native state.[5−8] However, the native set of disulfide bonds is the end result of an often complicated process involving covalent reactions such as oxidation, reduction, and reshuffling (Figure 1). The term *oxidative folding* describes the composite process by which a reduced, unfolded protein recovers both its native disulfide bonds (*disulfide bond regeneration*) and its native structure (*conformational folding*). This Account synthesizes a general picture of this process, based primarily on recent data from bovine pancreatic ribonuclease A (RNase A)[9−11] (Figure 2).

Mahesh Narayan received a B.Sc. degree in physics (1991) from Bombay University and a Ph.D. degree in biophysics (1997) at The Ohio State University. He is currently a research associate in the laboratory of Prof. Scheraga. His research interests include protein folding, structure−function relationships in proteins, and the application of milk proteins as carriers for nutritional and pharmaceutical compounds.

Ervin Welker received a B.S. degree in chemistry (1986) and a Ph.D. degree in biology (1997) from the Eötvös Loránd University in Budapest. He is currently a research associate in the laboratory of Prof. Scheraga. His research interests include protein folding, protein semisynthesis, and the mechanism of multidrug resistance proteins.

William J. Wedemeyer received B.S. degrees in music (1988) and physics (1988) from M.I.T. and M.S. and Ph.D. degrees in physics (1993, 1998) from Cornell University. He is currently an NIH postdoctoral fellow of chemistry in the laboratory of Prof. Scheraga. His research interests include computational and experimental studies of protein folding.

Harold A. Scheraga received a B.S. degree in chemistry (1941) from CCNY and Ph.D. (1946) and Sc.D. (honorary, 1961) degrees in chemistry from Duke University. He was a postdoctoral fellow with John Edsall at Harvard Medical School (1946−47) and joined the Chemistry faculty of Cornell University in 1947, where he is presently the George W. and Grace L. Todd Professor of Chemistry, emeritus. His research interests include theoretical and experimental studies of protein folding, hydrophobic interactions, polymer theory, and the mechanism of blood coagulation.

The regeneration of the native disulfide bonds is a critical step in the maturation of many extracellular proteins. (Disulfide bonds are relatively rare in intracellular proteins because the cytosol is a strongly reducing environment.[1]) During protein synthesis, such proteins are threaded into the endoplasmic reticulum (an oxidizing environment[12]) where the native disulfide bonds may form either co- or posttranslationally.[13] Oxidative folding is also an important stumbling block in the preparation of proteins for pharmaceutical and biotechnological applications, particularly multidomain proteins with several disulfide bonds.[14] Controlled in vitro experiments on relatively simple proteins such as RNase A may help to determine the factors affecting the regeneration of the native disulfide bonds both in vivo and for complex proteins.

Stable tertiary structure is one such factor.[11] Oxidation, reduction, and reshuffling are based on a single chemical reaction, *thiol/disulfide exchange*,[15,16] in which a thiolate anion approaches a disulfide bond closely enough to delocalize its negative charge between the three sulfur atoms (Figure 1). Stable tertiary structure can inhibit such contacts between thiolates and disulfide bonds by controlling the accessibility, reactivity, and proximity of these reactive groups. For example, the burial of the disulfide bonds in stable tertiary structure inhibits their reduction and reshuffling, while the sequestering of thiol groups can inhibit their oxidation, i.e., the formation of disulfide bonds. Other structural factors include local electrostatic interactions (which may affect the local concentration, pK_a, and reactivity of the negatively charged thiolate groups) and structural propensities (which may alter the effective concentrations of the reactive groups for each other). The role of stable tertiary structure in the regeneration of the native disulfide bonds will be emphasized in this review. However, the focus of this Account will be on the pathways for regenerating the native disulfide bonds, and not on the implications of disulfide bond experiments for conformational folding, which have been reviewed previously.[11] For brevity, the reader is also referred to that review for the experimental methods used in disulfide regeneration studies.

This Account is organized as follows. The first section introduces some technical terms needed for precise discussions, while the second reviews experimental data on the oxidative folding of RNase A. From these data, the third section distills four generic types of pathways for oxidative folding, depending on the role of tertiary structure and the relative protection of disulfide bonds and thiol groups. The fourth section shows that these generic pathways are applicable to the oxidative folding of other proteins. The conclusion summarizes these results and discusses their relevance for disulfide bond regeneration in vivo and for biotechnology.

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

10.1021/ar000063m CCC: $19.00
Published on Web 08/12/2000

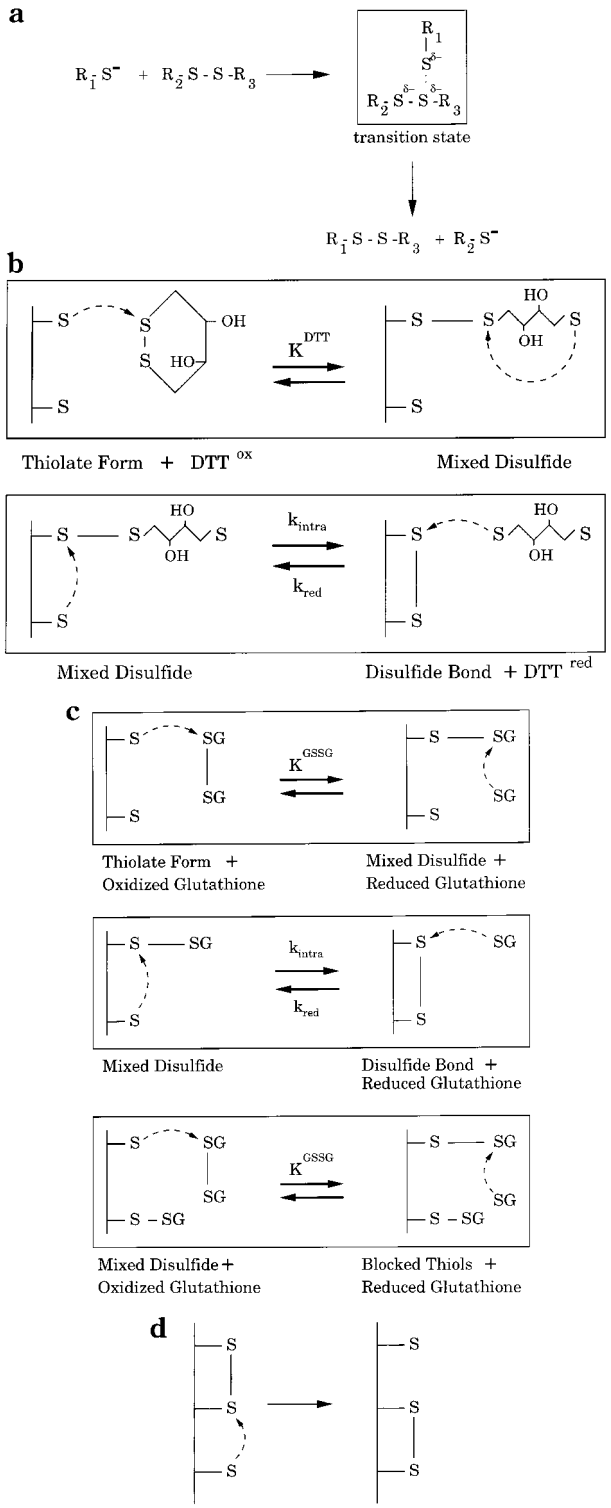

FIGURE 1. Chemistry of disulfide bond reactions. (a) In thiol/disulfide exchange,[15,16] a thiolate anion R_1S^- displaces one sulfur of a disulfide bond R_2SSR_3. In the transition state, the negative charge of the thiolate appears to be delocalized among the three sulfur atoms. (b and c) Protein disulfide bonds are formed and reduced by two such thiol/disulfide exchange reactions with a redox reagent, the first of which involves the formation of a *mixed disulfide bond* between the protein and the redox reagent. For completeness, the reactions are illustrated with cyclic (DTT^{red}/DTT^{ox}) and linear (GSSG/GSH) redox reagents. (d) Thiol/disulfide exchange reactions can also occur intramolecularly; e.g., a protein thiolate group may attack a disulfide bond of the same protein, leading to *disulfide reshuffling*.

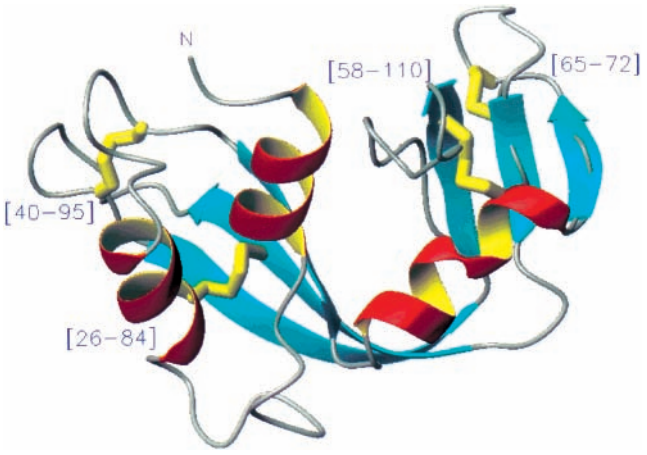

FIGURE 2. Structure of bovine pancreatic ribonuclease A (RNase A),[5] a small but stable protein of 124 residues with four disulfide bonds: 26–84, 40–95, 58–110, and 65–72. The disulfide bonds 58–110 and 26–84 are fully buried in the major and minor hydrophobic cores, respectively, and join well-defined secondary structures (an α-helix and β-sheet in both cases). By contrast, the disulfide bonds 65–72 and 40–95 occur near the surface of the protein in turn regions, although the 65–72 disulfide bond is part of a small hydrophobic core involving residues Val63, Ile107, Ala122, and Val124. The relatively large number of cysteines (eight) makes the oxidative folding of RNase A one of the most complex ever studied; there are 7193 possible disulfide species, if mixed disulfide species are considered. The well-studied conformational folding of disulfide-intact RNase A[9,10] aids in structural interpretations of oxidative folding experiments.

Definitions of Technical Terms

We adopt the following terminology to describe the covalent structure of the disulfide intermediates. A *disulfide species* refers to a protein with a particular pairing of cysteines in disulfide bonds, while a *disulfide ensemble* is any collection of disulfide species. A *des species* is a disulfide species with all but one of the native disulfide bonds; for example, the des[65–72] species of RNase A has three native disulfide bonds but lacks the 65–72 disulfide bond. The term *des* signifies only a covalent property (the absence of one native disulfide bond) and does not specify whether the species is folded. A subscript is added to characterize the conformational state; thus, des_U and des_N represent the unfolded and folded states of a des species, respectively. *n*S represents an ensemble of (typically unstructured) disulfide species with *n* disulfide bonds that are in a rapid reshuffling equilibrium among themselves; for example, 2S represents an ensemble of rapidly interconverting disulfide species with two intraprotein disulfide bonds. The rapid interconversion within such ensembles allows them to be treated as single kinetic species, although they are a heterogeneous mixture of disulfide species.[17] The ensemble of disulfide species with no protein disulfide bonds is denoted as R, the fully reduced ensemble (the multiplicity of R is due to the possibility of mixed disulfide bonds). Fully oxidized, nonnative species are denoted as *scrambled* disulfide species, e.g., those of the 4S ensemble in RNase A. A more general nomenclature of disulfide species (including mixed disulfide species) has been published.[18]

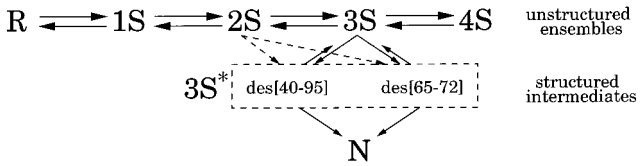

FIGURE 3. Kinetic model of the oxidative folding of RNase A.[19,20] A preequilibrium (quasi-steady-state) is established among the reduced protein R and the unstructured 1S−4S ensembles. Under typical conditions, the native protein is regenerated principally through two pathways, corresponding to the reshuffling of unstructured 3S species to form the structured 3S* species, i.e., des[40−95] and des-[65−72]; these latter two species contain three native disulfide bonds but lack the 40−95 and 65−72 disulfide bonds, respectively. The des[40−95] pathway is the major pathway (∼80%) for regenerating the native protein under typical conditions. A small fraction of these two structured species is also generated by oxidation from the 2S ensemble,[21,22] as indicated by dashed lines; however, these oxidation pathways are significant only in the regeneration of the mutant analogues, [C65S,C72S] and [C40A,C95A].

We introduce here the following terminology to describe the protection of disulfide bonds and thiol groups in structured disulfide species. In *disulfide-protected* species, the reduction and reshuffling of disulfide bonds are inhibited by stable tertiary structure. Such species often accumulate, since disulfide reshuffling reactions are generally much faster than oxidation reactions; thus, the inhibition of reshuffling in these species reduces their rate of disappearance significantly. A subset of disulfide-protected species are the *metastable* disulfide species, in which both the disulfide bonds and thiol groups are protected, inhibiting all disulfide bond reactions (oxidation, reduction, and reshuffling). Thus, a local or global unfolding step must precede disulfide bond reactions in such metastable species. It suffices to bury all but one thiol group in such species, since the remaining thiol group would have no partner with which to react. Two other terms (*disulfide-secure/insecure*) are introduced below to describe the relative protection of the disulfide bonds and thiol groups.

The Oxidative Folding of RNase A

Under typical conditions (pH 8, 25 °C, with DTT^{ox} as the redox reagent), the oxidative folding of RNase A is fit by the kinetic model of Figure 3.[19−22] In the early stages of oxidative folding, the disulfide ensembles 1S−4S are populated by successive oxidations, until a preequilibrium (quasi-steady-state condition) is established among these ensembles. No stable structure has been detected in these ensembles, although they do exhibit some conformational order.[11] The distribution of disulfide species within the 1S and 2S ensembles is enthalpically biased toward native disulfide bonds, relative to the populations predicted from loop entropy considerations alone.[3,23,24] This preferential population of native disulfide bonds accelerates the oxidative folding relative to that of an unbiased ensemble.

The rate-determining step in the regeneration of native RNase A is the formation of two des species, des[40−95] and des[65−72], with native-like structure.[25−28] These species are formed largely by reshuffling from the 3S ensemble,[19,20] although a small fraction (no more than 5%) may be formed by oxidation from the 2S ensemble[21,22] (Figure 3). Upon formation and folding of either des[40−95] or des[65−72], their three native disulfide bonds become protected from reduction and reshuffling ("locked in"), causing these des species to accumulate to much higher levels. However, their thiol groups remain accessible to the solvent and, hence, these species oxidize relatively rapidly to the native protein. These two pathways appear to account for nearly all of the native RNase A regenerated under typical oxidative folding conditions.[20]

Stable conformational structure appears to be essential for this acceleration of oxidative folding. The regeneration becomes drastically slower when carried out under conditions that destabilize the conformational structure of the folded des species.[26,27,29] Moreover, the original rate of regeneration can be restored by adding other agents that restabilize the conformational structure (e.g., phosphate salts).[30] Thus, the protective structure of these des species is the critical factor in promoting oxidative folding.

The oxidative folding of mutant analogues of des[40−95] and des[65−72] (the three-disulfide mutants [C40A,-C95A] and [C65S,C72S]) has also been characterized.[21,22] Consistent with the absence of structured two-disulfide intermediates in wild-type RNase A, neither mutant exhibits any folded disulfide intermediates prior to forming all three native disulfide bonds. In both mutants, the rate-determining step is the formation of the native protein by oxidation from the unstructured 2S ensemble and conformational folding.

The other two des species, des[58−110] and des[26−84], have stable conformational structure only under strongly stabilizing conditions, e.g., in the presence of stabilizing salts[30] or at low temperatures (≤15 °C).[31] Preliminary data suggest that, under such conditions, these des species are metastable dead-end species that reshuffle preferentially to the 3S ensemble rather than oxidizing to the native protein (unpublished data). Presumably, des[58−110] and des[26−84] bury both their thiol groups and disulfide bonds, thus inhibiting oxidation as well as reduction and reshuffling. This burial of thiol groups is structurally plausible, since these des species probably have native-like structures (judging from enzymatic activity measurements[32] and their HPLC elution properties), and the 26−84 and 58−110 disulfide bonds are fully buried in hydrophobic cores of the native protein.

Even mixed disulfide bonds can become buried in stable conformational structure. This is observed in the oxidative folding of RNase A through the covalent adduct N′, in which a DTT molecule cross-links cysteines 65 and 72.[33] Kinetic data indicate that N′ is formed through the burial of a mixed disulfide bond between a DTT molecule and one cysteine, probably Cys72.[25] The remaining thiolate of the covalently bound DTT molecule forms a disulfide bond with a second DTT molecule, which is attacked by the remaining protein thiolate to form the adduct. Such protected mixed disulfide bonds seem more likely to form with linear redox reagents such as glutathione[34] than with rapidly recyclizing reagents such as DTT.

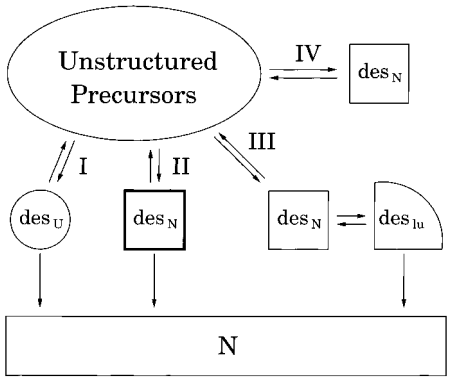

FIGURE 4. Four generic types of oxidative folding pathways for a typical disulfide-containing protein. For simplicity, we assume that only des species are capable of folding to a native-like structure (indicated by the N subscript). The U and lu subscripts indicate that the corresponding species are unfolded or locally unfolded, respectively (although such species may have some conformational order, as observed in RNase A[11]). In pathway I, the native protein N regenerates directly from an unfolded precursor des_U (a des_U pathway). Pathway II involves the formation of a folded intermediate des_N that is not metastable; des_N is oxidized directly to the native state (a des_N pathway). Pathway III likewise involves a folded intermediate des_N that oxidizes preferentially to the native protein but is metastable (a metastable des_N pathway). To become oxidized to the native protein, this metastable intermediate must undergo a local unfolding step (des_{lu}). Finally, pathway IV involves a folded, metastable species des_N that preferentially reshuffles rather than being oxidized to the native protein (a metastable dead-end pathway). The rate-determining steps of pathways I and II are associated with the formation of the structured species, while that of the metastable dead-end pathway IV is associated with conformational unfolding of the metastable species. In this respect, pathway III is intermediate between pathways II and IV; the rate-determining step may be either the formation or the (local) unfolding of the structured intermediate, depending on the conditions and the $\Delta G°$ of local unfolding required to oxidize the metastable species.

Generic Pathways of Oxidative Folding

Based on these oxidative folding studies of RNase A, four generic types of oxidative folding pathways may be identified (Figure 4). These four pathways are distinguished by three factors: whether any disulfide intermediates are folded, whether the folded species are metastable, and whether the metastable species preferentially oxidize or reshuffle. Throughout our discussion, we assume that the experimental conditions favor the folding of the native protein; thus, the full set of native disulfide bonds should be protected from reduction.

The simplest oxidative folding pathway has no structured intermediates (pathway I in Figure 4). This type of pathway is denoted as a *des_U pathway*, since the des species from which the native protein forms (by oxidation) is unfolded. The three-disulfide mutants of RNase A, [C65S,C72S] and [C40A,C95A], follow such a des_U pathway, in that their des species appear to be all unfolded. Since these unfolded (two-disulfide) des_U species are in a rapid reshuffling equilibrium with the unstructured 2S ensemble, these des_U species are preferentially populated only to a level determined by the bias toward the (three) native disulfide bonds in that ensemble. Therefore, such

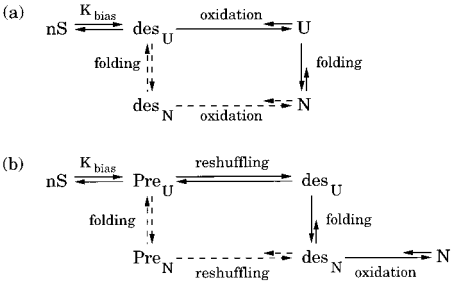

FIGURE 5. More detailed model of the rate-determining step in the des_U and des_N pathways. (a) In the des_U pathway (pathway I of Figure 4), the des_U species are in a rapid reshuffling equilibrium with their *n*S ensemble (K_{bias}). In the rate-determining step, the des_U species are oxidized to U, an unfolded protein with the native disulfide bonds, which then folds conformationally to the native protein N. (b) In the des_N pathway (pathway II of Figure 4), the immediate precursors of the des_U species (denoted Pre_U) are in a rapid reshuffling equilibrium with their *n*S ensemble (K_{bias}). (By definition, these Pre_U species have the same total number of disulfide bonds as the des_U species, with one nonnative disulfide bond replacing a native disulfide bond.) In the rate-determining step, the Pre_U species reshuffle to des_U, which folds conformationally to form the disulfide-protected des_N species, which then oxidizes to the native protein. In both (a) and (b), conformational folding *follows* the formation of a set of disulfide bonds that render the folded state thermodynamically more stable than the unfolded state; in this respect, our model differs from the traditional view,[35] in which these processes are reversed. (The traditional view is indicated by dashed lines.) Experiments are underway to determine the relative contributions of these two mechanisms; however, the generic pathways of Figure 4 are independent of the relative order of conformational folding and the formation of the set of disulfide bonds that favors folding.

des_U species may constitute only a minor fraction of that ensemble, having a thermodynamic weight comparable to those of the (many) disulfide species with nonnative disulfide bonds. In such cases, the major fraction of that ensemble will oxidize to scrambled (three-disulfide) species rather than to the native protein, since the oxidation rates of the unfolded species should be comparable. As another example, the oxidative folding of *wild-type* RNase A should also proceed through a des_U pathway at temperatures or denaturant concentrations above the unfolding transition of its (three-disulfide) des species. Under such conditions, the oxidation of the unstructured species of the 3S ensemble is much more likely to result in scrambled 4S species, rather than the native protein. Such scrambled species must then be reduced and reshuffled before having another chance to form the native disulfide bonds. In such cases, a typical protein molecule may be oxidized to (and reduced from) the scrambled ensemble several times before finally reshuffling to a des_U species that oxidizes to the native protein. The rate-determining step in such des_U pathways appears to be the formation of the native protein N by oxidation from an unfolded *n*S precursor ensemble, followed by conformational folding (Figure 5a), contrary to a previously proposed mechanism.[35]

Oxidative folding can also proceed through *folded* intermediates (pathways II−IV in Figure 4). Such folded disulfide species may become preferentially populated to

levels well above that dictated by the bias toward native disulfide bonds in these unstructured ensembles, because their folded structure inhibits the reduction and reshuffling of their disulfide bonds. (To simplify the discussion, we will assume that such folding occurs only in the des species, although our conclusions do not depend on this assumption.) Each folded intermediate can be considered as giving rise to a separate pathway from the preequilibrated unstructured ensembles to the native protein, since these intermediates are not generally in equilibrium with each other, e.g., des[40−95] and des[65−72] in RNase A (Figure 3).

The most efficient structured pathway is the *des$_N$ pathway*, which involves folded intermediates that protect their disulfide bonds from reduction/reshuffling but leave their thiol groups readily accessible for oxidization to the native protein (pathway II in Figure 4). In RNase A, the pathways involving the formation and folding of des[40−95] and des[65−72] at 25 °C are representative of these des$_N$ pathways. Such species are effective in regenerating the native protein because they "lock in" the native disulfide bonds without inhibiting the subsequent oxidation step(s) leading to the native protein. The rate-determining step in such des$_N$ pathways is the formation of the structured des species by reshuffling from an unstructured *n*S precursor ensemble and conformational folding (Figure 5b).

By contrast, pathways III and IV of Figure 4 involve metastable des species. Such species may accumulate to significant levels, since they protect their disulfide bonds from rearrangement; however, rearrangements of such a species must be preceded by an unfolding step that allows its thiol groups and/or disulfide bonds to react. Pathways III and IV are distinguished by the ultimate fate of the metastable des species. In pathway III, the metastable des species preferentially oxidizes to the native state; such pathways are denoted as *metastable des$_N$* pathways, owing to their similarity to pathway II in Figure 4 (the normal des$_N$ pathway). In pathway IV, the metastable des species preferentially reshuffles rather than oxidizing; such pathways are denoted as *metastable dead-end* pathways.

At low temperatures, the des[65−72] species of RNase A may provide an example of a metastable des$_N$ pathway (pathway III in Figure 4). At least one thiol group (probably Cys72[25]) appears to be buried in des[65−72] under these conditions and may form a metastable, on-pathway species (unpublished data). This is consistent with the formation of N′ cited above, in which Cys72 (or, possibly, Cys65) forms a buried mixed disulfide bond with a DTT molecule. By contrast, the structured des[26−84] and des-[58−110] observed under strongly stabilizing conditions appear to be metastable, dead-end species, in that they do not oxidize directly to the native protein but preferentially reshuffle to the unstructured 3S ensemble.

A key structural difference should be noted between the two types of metastable species. Des[65−72] appears to be *disulfide-secure*, meaning that its thiol groups can be exposed by a local unfolding that maintains the burial (and protection) of its three native disulfide bonds. By contrast, des[26−84] and des[58−110] appear to be *disulfide-insecure*, meaning that their thiol groups are as well protected as their disulfide bonds; structural fluctuations that expose the thiols are likely to expose the disulfide bonds as well. Under typical oxidative folding conditions, disulfide-secure species preferentially oxidize to the native protein (forming metastable des$_N$ pathways, pathway III of Figure 4), whereas disulfide-insecure species preferentially reshuffle to the unstructured ensemble (forming metastable dead-end pathways, pathway IV of Figure 4). The latter preference arises because the reshuffling reaction typically competes effectively with the oxidation reaction when the protein disulfide bonds and mixed disulfide bonds are equally accessible, due to the conformational similarity between a reshuffling reaction (in which a protein thiolate attacks a protein disulfide bond) and the second step of an oxidation reaction (in which a protein thiolate attacks a mixed disulfide bond).

The rate-determining steps of oxidative folding are usually associated with either the formation of a structured disulfide species or the loss of that structure (Figure 5). Thus, the regeneration of the native protein along des$_N$ pathways (pathway II in Figure 4) should be accelerated (decelerated) by factors that promote the folding (unfolding) of the structured intermediate. By contrast, the eventual regeneration of the native protein along a metastable dead-end pathway (pathway IV in Figure 4) is clearly promoted (inhibited) by factors that destabilize (stabilize) the folded metastable species. Hence, the overall kinetics of regeneration may show a complex dependence on experimental conditions influencing the conformational stability of the structured intermediates; changes in such conditions can alter the relative rates along the various pathways and, thus, affect the kinetic partitioning between the various pathways. Ideally, the oxidative folding conditions should be sufficiently stabilizing to favor the folding of disulfide-secure intermediates, but not so strong as to inhibit their subsequent oxidation to the native protein (e.g., des[65−72] of RNase A at low temperatures) and/or the reshuffling of disulfide-insecure species (e.g., des[26−84] and des[58−110]). The redox conditions may also affect the relative populations of the pathways, e.g., by altering the relative flux through the des$_U$ and des$_N$ pathways.

The Oxidative Folding of Other Proteins

These generic pathways appear to apply to other proteins and other redox conditions, as illustrated here by a few examples. The reader is also referred to recent reviews,[36,37] since space limitations preclude a comprehensive survey of the oxidative folding of other proteins.

The three-disulfide protein hirudin seems to follow a des$_U$ pathway (pathway I of Figure 4).[38] No structured intermediates appear to be populated in its regeneration; the rate-determining step appears to be an oxidation from an unstructured 2S ensemble to the native protein. This oxidative folding behavior is similar to that observed for

the RNase A mutants [C65S, C72S] and [C40A, C95A], as well as that of several other small proteins.[37]

The two-disulfide protein ribonuclease T1 appears to regenerate primarily through a des_N pathway (pathway II of Figure 4).[4,39] After initial population of the unstructured R, 1S, and 2S ensembles, the rate-determining step is a slow reshuffling from the 1S ensemble to a structured des species that oxidizes directly to the native protein.

Fully folded des species are not the only possible type of structured disulfide intermediate. Stable conformational structure can also appear in a precursor of a des species; moreover, native disulfide bonds can induce partial folding as well as global folding of the protein. Nevertheless, the effects of "locking in" the native disulfide bonds and the disulfide-secure/insecure distinction remain the same, as illustrated by the well-studied, three-disulfide protein bovine pancreatic trypsin inhibitor (BPTI).[36,37,40,41] Stable native structure can appear in BPTI even after the formation of a single native disulfide bond,[42] while unstructured intermediates are poorly populated (presumably because they reshuffle to the structured species).[36,37] Despite its small size, BPTI also has two hydrophobic cores,[43,44] which may fold semi-independently. Specifically, formation of the 5−55 disulfide bond causes both cores to fold (global folding), while formation of the 30−51 disulfide bond causes only one core to fold (partial folding).[42] In both the (5−55) and (30−51) intermediates, the native 14−38 disulfide bond forms quickly, producing the des[30−51] and des[5−55] species, respectively. However, these des species appear to be disulfide-insecure metastable species, preferentially reshuffling rather than oxidizing; thus, the immediate precursor of native BPTI is almost always the disulfide-secure des[14−38] species. Hence, the rate-determining step in the oxidative folding of BPTI is either a reshuffling to the structured des[14−38] (pathway II of Figure 4) or the escape from the dead-end metastable species des[5−55] and des[30−51] (pathway IV of Figure 4), depending on the conditions.

The four-disulfide protein hen egg white lysozyme (HEWL) illustrates several features of the oxidative folding of larger, more complex proteins.[45,46] HEWL appears to have two folding domains, denoted as the α- and β-domains.[47] The unstructured 1S and 2S disulfide ensembles are populated in the early stages of oxidative folding and exhibit some conformational ordering and bias toward the native disulfide bonds. Three structured des species are observed: des[64−80], des[6−127], and des[76−94], of which des[76−94] appears to be metastable, although it has not been established whether the first two species belong to a type II pathway nor whether des[76−94] belongs to a type III or type IV pathway. Stable structure may even develop in the 2S ensemble, perhaps corresponding to the oxidative folding of the more stable α-domain of HEWL. Aggregation can also be a problem in HEWL, which may be ameliorated by factors that accelerate oxidative folding, such as protein disulfide isomerase.[46] Similar observations have been made for the homologous protein α-lactalbumin,[48,49] although the oxidative folding of that protein is complicated by calcium binding and the formation of the molten globule state.[50−52]

Conclusions

The essential points of this Account are threefold. First, the formation of stable structure in an intermediate of oxidative folding can protect its disulfide bonds and/or thiol groups from further rearrangement, significantly altering the subsequent course of oxidative folding. Experiments on RNase A and other proteins indicate that only native disulfide bonds provoke this folding and that only native-like structure is formed. Consequently, only native disulfide bonds should become protected in this fashion, which should promote the regeneration of the native protein in the absence of metastable species. Second, two types of metastable species were discerned. In disulfide-secure species, the thiol groups may be exposed by a local unfolding event without exposing the disulfide bonds, causing such species to preferentially oxidize rather than reshuffle. By contrast, in disulfide-insecure species, the thiol groups and disulfide bonds tend to become exposed in concert, leading such species to preferentially reshuffle rather than oxidize. (The term *disulfide-secure* may also be applied to species that are not metastable, to indicate the relative protection of the thiol groups and disulfide bonds.) Third, on the basis of these considerations, four generic types of pathways were identified in the oxidative folding of single-domain proteins.

The "locking in" of native disulfide bonds by conformational folding may particularly assist the oxidative folding of large, multidomain proteins. Such proteins can have numerous cysteines and, thus, an exorbitant number of disulfide intermediates. In the absence of structured intermediates or a strong native bias, oxidative folding would be difficult, since disulfide intermediates with nonnative disulfide bonds would far outnumber (and have comparable thermodynamic weight to) the disulfide species with only native disulfide bonds. However, the independent folding of domains may help in overcoming this entropic barrier, by successively locking in the disulfide bonds of each domain. Once a sufficiently stabilizing complement of native disulfide bonds forms in any particular domain, that domain will fold, protecting its native disulfide bonds from further rearrangements. Assuming that this folded domain remains stable during the subsequent regeneration, the regeneration of the remaining domains should continue until all the domains have folded and all disulfide bonds are protected.

This piecemeal approach to the oxidative folding of large, multidomain proteins may also operate in vivo, where oxidative folding can occur cotranslationally. If one domain has finished its oxidative folding while the next is being synthesized, the oxidative folding of even a complex protein may be straightforward. In the future, it may be technically possible to construct an in vitro system for regenerating large, multidomain proteins in this piecemeal fashion.

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (Grant No. GM-24893). Support was also received from the National Foundation for Cancer Research.

References

(1) Thornton, J. M. Disulphide Bridges in Globular Proteins. *J. Mol. Biol.* **1981**, *151*, 261−287.

(2) Kauzmann, W. Relative Probabilities of Isomers in Cystine-Containing Randomly Coiled Polypeptides. In *Symposium on Sulfur in Proteins*; Benesch, R., Benesch, R. E., Boyer, P. D., Klotz, I. M., Middlebrook, W. R., Szent-Györgyi, A. G., Schwarz, D. R., Eds.; Academic Press: New York, 1959; Chapter II.2.

(3) Poland, D. C.; Scheraga, H. A. Statistical Mechanics of Noncovalent Bonds in Polyamino Acids. VIII. Covalent Loops in Proteins. *Biopolymers* **1965**, *3*, 379−399.

(4) Pace, C. N.; Grimsley, G. R.; Thomson, J. A.; Barnett, B. J. Conformational Stability and Activity of Ribonuclease T_1 With Zero, One, and Two Intact Disulfide Bonds. *J. Biol. Chem.* **1988**, *263*, 11820−11825.

(5) Wlodawer, A.; Svensson, L. A.; Sjölin, L.; Gilliland, G. L. Structure of Phosphate-Free Ribonuclease A Refined at 1.26 Å. *Biochemistry* **1988**, *27*, 2705−2717.

(6) Milburn, P. J.; Scheraga, H. A. Local Interactions Favor the Native 8-Residue Disulfide Loop in the Oxidation of a Fragment Corresponding to the Sequence Ser50-Met79 Derived From Bovine Pancreatic Ribonuclease A. *J. Protein Chem.* **1988**, *7*, 377−398.

(7) Altmann, K.-H.; Scheraga, H. A. Local Structure in Ribonuclease A. Effect of Amino Acid Substitutions on the Preferential Formation of the Native Disulfide Loop in Synthetic Peptides Corresponding to Residues Cys58-Cys72 of Bovine Pancreatic Ribonuclease A. *J. Am. Chem. Soc.* **1990**, *112*, 4926−4931.

(8) Talluri, S.; Falcomer, C. M.; Scheraga, H. A. Energetic and Structural Basis for the Preferential Formation of the Native Disulfide Loop Involving Cys-65 and Cys-72 in Synthetic Peptide Fragments Derived From the Sequence of Ribonuclease A. *J. Am. Chem. Soc.* **1993**, *115*, 3041−3047.

(9) Neira, J. L.; Rico, M. Folding Studies on Ribonuclease A, a Model Protein. *Folding Des.* **1997**, *2*, R1−R11.

(10) Raines, R. T. Ribonuclease A. *Chem. Rev.* **1998**, *98*, 1045−1065.

(11) Wedemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A. Disulfide Bonds and Protein Folding. *Biochemistry* **2000**, *39*, 4207−4216.

(12) Hwang, C.; Sinskey, A. J.; Lodish, H. F. Oxidized Redox State of Glutathione in the Endoplasmic Reticulum. *Science* **1992**, *257*, 1496−1502.

(13) Molinari, M.; Helenius, A. Glycoproteins Form Mixed Disulphides With Oxidoreductases During Folding in Living Cells. *Nature* **1999**, *402*, 90−93.

(14) Thatcher, D. R.; Hitchcock, A. Protein Folding in Biotechnology. In *Mechanisms of Protein Folding*; Pain, R. H., Ed.; IRL Press: New York, 1994; Chapter 9.

(15) Huggins, C.; Tapley, D. F.; Jensen, E. V. Sulphydryl−Disulfide Relationships in the Induction of Gels in Proteins by Urea. *Nature* **1951**, *167*, 592−593.

(16) Gilbert, H. F. Molecular and Cellular Aspects of Thiol-Disulfide Exchange. *Adv. Enzymol.* **1990**, *63*, 69−172.

(17) Scheraga, H. A.; Konishi, Y.; Rothwarf, D. M.; Mui, P. W. Toward an Understanding of the Folding of Ribonuclease A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 5740−5744.

(18) Konishi, Y.; Ooi, T.; Scheraga, H. A. Regeneration of Ribonuclease A From the Reduced Protein. Isolation and Identification of Intermediates, and Equilibrium Treatment. *Biochemistry* **1981**, *20*, 3945−3955.

(19) Rothwarf, D. M.; Li, Y.-J.; Scheraga, H. A. Regeneration of Bovine Pancreatic Ribonuclease A: Identification of Two Nativelike Three-Disulfide Intermediates Involved in Separate Pathways. *Biochemistry* **1998**, *37*, 3760−3766.

(20) Rothwarf, D. M.; Li, Y.-J.; Scheraga, H. A. Regeneration of Bovine Pancreatic Ribonuclease A: Detailed Kinetic Analysis of Two Independent Folding Pathways. *Biochemistry* **1998**, *37*, 3767−3776.

(21) Xu, X.; Scheraga, H. A. Kinetic Folding Pathway of a Three-Disulfide Mutant of Bovine Pancreatic Ribonuclease A Missing the [40−95] Disulfide Bond. *Biochemistry* **1998**, *37*, 7561−7571.

(22) Iwaoka, M.; Juminaga, D.; Scheraga, H. A. Regeneration of Three-Disulfide Mutants of Bovine Pancreatic Ribonuclease A Missing the 65−72 Disulfide Bond: Characterization of a Minor Folding Pathway of Ribonuclease A and Kinetic Roles of Cys65 and Cys72. *Biochemistry* **1998**, *37*, 4490−4501.

(23) Xu, X.; Rothwarf, D. M.; Scheraga, H. A. Nonrandom Distribution of the One-Disulfide Intermediates in the Regeneration of Ribonuclease A. *Biochemistry* **1996**, *35*, 6406−6417.

(24) Volles, M. J.; Xu, X.; Scheraga, H. A. Distribution of Disulfide Bonds in the Two-Disulfide Intermediates in the Regeneration of Bovine Pancreatic Ribonuclease A: Further Insights into the Folding Process. *Biochemistry* **1999**, *38*, 7284−7293.

(25) Talluri, S.; Rothwarf, D. M.; Scheraga, H. A. Structural Characterization of a Three-Disulfide Intermediate of Ribonuclease A Involved in Both the Folding and Unfolding Pathways. *Biochemistry* **1994**, *33*, 10437−10449.

(26) Shimotakahara, S.; Rios, C. B.; Laity, J. H.; Zimmerman, D. E.; Scheraga, H. A.; Montelione, G. T. NMR Structural Analysis of an Analog of an Intermediate Formed in the Rate-Determining Step of One Pathway in the Oxidative Folding of Bovine Pancreatic Ribonuclease A: Automated Analysis of ^{1}H, ^{13}C and ^{15}N Resonance Assignments for Wild-Type and [C65S,C72S] Mutant Forms. *Biochemistry* **1997**, *36*, 6915−6929.

(27) Laity, J. H.; Lester, C. C.; Shimotakahara, S.; Zimmerman, D. E.; Montelione, G. T.; Scheraga, H. A. Structural Characterization of an Analog of the Major Rate-Determining Disulfide Folding Intermediate of Bovine Pancreatic Ribonuclease A. *Biochemistry* **1997**, *36*, 12683−12699.

(28) Pearson, M. A.; Karplus, P. A.; Dodge, R. W.; Laity, J. H.; Scheraga, H. A. Crystal Structures of Two Mutants That Have Implications for the Folding of Bovine Pancreatic Ribonuclease A. *Protein Sci.* **1998**, *7*, 1255−1258.

(29) Rothwarf, D. M.; Scheraga, H. A. Regeneration of Bovine Pancreatic Ribonuclease A. 4. Temperature Dependence of the Regeneration Rate. *Biochemistry* **1993**, *32*, 2698−2703.

(30) Low, L. K.; Shin, H.-C.; Narayan, M.; Wedemeyer, W. J.; Scheraga, H. A. Acceleration of Oxidative Folding in Bovine Pancreatic Ribonuclease A by Anion-Induced Stabilization and Formation of Structured Native-Like Intermediates. *FEBS Lett.* **2000**, *472*, 67−72.

(31) Welker, E.; Narayan, M.; Volles, M. J.; Scheraga, H. A. Two New Structured Intermediates in the Oxidative Folding of Ribonuclease A. *FEBS Lett.* **1999**, *460*, 477−479.

(32) Klink, T. A.; Woycechowsky, K. J.; Taylor, K. M.; Raines, R. T. Contribution of Disulfide Bonds to the Conformational Stability and Catalytic Activity of Ribonuclease A. *Eur. J. Biochem.* **2000**, *267*, 566−572.

(33) Li, Y.-J.; Rothwarf, D. M.; Scheraga, H. A. An Unusual Adduct of Dithiothreitol With a Pair of Cysteine Residues of a Protein As a Stable Folding Intermediate. *J. Am. Chem. Soc.* **1998**, *120*, 2668−2669.

(34) Weissman, J. S.; Kim, P. S. A Kinetic Explanation for the Rearrangement Pathway of BPTI Folding. *Nature Struct. Biol.* **1995**, *2*, 1123−1130.

(35) Creighton, T. E. Protein Folding Coupled to Disulphide Bond Formation. *Biol. Chem. Hoppe-Seyler* **1997**, *378*, 731−744.

(36) Gilbert, H. F. The Formation of Native Disulfide Bonds. In *Mechanisms of Protein Folding*; Pain, R. H., Ed.; IRL Press: New York, 1994; Chapter 5.

(37) Dadlez, M. Disulfide Bonds in Protein: Folding Studies: Friends or Foes? *Acta Biochim. Polon.* **1997**, *44*, 433−452.

(38) Thannhauser, T. W.; Rothwarf, D. M.; Scheraga, H. A. Kinetic Studies of the Regeneration of Recombinant Hirudin Variant 1 With Oxidized and Reduced Dithiothreitol. *Biochemistry* **1997**, *36*, 2154−2165.

(39) Pace, C. N.; Creighton, T. E. The Disulphide Folding Pathway of Ribonuclease T1. *J. Mol. Biol.* **1986**, *188*, 477−486.

(40) Creighton, T. E. The Disulfide Folding Pathway of BPTI. *Science* **1992**, *256*, 111−112.

(41) Weissman, J. S.; Kim, P. S. Response to Creighton 1992 Science Letter. *Science* **1992**, *256*, 112−114.

(42) Oas, T. G.; Kim, P. S. A Peptide Model of a Protein Folding Intermediate. *Nature* **1988**, *336*, 42−48.

(43) Richardson, J. S. Schematic Drawings of Protein Structure. *Methods Enzymol.* **1985**, *115*, 359−380.

(44) Yu, M. H.; Weissman, J. S.; Kim, P. S. Contributions of Individual Side-Chains to the Stability of BPTI Examined by Alanine-Scanning Mutagenesis. *J. Mol. Biol.* **1995**, *249*, 388−397.

(45) Roux, P.; Ruoppolo, M.; Chaffotte, A.-F.; Goldberg, M. E. Comparison of the Kinetics of SS-Bond, Secondary Structure, and Active Site Formation During Refolding of Reduced Denatured Hen Egg White Lysozyme. *Protein Sci.* **1999**, *8*, 2751−2760.

(46) van den Berg, B.; Chung, E. W.; Robinson, C. V.; Mateo, P. L.; Dobson, C. M. The Oxidative Refolding of Hen Lysozyme and Its Catalysis by Protein Disulfide Isomerase. *EMBO J.* **1999**, *18*, 4794−4803.

(47) Radford, S. E.; Dobson, C. M.; Evans, P. A. The Folding of Hen Lysozyme Involves Partially Structured Intermediates and Multiple Pathways. *Nature* **1992**, *358*, 302−307.

(48) Ewbank, J. J.; Creighton, T. E. Pathway of Disulfide-Coupled Unfolding and Refolding of Bovine α-Lactalbumin. *Biochemistry* **1993**, *32*, 3677−3693.
(49) Ewbank, J. J.; Creighton, T. E. Structural Characterization of the Disulfide Folding Intermediates of Bovine α-Lactalbumin. *Biochemistry* **1993**, *32*, 3694−3707.
(50) Creighton, T. E.; Ewbank, J. J. Disulfide-Rearranged Molten Globule State of α-Lactalbumin. *Biochemistry* **1994**, *33*, 1534−1538.
(51) Peng, Z.-Y.; Wu, L. C.; Kim, P. S. Local Structural Preferences in the α-Lactalbumin Molten Globule. *Biochemistry* **1995**, *34*, 3248−3252.
(52) Wu, L. C.; Kim, P. S. A Specific Hydrophobic Core in the α-Lactalbumin Molten Globule. *J. Mol. Biol.* **1998**, *280*, 175−182.

AR000063M