Acta Biochimica Polonica

Vol. 44 No. 3/1997

433–452

QUARTERLY

Critical review

Disulfide bonds in protein folding studies: friends or foes?

Michał Dadlez

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5A, 02-106 Warsaw, Poland

Key words: protein folding, disulfide bonds, BPTI

The studies on protein folding pathways utilizing disulfide bonds as reporter groups in several protein model systems are reviewed. Implications for a general mechanism of protein folding are discussed. An updated folding pathway for bovine pancreatic trypsin inhibitor (BPTI) based on recent data is proposed.

Unique properties of proteins and their enormous functional diversity depend on precise folding of the protein chains to proper three dimensional structures. Breaking the rules by which the sequence determines the structure is sometimes called deciphering the second half of the genetic code [1]. The potential diversity of primary sequences leads to an astronomic number of possible conformations exceeding the number of atoms in the Universe. It has been argued [2] that the search of the native state through all possible conformations would require unrealistic timescale, so well defined pathways are necessary since the folding process can proceed on a physiological timescale [3]. Despite a significant effort put into the studies on protein folding during the last half of a century the problem of the mechanism of this process is yet unsolved. Analysis of the folding process acquires additional importance due to the recently growing recognition that protein misassembly, in many cases leading to serious human diseases, may start from structured folding intermediates and sometimes therapeutics could be designed provided a sound knowledge of the nature of these species is obtained [4].

To describe the protein folding pathway it is necessary to find which conformations, intermediate between the unfolded and native ones, are formed in the most significant fraction of the molecules at each stage of folding. Although it is easy to induce the folding-unfolding transition at will, by changing solvent parameters like: temperature, composition, and pressure, and to trace the folded-unfolded equilibrium by a variety of physicochemical methods, it is difficult to isolate and characterize the intermediate states. First, since folding starts from heterogeneous, unfolded state the intermediate structures are also heterogeneous to some extent and are difficult to characterize in terms of positions of the atoms in the structure. The main difficulty, however, is that folding is close to an all-or-none process, or in other words it is cooperative, which means

Abbreviations: Abu, α-amino-*n*-butyric acid; ANS, 1-anilino-8 naphthalenesulfonic acid; BPTI, bovine pancreatic trypsin inhibitor; DsbA, periplasmic thiol/disulfide oxidoreductase of *E. coli*; ER, endoplasmic reticulum.

that, in folding conditions, intermediate structures are transient and completely folded or unfolded molecules dominate the population of folding molecules.

Nevertheless, the efforts to separate and characterize folding intermediates are being undertaken, especially since some of them have been observed to accumulate during folding. A mere accumulation, however, of an intermediate is not a sufficient argument to assess its importance in the folding process. An intermediate is on the main folding pathway only if a significant fraction of folding molecules proceeds through this intermediate. Intermediates that accumulate may be easiest to find, but also they may be kinetic traps, i.e. intermediates in which further folding is blocked for some reason, as illustrated in Fig. 1. These are dead ends of folding, not intermediates, through which a major flux of molecules proceeds. It is thus necessary not only to describe the conformation of a given intermediate but also its kinetic properties, its rate of formation, to estimate the flux of molecules through this intermediate state, and its importance for folding. The major flow of folding molecules in the absence of kinetic barriers should proceed through fast forming and fast decaying intermediate states which do not accumulate. This underscores the importance of kinetic studies which enable to design a flow chart of the folding pathway. Recent experiments show [5–9] that many intermediates isolated in the process of folding of different proteins are actually kinetic traps, and relaxing of kinetic barriers gives fast folding proteins without accumulation of intermediates. Thus a careful kinetic analysis is necessary to assess the role of an intermediate in folding.

The folding of disulfide bonded proteins gives a unique opportunity to trap and isolate folding intermediates. Intermediates containing different quantity and arrangement of the disulfide bonds can be separated, maintained at equilibrium and their structures studied by different methods. The kinetics of the formation of these intermediates is also relatively easy to measure. For these reasons cysteine thiols were often used as reporter groups for the studies of the folding process.

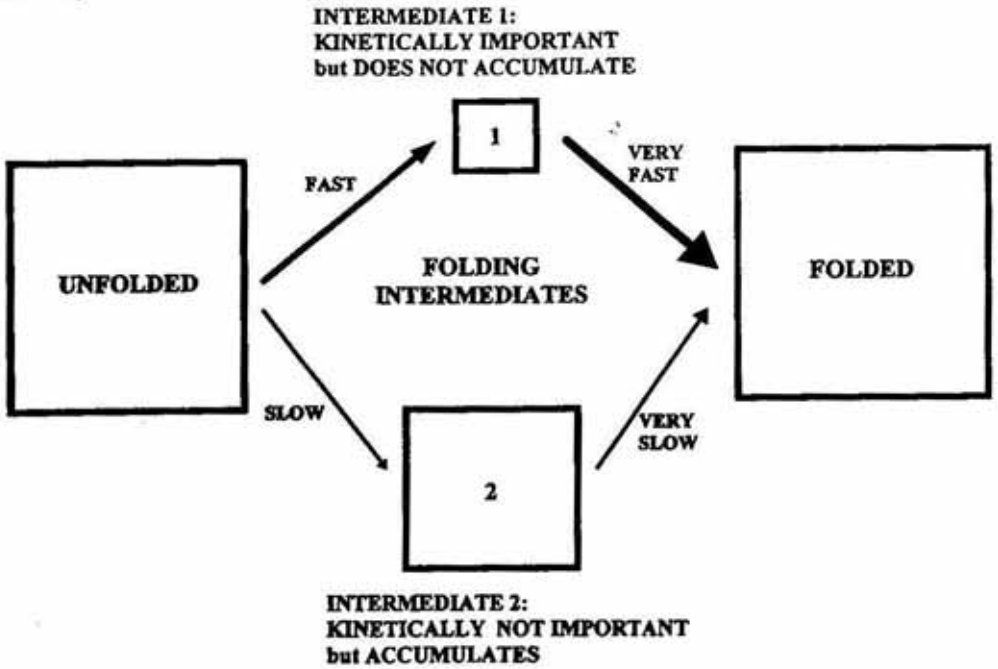

Figure 1. Hypothetical folding of a protein with two folding intermediates: a productive one (1) and a kinetic trap (2).

Since the rearrangement of (2) to the native state is very slow it will accumulate on the folding pathway, so in the spectrum of folding intermediates (2) may be more populated than (1); thus (2) may be easier to find, isolate and characterize. More molecules populate (2) than (1) at each moment of folding, however the main flux of folding molecules proceeds through (1) and (1) is the main folding intermediate, not (2).

Oxidation of the free thiol groups of cysteines to form a disulfide bond can be carried out by molecular oxygen dissolved in water, but this process is hard to control and should be avoided [10, 11] in protein folding studies. Instead, low molecular thiols are used as oxidants. A major potential source of oxidizing power in endoplasmic reticulum (ER) of eukaryotic cells is oxidized glutathione (GSSG) [12] which is present in ER in equilibrium with reduced glutathione (GSH). Overall glutathione concentration in ER is 5–10 mM. GSSG is widely used for *in vitro* folding experiments. Glutathione-mediated formation of an intramolecular disulfide bond proceeds in two steps (Fig. 2). First, a free thiol in the protein attacks the disulfide bond of the oxidized glutathione, releasing a molecule of reduced glutathione. This step yields a species with a disulfide bond between a protein cysteine residue and a glutathione molecule, referred to as a mixed disulfide. In the second step, a different protein thiol attacks the protein-glutathione mixed disulfide, resulting in the formation of a disulfide bond between two protein cysteine residues. Such an intramolecular disulfide bond may subsequently be rearranged by other free thiols within the same molecule. The key feature of the disulfide rearrangement is that it is strongly pH dependent, being very slow at low pH. This permits to quench the rearrangement at a chosen time by rapid lowering of pH [13]. Trapped species may be purified to homogeneity by chromatographic methods and subsequently analysed in terms of their disulfide pairing or structural properties. Moreover, the rearrangement may be resumed starting from the purified species, provided the isolated intermediates are dissolved in a neutral pH buffer. Irreversibly blocking agents, such as iodoacetate, are also sometimes used to trap the intermediates, but may distort the spectrum of intermediates [13].

Numerous proteins were used both to study folding and to obtain active proteins from reduced unfolded proteins aggregated in the form of inclusion bodies often obtained as a product of overexpression of recombinant proteins in bacteria [14]. The most thoroughly studied are bovine pancreatic trypsin inhibitor (BPTI), ribonuclease A, ribonuclease T1, α-lactalbumin. Other studies con-

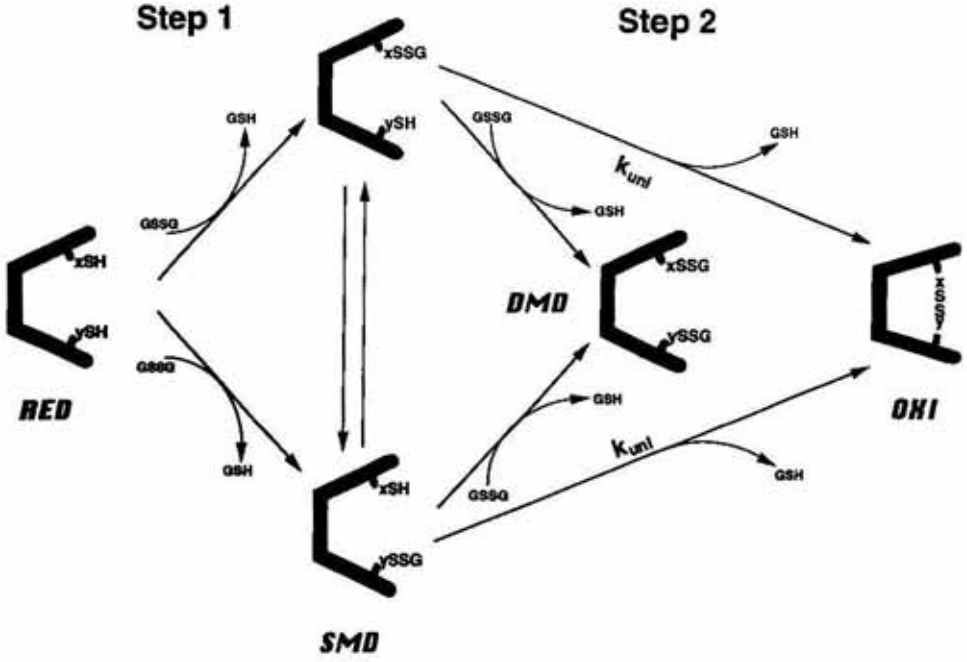

Figure 2. Scheme of the reaction of a protein containing two cysteines with oxidized glutathione (GSSG) leading to formation of an intramolecular disulfide bond.

The reduced protein (RED) dissolved in a neutral pH buffer containing GSSG undergoes oxidation first (Step 1) to the single-mixed disulfide species (SMD) and then (Step 2) to either the intramolecularly oxidized species (OXI) or the double mixed-disulfide species (DMD). This reaction is most often used to study the oxidative refolding of proteins.

cerned hen lysozyme [15], insulin [16], insulin-like growth factor [17], immunoglobulin light chain [18], hirudin [19], human epidermal growth factor [20], carboxypeptidase inhibitor [21], tick anticoagulant peptide [22], ω-conotoxins [23], ovalbumin [24], porcine growth hormone [25], bovine growth hormone [26], apamin [27], alkaline phosphatase [28], and β-lactamase [29]. Also, nonnative disulfide bonds have been engineered into proteins in order to facilitate folding studies [30, 31]. The folding studies on some of these proteins are described below in some detail. Only *in vitro* studies are taken into account, *in vivo* folding studies [32–34] or enzyme catalyzed folding [35–37] remain beyond the scope of this minireview.

BOVINE PANCREATIC TRYPSIN INHIBITOR

Bovine pancreatic trypsin inhibitor is a small, 58 residue protein. It is one of the most thoroughly studied proteins in terms of function, structure and folding. BPTI is a potent inhibitor of numerous serine proteinases including trypsin, plasmin, kallikrein and chymotrypsin [38]. The inhibitor binds proteinase *via* a continuous epitope, a proteinase binding loop of the conformation canonical for the entire family of inhibitors [39]. The structure of BPTI, many of its variants and homologous proteins have been studied both by crystallographic [40–47], and nuclear magnetic resonance methods [48–55]. The structure of native BPTI is known to be extremely stable, with the midpoint temperature of thermal unfolding of 103°C [56, 57]. Such extraordinary stability is necessary for BPTI function of tight binding to the enzyme. The binding loop of the inhibitor is kept very rigid by the protein structure, so its conformation does not change upon binding to enzyme [39].

The crystal structure of BPTI is shown schematically in Fig. 3, along with the three disulfide bonds, linking six cysteines in the following pattern 5–55, 14–38, 30–51, where the numbers denote positions of the cysteines in the sequence. These disulfides are a key feature of the BPTI structure; when the disulfide bonds are reduced, the protein unfolds. Fully native structure of BPTI with its strong inhibitory properties can be reestablished in approximately 50% of molecules in the process of oxidative refolding [13, 59, 60]. The remaining 50% form a stable intermediate (named N*) lacking one of the disulfides (30–51). The process of oxidative refolding of BPTI has been studied for two decades and is one of the best characterized in terms of intermediates populating the folding pathway and their kinetic and structural properties [13, 59–69]. The issue of the importance of different intermediates found on the BPTI folding pathway and the mechanisms involved in the formation of the native state have led to significant controversies [70–73] which most clearly exemplify the difficulties accompanying the interpretation of results of the oxidative refolding studies in relation to the protein folding problem in general.

The possibility of trapping folding intermediates by blocking the disulfide rearrangement at different stages of folding of disulfide

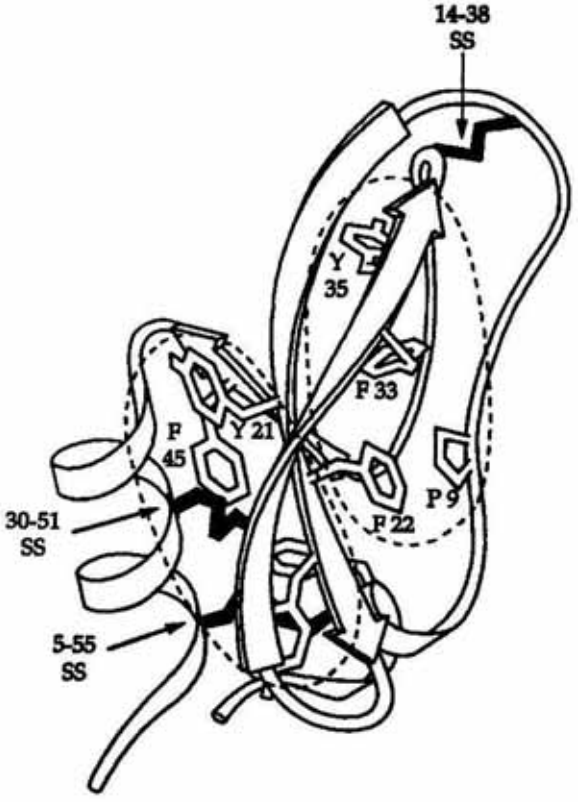

Figure 3. Schematic representation of the crystal structure of bovine pancreatic trypsin inhibitor (BPTI) [58].

Disulfide bonds between cysteines, some aromatic side chains, and two hydrophobic cores are shown.

bonded proteins has been recognized for the first time by T. Creighton in his pioneering work on BPTI folding in 1974 [61]. During further work by the same group [59, 60, 62] carried out in the years 1975–1990, some intermediates which accumulate on the BPTI folding pathway were assigned and the average rates of transition from intermediates of different quantity of disulfides were measured. In this work iodoacetate was used to trap intermediates and ion-exchange electrophoresis and paper electrophoresis were used to separate these intermediates [74].

Based on these experiments a preferred pathway has been worked out for disulfide pattern formation in BPTI, at pH 8.7, 25°C. Although it was recognized that a mere accumulation of an intermediate is insufficient to treat it as an important folding intermediate, no precise measurements of the flux through each of the intermediates could be made due to crude separation methods available at that time. In the most recent version [75] this pathway proceeds as follows (Fig. 4). The starting species of the pathway is the reduced species and the formation of the first disulfide is assumed to be random [59, 75, 76] since the structure of BPTI is thought to be close to random-coil [60, 77–80] and all six cysteines are believed to be equally reactive [81]. Of all 15 possible single-disulfide intermediates only (30–51), i.e. the form containing one disulfide bond linking cysteines 30 and 51 and the remaining four cysteines in free thiol form, accumulates because it is thermodynamically more stable. This intermediate is then assumed to oxidize preferentially to two non-native two-disulfide species (5–14, 30–51) and (5–38, 30–51) and to a natively paired species (14–38, 30–51) (named N′). N′ is a kinetic trap since it is a very stable molecule and it interconverts only slowly to other species, but the sequential addition of the third, missing native disulfide 5–55 is impossible directly. N′ is a completely folded protein of the structure indistinguishable from native, and free thiols 5 and 55 are fixed and buried inside the protein, inaccessible to oxidizing agents and unable to form a transition state for formation of a disulfide bond. The only way out of this trap is by unfolding, and rearrangement to non-native species (like (5–14, 30–51) or (5–38, 30–51)). These non-native species are more flexible and may interconvert fast to the only productive two-disulfide intermediate (30–51,

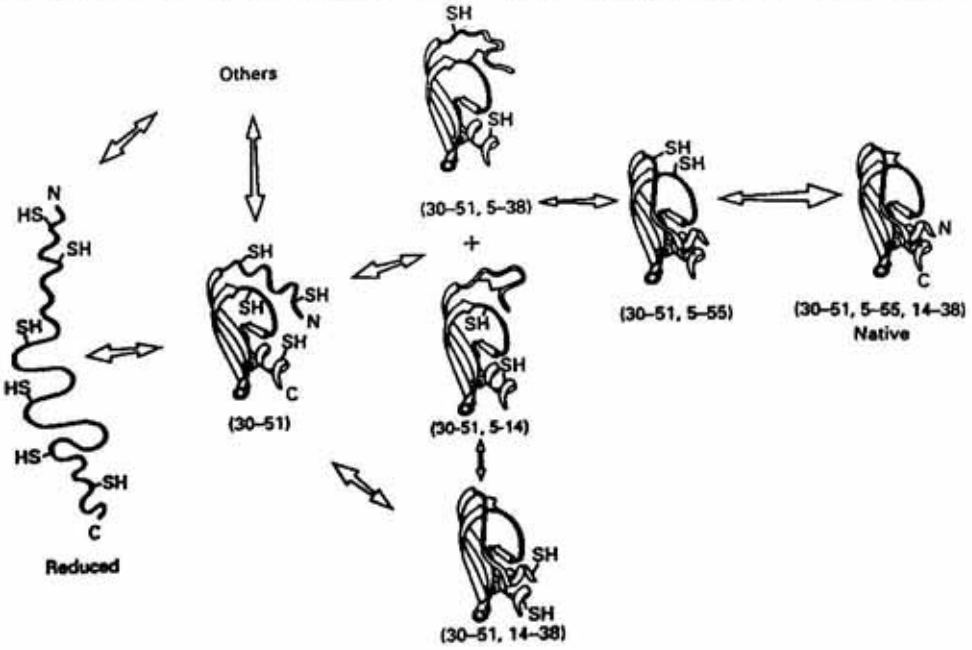

Figure 4. The disulfide folding pathway of bovine pancreatic trypsin inhibitor (BPTI) at pH 8.7 as proposed in [75].

Numbers in brackets denote disulfides formed in a given intermediate (for instance: (30–51) denotes an intermediate with a single disulfide linking cysteines 30 and 51, and four remaining cysteines in a free thiol form). See discussion in the text.

5–55). This is also a very stable, native-like molecule but, unlike cysteines 5 and 55, the two free thiols 14 and 38 are exposed to solvent and easily oxidize to form the native state. Closing of 14–38 as the last disulfide is the only way to reach native BPTI, so the trick in BPTI folding is to get (30–51, 5–55) and avoid all kinetic traps. Whereas the above described pathway is certainly the most efficient in avoiding the kinetic traps and reaching the native state in shortest time, BPTI is not necessarily selected for optimal folding, but rather for extraordinary stability. The productivity of the pathway is thus not an argument that the major flux of folding molecules actually proceeds along the proposed pathway.

The results obtained in the early studies were astonishing in several aspects. First, the amount of intermediates found to accumulate was small. Only five (two single-disulfide and three double disulfide intermediates), a small fraction of possible intermediates, were found. This indicated that folding proceeds *via* a well defined specific pathway and not by random search of native conformation. The second, much more surprising finding, was that three of the five intermediates contained non-native disulfides i.e. cysteines paired in a way different from that found in native protein. This indicated that acquisition of native disulfides in BPTI is not sequential [13, 62]. On the contrary, the folding pathway involves intramolecular rearrangements of intermediates containing non-native disulfides. Moreover, the third striking feature was that the two non-native intermediates, namely (5–14, 30–51) and (5––38, 30–51) were detected at similar levels as a native one (14–38, 30–51). The conclusion has been drawn that the thermodynamic stability of these disulfides must be comparable. On this basis some authors suggested that non-native intermediates could be stabilized by specific interactions not existing in the native protein and that these interactions are necessary for folding [72, 73]. Such a counterintuitive conclusion that non-native

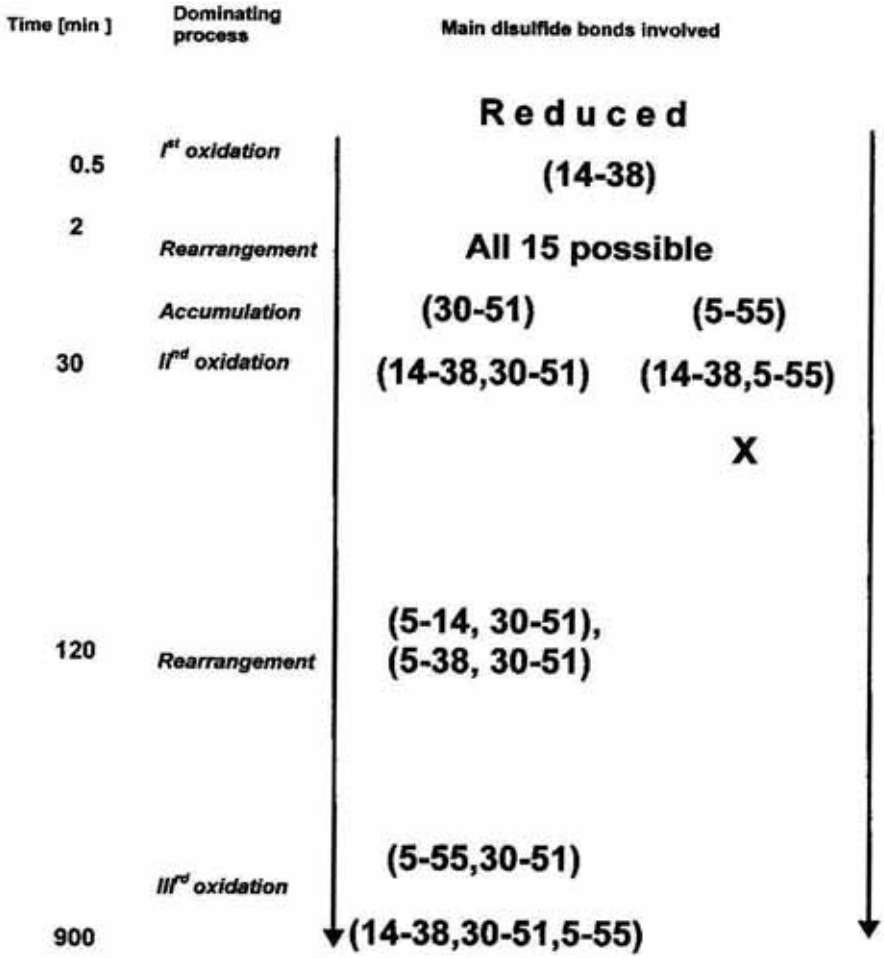

Figure 5. The disulfide folding pathway of BPTI at pH 7.3 (from [13], modified to include recent data).

The presented flow chart is based on available kinetic data and shows intermediates through which the largest fraction of molecules proceeds at each stage of folding. X denotes a dead end of folding, in which a two-disulfide intermediate (14–38, 5–55) is unable to proceed to the native form. See discussion in the text.

structures help in folding of proteins made the folding problem seem much more complicated. Whereas the first two conclusions were later fully confirmed, the third was not [13, 63].

The above results prompted an effort of other groups to understand these striking results. The studies carried out in the period 1990–1997 showed that all features of BPTI folding can be accounted for by the presence of native-like, or native structure in the folding intermediates. At the beginning of 90-ies, the advance in separation techniques allowed a reexamination of the BPTI folding pathway, carried out by Weissman & Kim [13]. They have found that all well populated intermediates contained native disulfides and proposed a different pathway [13]. This pathway, in the form of a protein folding flow chart, slightly modified to include most recent results, is presented in Fig. 5.

The pathway represents BPTI folding at pH shifted to the more physiological value of 7.3. Since pK's of cysteines are close to 8.8, at pH 8.7 cysteine residues, hydrophobic at neutral pH's, become predominantly deprotonated and charged, which destabilizes the structure forming interactions in folding molecules. As oxidant, 0.15 mM oxidized glutathione was used; at higher glutathione concentrations too many mixed disulfides with glutathione may form at later stages of folding, blocking the intramolecular disulfide formation. Reaction times are taken from Fig. 9 in [13] and Fig. 1 in [67]. The proposed pathway presents intermediates through which a significant flux of the molecules proceeds at different stages of folding. Wherever possible precise kinetic data were used to estimate the flux *via* different intermediates. Some of these intermediates either form fast (in a significant quantity of molecules), decay fast, and do not accumulate, the others accumulate to significant levels due to their thermodynamic stability and inability to proceed on the folding pathway. The presented pathway is based on a wealth of recent kinetic and structural data.

The starting point for folding — the reduced protein — is no longer believed to be a random coil, and cysteines in reduced protein to be equally reactive [68]. Different techniques, like NMR of reduced BPTI or its models [82–86], fluorescence energy transfer [87–90], disulfide formation studies [68], or ANS binding study [91] indicate a collapsed compact structure with significant conformational preferences.

Moreover, the structure of the reduced state specifically promotes the native pairing of cysteines 14 and 38 [68] even at the first step of folding. It is achieved by hydrophobic interactions of seven residues located in between positions 18 and 35 in BPTI sequence [69], which form a folding nucleus and accelerate formation of the 14–38 disulfide fourfold. This effect was measured in molecules with four cysteines replaced by alanines. Since cysteines are more hydrophobic this conformational effect in native BPTI may be significantly enhanced [86, 91–93]. The hydrophobic cluster persists even in 8 M urea [69]. Combined with higher reactivities of cysteines 14 and 38 *versus* cysteines 5, 30, 51 or 55 [68], the hydrophobic folding nucleus leads to much faster formation of 14–38, even at the first step of folding, than of any other of the 14 possible single disulfide intermediates. 14–38 forms as a first disulfide in 30% of molecules [68]. The major flux of folding molecules at 30 s after the start of folding proceeds through the 14–38 disulfide.

(14–38) is a preferred kinetic intermediate at the first step of folding but it is not thermodynamically preferred; other single disulfide intermediates are more stable. A synthetic model of (14–38) in which the remaining cysteines are replaced by α-amino-*n*-butyric acid (Abu) folds into an unstable molten globule [86, 91, 93]. A recombinant model of (14–38) in which the remaining four cysteines are replaced by alanines does not show folded structure (Dadlez, M. & Kim, P.S., unpublished). On the contrary, models of (30–51) and (5–55), the remaining two native single disulfide intermediates, fold into native or partly native stable structures [50–52, 94–98]. Since (14–38) is relatively unstable it readily interconverts *via* intramolecular disulfide rearrangements to other single-disulfide intermediates [68]. The rate of this interconversion is approximately 100-fold faster than the rate of oxidation by 0.15 mM oxidized glutathione [68]. Intramolecular di-

sulfide rearrangement is very efficient at this stage of folding and allows to approach the thermodynamic equilibrium (compare Fig. 9 in [13] and Fig. 2 in [68]), where (30–51) and (5–55) dominate in the flux of folding molecules, 2 min from the start of folding.

(5–55) is in fact a completely folded molecule [50, 97], it specifically binds to trypsin [92] in spite of the lack of two disulfides. BPTI can thus fold to native structure provided one and only one covalent link is formed between the N- and C-terminal part of the molecule. Folding is completed at this stage; all subsequent steps are necessary only to support the existing structure with additional disulfide links. These additional links are necessary to obtain an extraordinarily stable protein, which after cleavage by trypsin will not dissociate from the enzyme and will continue to inhibit its function. The analysis of all further steps of folding, involving the non-native intermediates, is thus less relevant to protein folding studies.

The second oxidation step (15 min) starts from a close to thermodynamic equilibrium mixture of single disulfide intermediates in which (30–51) and (5–55) predominate. Again, as previously, the rate of formation of 14–38 disulfide is the fastest, so a major flux of molecules proceeds to the two-disulfide native-like intermediates (14–38, 30–51) (N′) and (14–38, 5–55) (N*) [62, 63, 65]. Non-native intermediates (5–14, 30–51) and (5–38, 30–51) which were suggested to play a dominant role in this step of folding [62, 75, 99] do not seem to be preferred either kinetically or thermodynamically [63], though the precise rates of formation of disulfide bonds 14–38, 5–14 and 5–38 in (30–51) of native BPTI have not been compared directly. Only the thermodynamical stability of the intermediates N′, (5–14, 30–51) and (5–38, 30–51) has been compared directly and it has been shown that N′ is the most stable by a factor of at least 30 [63] and that these three intermediates exist in rapid equilibrium. This rules out the direct pathway (30–51) to ((5–14, 30–51) or (5–38, 30–51)) and then directly to the productive two-disulfide intermediate (5–55, 30–51). Instead, N′ and N* will dominate. The average rates of formation of a second disulfide in (30–51) were measured in three model molecules corresponding to the three two-disulfide intermediates, namely N′, (5–14, 30–51) and (5–38, 30–51). In these three molecules serines replaced two cysteines not engaged in these disulfides and experiments were carried out at pH 8.7. Both the use of serines [97] and high pH destabilizes the native-like structural tendencies in the folding molecules. Still, the rate of formation of 14–38 was found to be 4 times faster than of either of the two non-native disulfides, and the difference can be expected to increase in the native BPTI, at pH 7.3. Thus a major flux of the folding molecules again proceeds through 14–38 disulfide at this stage of folding (15 min).

N′ and N* are very stable molecules [13, 45, 49, 100, 101] and their structures are indistinguishable from the native structure, with the remaining free thiol groups buried inside the folded protein, and thus inaccessible to oxidizing agents [13, 59, 66]. This prevents their direct oxidation to native BPTI which is (14–38, 30–51, 5–55). The only productive two disulfide intermediate is (30–51, 5–55). This is also a very stable molecule [102] but, contrary to N′ and N*, free cysteines (14 and 38) are not buried and easily oxidize to form the native three disulfide bonded structure. Remarkably, the direct formation of (30–51, 5–55) at the preceding step of folding, i.e. from (30–51) or (5–55), is also very slow [62, 103] for similar reasons. Both (30–51) and (5–55) have native structures which fix at least Cys-55 in (30–51) and bury Cys-30 and 51 in (5–55). A mixed disulfide with glutathione of Cys-55 has been observed to accumulate in (30–51) [62], as no other mixed disulfide does, which indicates problems with intramolecular rearrangements in (30–51) on the way to (30–51, 5–55). In (5–55) cysteines 30 and 51 are inaccessible to oxidizing agents [13, 97] and formation of (30–51, 5–55) is inefficient.

Due to hyperreactivity of Cys-14 and -38 and burial of other four cysteines, the sequential formation of disulfides in BPTI becomes blocked. The major flux of molecules proceeds to one after another kinetic trap. N* for instance is stable for weeks and therefore it is a dead end for folding. N′ is less stable so it can undergo slow rearrangements which

finally (900 min) allow its interconversion to native state. When these kinetic barriers are released, for instance by mutation [104] or addition of a chaotropic agent, e.g. urea [13] (Dadlez, M. & Otlewski, J., unpublished), folding is much more efficient in terms of acquiring the native disulfide pairing.

Non-native disulfides, necessary to get out of the kinetic traps, form according to the rules of disulfide rearrangements and not due to their thermodynamic properties. Most probably free and relatively flexible Cys-5 in N′ can slowly rearrange with 14–38 disulfide to form (5–14, 30–51) or (5–38, 30–51) thus, the molecules become less stable [63] and can further rearrange to (30–51, 5–55),which is a thermodynamically most stable two disulfide intermediate. Again, in (30–51, 5–55) 14–38 disulfide forms readily, leading to native BPTI.

Fast formation of solvent exposed 14–38 at each step of folding is a prominent feature of the folding pathway of BPTI. Combined with strong native-like folding tendencies and extraordinary stability of this molecule, this factor may be the reason for the complicated folding pathway of BPTI. The problems of BPTI folding are the problems of unfolding of folded structure (900 min) and not the problems of folding (2 min). Folding itself proceeds easily and is completed at the stage of a single disulfide. Clearly, a folding pathway in which the buried disulfides form first would be more efficient in terms of acquiring the complete set of disulfides, but it is not realized in the case of BPTI.

The oxidative folding pathway of BPTI has been studied in considerable detail. These studies have shown that native-like tendencies are very strong even at first steps of folding, proved that partly folded native-like structures can be encountered on the folding pathway, and that abundant kinetic traps may decrease the rate of folding if the structure in folding intermediates is prematurely stable. The folding process itself, however, seems to be completed early in folding of BPTI, so further studies focusing on the early events may bring new information on protein folding in general. Thorough knowledge of the BPTI folding pathway helped to study the folding of BPTI with a pro sequence [64], catalysed by protein disulfide isomerase [65], and its folding *in vivo* [33, 34, 105].

RNase T1

This is a 104 amino-acids ribonuclease from *Aspergillus oryzae* containing four cysteines at positions 2, 6, 10, 103. In native structure [106] these cysteines form two disulfide bonds 2–10 and 6–103. Unlike BPTI, native RNase T1 with intact disulfides is unfolded by moderate concentrations of chaotropic agents or temperature [107]. When folding conditions are restored after prolonged incubation in unfolding conditons only a small fraction of molecules refold fast to the native state, the recovery of the remaining majority is very slow [108, 109]. Nevertheless, most of the secondary structure is regained in milliseconds in the entire population of folding molecules. Species which interconvert slowly to native form, contain intact disulfides and native-like structure but improper isomers of the two proline peptide bonds [110, 111], which in the native molecule are *cis*. All slow folding steps are catalysed by peptidyl-prolyl-*cis*/*trans*-isomerase [112, 113] and thus can be attributed to formation of energetically unfavorable proline *cis* isomer in the two prolines 39 and 55. Moreover, it has been shown that premature structure, formed in folding, additionally slows down prolyl isomerization, leading to formation of kinetic traps in folding of RNase T1 [114]. The structure affects either proline in a different way and the flux of folding molecules is directed by the order in which the two amide bonds isomerize [108, 109]. Removal of one kinetic trap by replacement of Pro-55 by alanine accelerates folding [115]. However, removal by mutation of the second (Pro-39 by Ala replacement) does not accelerate folding [116] and it has been shown that the *cis* isomer of the peptide bond Tyr38-Ala39 is retained in the folded form [117]. Proline dependent folding of disulfide intact RNase T1 has been studied thoroughly, which helped to understand the oxidative folding of RNase T1 from reduced molecule.

Disruption of both disulfides at low ionic strength leads to unfolding of the protein and

isomerization of the majority of prolines to *trans* conformer. The refolding pathway has been studied by a classical approach [118] of isolation of disulfide bonded intermediates. Initially three disulfides 2–10, 6–10, and 2–6 form, and Cys-103 is not involved in any disulfide bond. Rearrangements follow, leading to intermediate (6–103), which already has a native-like conformation and which is readily oxidized to the native protein (6–103, 2–10). Replacement of cysteines 2 and 10 by alanines [119] shows that this disulfide bond is important for stability but not for folding.

Remarkably, in a high ionic strength buffer even the molecules with both disulfides reduced fold to the native structure [107], [120, 121]. This unique feature shows that no disulfides at all are necessary either for maintaining structure and directing folding [119], and allows to study folding kinetics both with intact and broken disulfides. A careful comparison of folding kinetics [8, 122] shows that the disulfide bonds formed actually decelerate the folding of RNase T1 by hindering proline isomerization. Isomerization of prolines and disulfide formation have to be coupled with the formation of structure, so proper proline isomers and disulfides can be selected upon folding. Native like tendencies are strong and the structure formation process starts early in folding, being decelerated at later steps both by disulfides and proline residues.

Thiols can serve as reporter groups for early folding events and salt induced folding of RNase T1 [123, 124]. It has been found that the average distance of the cysteines 6 and 103 is a hundred times shorter in reduced unfolded protein than in a high ionic strength folded molecule. At the same time the salt induced native structure does not fix cysteines and the disulfide forms easily. Thus, these are the conditions for formation of a productive intermediate, efficiently directing the native pairing of disulfides. In RNase T1 the pre-formed structure does not lead to a kinetic trap blocking disulfide formation, as it was observed for BPTI or immunoglobulin light chain [18].

RNase A

This protein consists of 124 amino acids, contains four disulfide bonds and two prolines in *cis* isomer in the native structure [125]. It was the first protein that was shown to be able to reconstitute spontaneously its active, fully native conformation after reduction and unfolding [3, 126]. Its regeneration pathway is much more complicated than in the case of BPTI, since 8 cysteines can form in folding 764 different disulfide bonded species as compared to 74 for 6 cysteines in BPTI. An additional complication is brought by two *cis*-prolines [127–129]. Nevertheless, a considerable and long lasting effort in many laboratories, dating from 1974 [130] to 1993 [131], to elucidate the RNase A folding pathway has been undertaken. It brought only a general description of this process in terms of equilibrium constants between groups of intermediates [130–136]. During folding, the disulfide intermediates achieve a steady state distribution from which the native state emerges [131, 137]. The intermediates have been found to be disordered [138], and deprived of enzymatic activity [131] which indicates that more natively paired disulfides are necessary to stabilize the structure of RNase A than of BPTI. Six major rate determining steps have been found [134] indicating a multiple folding pathway. Other studies suggest a single rate determining step and a single folding pathway [133, 139]. The two theories lead to a significant controversy [10, 11, 131–133, 139, 140], centered on the interpretation of the obtained data and the kinetic role of the tentative major rate determining intermediate. This intermediate, lacking the disulfide 65–72, has been found as the major species on the unfolding pathway [141]. Its structure is native-like [142], but no special kinetic significance of this intermediate has been proven yet. It is not known either at which stage of folding the native-like structural tendencies lead to the appearance of the native structure. Non-random residual structure has been found in reduced, unfolded RNase A [143–146], but its role in folding has not been established.

α-LACTALBUMIN

α-Lactalbumin is a two-domain protein, consisting of an α-helical domain and β-sheet domain [147, 148]. The helical domain consists of N- and C-terminal parts of the molecule connected by the β-sheet domain. A single calcium ion is bound to the β-sheet domain and significantly stabilizes the native conformation. Four disulfides support the structure of α-lactalbumin [149], two in the α-helical domain linking cysteines 6–120 and 28–111, and two in the β-sheet domain. Without the calcium ion α-lactalbumin exists in the state of a molten globule [150–152]. Disulfide bonds are perfect reporter groups to help studying the structure of a molten globule, which is difficult to characterize using classical methods of structural analysis due to its flexibility and heterogeneity. The disulfide folding pathway of α-lactalbumin has been studied by a classical approach [153], but, as in the case of RNase A, only the most general description could be obtained. Based on unfolding studies, two late folding intermediates have been tentatively proposed as important folding intermediates. In the first, one of the α-helical domain disulfides 6–120 is broken and in the second both of them (6–120 and 28–111) are missing [154, 155]. The affinity to calcium is retained in both of these intermediates, and the structure in holo form is native-like (first intermediate) or locally native like in the region of the molecule surrounding the calcium binding site (second intermediate) [154]. In apo form the first intermediate has all the characteristics of a molten globule, and the second of a partly molten globule and partly of unfolded protein. When these intermediates in the apo form are allowed to rearrange their disulfides, multiple species differing in disulfide pairing appear [154, 155]. This seems to indicate that in the molten globule there are no apparent preferences for disulfide pairing, suggesting that α-lactalbumin in the molten globule state does not have a specific tertiary fold.

However, precise measurements of the equilibria between different disulfide pairings point to a different conclusion. A model of α-helical domain constructed by linking N- and C-terminal part of α-lactalbumin by a short Gly linker has been shown to form a molten globule state in which a native like disulfide pairing is thermodynamically preferred [156]. In denaturing conditions this preference disappears and non-native pairings are preferred. Also, local structural preferences have been found in the entire α-lactalbumin in the molten globule state. Molten globule of α-lactalbumin has a non-uniform structure [157]. Whereas the α-helical domain adopts a native-like topology, the β-sheet domain remains unstructured. Four cysteines of the α-helical domain can form six disulfides, and six single-disulfide α-lactalbumin variants have been studied [158] in terms of disulfide preferences both in denaturing and in non-denaturing conditions. Remarkably, the molten globule state seems to enhance 1000-fold the preference for formation of 28–111 disulfide but bears no effect on 6–120 disulfide. In conclusion these observations show that structural specificity in the molten globule of α-lactalbumin is very local, restricted to a region rich in hydrophobic residues, which is close to 28–111 disulfide. Disulfide bonds proved useful in the studies of the molten globule structure, which is thought to be an important folding intermediate for many proteins [159, 160].

OTHER PROTEINS

Studies of some other proteins also provided results which are interesting for protein folding studies. In immunoglobulin light chains a single disulfide bond is buried inside the protein structure. It has been shown [18] that this protein folds without the disulfides and that after folding the formation of a disulfide is slow. Remarkably, it is also slow in the urea unfolded state due to a large distance in sequence between the cysteines. In conclusion, in the course of folding both structure formation and disulfide formation have to be concerted for efficient formation of a native protein. Similarly, if β-lactamase is allowed to fold in reducing environment further disulfide formation is very slow [28]. DsbA protein, catalyzing disulfide formation in the periplasm of bacteria, has to compete with

inherent tendencies for structure formation to introduce a disulfide bond [29]. Very slow unfolding of a nerve growth factor was attributed to untangling of a cysteine knot [161] in which a peptide chain has to be threaded through a loop.

An interesting case of two different, thermodynamically equivalent, folded structures has been discovered on studying the folding of insulin-like growth factor [17, 162]. It was shown that a single sequence codes for two structures differing in disulfide pairing and tertiary packing, but retaining similar secondary structures. A compact, partially folded intermediate containing a single disulfide seems to be the branch point in a bifurcating folding pathway of this protein. In ω-conotoxins oxidative folding of a mature protein is modestly efficient, and forms with different disulfide pairing are also of significant thermodynamic stability [23]. However, addition of C-terminal glycine, which is a part of ω-conotoxins pro-sequence, enhances folding efficiency by stabilizing the native *versus* the non-native form [163].

In a series of studies on the folding of tick anticoagulant peptide [22], epidermal growth factor [20], carboxypeptidase inhibitor [21] or hirudin [19] by Chang and coworkers no native-like intermediates have been found to accumulate along the folding pathway. Since no kinetic properties of these intermediates were studied it is not known whether the lack of accumulation was caused by lack of thermodynamic stability in natively paired intermediates or by their efficient decay to later folding intermediates. Nevertheless, these studies indicate that in many proteins disulfide formation is well concerted with structure formation and no kinetic traps hinder the process of folding. This is especially interesting in the case of the tick anticoagulant peptide which is structurally homologous to BPTI but does not share the complicated folding scheme. Presence of scrambled species, i.e., fully oxidized proteins with non-native disulfide pairings, underscores the necessity for the presence of an external reducing agent and relative inefficiency of intramolecular rearrangements in the late steps of oxidative folding.

Also intermolecular folding events were studied using disulfides, as in the case of dimeric bovine seminal ribonuclease, where it has been shown that adventitious oxidation can lead to the formation of a misfolded but active dimer [164]. Disulfides helped to distinguish between homo- and heterodimerization of two oncogene products, transcription modulators *fos* and *jun* [165].

In general, formation of disulfide bonds proved useful in the studies of different aspects of the protein folding and assembly process. Since disulfide formation may lead to kinetic traps, careful kinetic studies are necessary to assess the importance of a given folding intermediate. Most clearcut results were obtained in the systems with a small number of thiol groups engaged; with more than two thiols the system contains too many possible forms to enable a precise analysis, which often leads to difficulties in interpretation and to controversies. Recent interest in folding intermediates suggests that in further work disulfides will prove to be an important tool for the studies on heterogeneous and flexible, partly folded protein structures.

REFERENCES

1. Gierash, L.M. & King, J. (1990) Deciphering the second half of the genetic code; in *Protein Folding*. AAAS. Washington.
2. Levinthal, C. (1968) Are there pathways for protein folding? *J. Chem. Phys.* **65,** 44–45.
3. Anfinsen, C.B., Haber, E., Sela, M. & White, F.H. (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* **47,** 1309–1314.
4. Wetzel, R. (1996) For protein misassembly, it is the "I" decade. *Cell* **86,** 699–702.
5. Baldwin, R.L. (1995) The nature of protein folding pathways: The classical *versus* the new view. *J. Biomol. NMR* **5,** 103–109.
6. Baldwin, R.L. (1996) Why is protein folding so fast? *Proc. Natl. Acad. Sci. U.S.A.* **93,** 2627–2628.

7. Weissman, J.S. (1995) All roads lead to Rome? The multiple pathways of protein folding. *Chem. Biol.* **2,** 255–260.

8. Mucke, M. & Schmid, X.F. (1994) Intact disulfide bonds decelerate the folding of ribonuclease T1. *J. Mol. Biol.* **239,** 713–725.

9. Sosnick, T.R., Mayne, L., Miller, R. & Englander, S.W. (1994) The barriers in protein folding. *Nature Struct. Biol.* **1,** 149–155.

10. Creighton, T.E. (1984) Disulfide bond formation in proteins. *Methods Enzymol.* **107,** 305–318.

11. Creighton, T.E. (1986) Disulfide bonds as probes of protein folding pathways. *Methods Enzymol.* **131,** 83–106.

12. Hwang, C., Sinskey, A.J. & Lodish, H.F. (1992) Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* **257,** 1496–1502.

13. Weissman, J.S. & Kim, P.S. (1991) Reexamination of the folding of BPTI: Predominance of native intermediates. *Science* **253,** 1386–1393.

14. Marston, F.A.O. (1986) The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem. J.* **240,** 1–12.

15. Eyles, S.J., Radford, S.E., Robinson, C.V. & Dobson, C.M. (1994) Kinetic consequences of the removal of a disulfide bridge on the folding of hen lysozyme. *Biochemistry* **33,** 13038–13048.

16. Tang, J.-G. & Tsou, C.-L. (1990) The insulin A and B chains contain structural information for the formation of the native molecule. *Biochem. J.* **268,** 423–455.

17. Miller, J.A., Owers Narhi, L., Hua, Q.-X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K.S., Tsai, L. & Weiss, M.A. (1993) Oxidative refolding of insulin-like growth factor 1 yields two products of similar thermodynamic stability: A bifurcating protein-folding pathway. *Biochemistry* **32,** 5203–5213.

18. Goto, Y. & Hamaguchi, K. (1981) Formation of the intrachain disulfide bond in the constant fragment of the immunoglobulin light chain. *J. Mol. Biol.* **146,** 321–340.

19. Chatrenet, B. & Chang, J.-Y. (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. *J. Biol. Chem.* **268,** 20988–20996.

20. Chang, J.-Y., Schindler, P., Ramsaier, U. & Lai, P.-H. (1995) The disulfide folding pathway of human epidermal growth factor. *J. Biol. Chem.* **270,** 9207–9216.

21. Chang, J.-Y., Canals, F., Schindler, P., Querol, E. & Aviles, F.X. (1994) The disulfide folding pathway of potato carboxypeptidase inhibitor. *J. Biol. Chem.* **269,** 22087–22094.

22. Chang, J.-Y. (1996) The disulfide folding pathway of tick anticoagulant peptide (TAP) a Kunitz-type inhibitor structurally homologous to BPTI. *Biochemistry* **35,** 11702–11709.

23. Price-Carter, M., Gray, W.R. & Goldenberg, D.P. (1996) Folding of ω-conotoxins. 1. Efficient disulfide coupled folding of mature sequences *in vitro*. *Biochemistry* **35,** 15537–15546.

24. Takahashi, N. & Hirose, M. (1992) Reversible denaturation of disulfide-reduced ovalbumin and its reoxidation generating the native cysteine cross-link. *J. Biol. Chem.* **267,** 11565–11572.

25. Cardamone, M., Puri, N.K. & Brandon, M.R. (1995) Comparing the refolding and reoxidation of recombinant porcine growth hormone from urea denatured state and from *Escherichia coli* inclusion bodies. *Biochemistry* **34,** 5773–5794.

26. Langley, K.E., Berg, T.F., Strickland, T.W., Fenton, D.M., Boone, T.C. & Wypych, J. (1987) Recombinant-DNA-derived bovine growth hormone from *Escherichia coli*. *Eur. J. Biochem.* **163,** 313–321.

27. Chau, M.-H. & Nelson, J.W. (1992) Cooperative disulfide bond formation in apamin. *Biochemistry* **31,** 4445–4450.

28. Walker, K.W. & Gilbert, H. (1994) Effect of redox environment on the *in vitro* and *in vivo* folding of RTEM1 β-lactamase and *Escherichia coli* alkaline phosphatase. *J. Biol. Chem.* **269,** 28487–28493.

29. Frech, C., Wunderlich, M., Glockshuber, R. & Schmid, F.X. (1996) Competition between

DsbA-mediated oxidation and conformational folding of RTEM1 β-lactamase. *Biochemistry* **35,** 11386–11395.

30. Clarke, J. & Fersht, A.R. (1993) Engineered disulfide bonds as probes of the folding pathway of barnase: Increasing the stability of proteins against the rate of denaturation. *Biochemistry* **32,** 4322–4329.

31. Strausberg, S., Alexander, P., Wang, L., Gallagher, T., Gilliland, G. & Bryan, P. (1993) An engineered disulfide cross-link accelerates the refolding rate of calcium-free subtilisin by 850-fold. *Biochemistry* **32,** 10371–10377.

32. Freedman, R.B. (1995) The formation of protein disulphide bonds. *Curr. Opin. Struct. Biol.* **5,** 85–91.

33. Creighton, T.E. & Charles, I.G. (1987) Biosynthesis, processing, and evolution of bovine pancreatic trypsin inhibitor. *Cold Spring Harbor Symp. Quant. Biol.* **LII,** 511–519.

34. Creighton, T.E., Bagley, C.J., Cooper, L., Darby, N.J., Freedman, R.F., Kemmink, J. & Sheikh, A. (1993) On the biosynthesis of bovine pancreatic trypsin inhibitor (BPTI). *J. Mol. Biol.* **232,** 1176–1196.

35. Frech, C. & Schmid, F.X. (1995) DsbA-mediated disulfide bond formation and catalyzed prolyl isomerisation in oxidative protein folding. *J. Biol. Chem.* **270**, 5367–5374.

36. Bardwell, J.C.A. & Beckwith, J. (1993) The bonds that tie: Catalyzed disulfide bond formation. *Cell* **74,** 769–771.

37. Zapun, A. & Creighton, T.E. (1994) Effects of DsbA on the disulfide folding of BPTI and α-lactalbumin. *Biochemistry* **33**, 5202– 5211.

38. Fritz, H. & Wunderer, G. (1983) Biochemistry and applications of aprotinin, the kallikrein inhibitor from bovine organs. *Arzneim. Forsch./Drug Res.* **33,** 479–494.

39. Bode, W. & Huber, R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204,** 433–451.

40. Eigenbrot, C., Randal, M. & Kossiakoff, A.A. (1992) Structural effects induced by mutagenesis affected by crystal packing factors: The structure of a 30–51 disulfide mutant of basic pancreatic trypsin inhibitor. *Proteins Struct. Funct. Genet.* **14,** 75–87.

41. Deisenhofer, J. & Steigemann, W. (1975) Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1.5 Å resolution. *Acta Crystallogr.* **B31,** 238–250.

42. Wlodawer, A., Walter, J., Huber, R. & Sjolin, L. (1984) Structure of bovine pancreatic trypsin inhibitor. Results of joint neutron and X-ray refinement of crystal form II. *J. Mol. Biol.* **180,** 301–329.

43. Wlodawer, A., Nachman, J., Gilliland, G.L., Gallagher, W. & Woodward, C. (1987) Structure of form III crystals of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **198,** 469–480.

44. Hynes, T.R., Randal, M., Kennedy, L.A., Eigenbrot, C. & Kossiakoff, A.A. (1990) X-ray crystal structure of the inhibitor domain of Alzheimer's amyloid β-protein precursor. *Biochemistry* **29,** 10018–10022.

45. Eigenbrot, C., Randal, M. & Kossiakoff, A.A. (1990) Structural effects induced by removal of a disulfide-bridge: The X-ray structure of the C30A/C51A mutant of basic pancreatic trypsin inhibitor at 1.6 Å. *Protein Eng.* **3,** 591–598.

46. Housset, D., Kim, K.-S., Fuchs, J., Woodward, C. & Wlodawer, A. (1991) Crystal structure of a Y35G mutant of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **220,** 757–770.

47. Arnoux, B., Merigeau, K., Saludjian, P., Norris, F., Norris, K., Bjorn, S., Olsen, O., Petersen, L. & Ducruix, A. (1995) The 1.6 Å structure of Kunitz-type domain from the α3 chain of human type VI collagen. *J. Mol. Biol.* **246,** 609–617.

48. Wagner, G. & Wuthrich, K. (1982) Amide proton exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution. *J. Mol. Biol.* **160,** 343–361.

49. van Mierlo, C.P.M., Darby, N.J., Neuhaus, D. & Creighton, T.E. (1991) (14–38, 30–51) double disulphide intermediate in folding of bovine pancreatic trypsin inhibitor: A two dimensional ^{1}H nuclear magnetic resonance study. *J. Mol. Biol.* **222,** 353–371.

50. van Mierlo, C.P.M., Darby, N.J., Neuhaus, D. & Creighton, T.E. (1991) Two-dimensional ^{1}H NMR study of the (5–55) single disulphide intermediate of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **222**, 373–390.

51. van Mierlo, C.P.M., Darby, N.J. & Creighton, T.E. (1992) The partially folded conformation of the Cys30–Cys51 intermediate in the disulphide folding pathway of bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 6775–6779.

52. van Mierlo, C.P.M., Darby, N.J., Keeler, J., Neuhaus, D. & Creighton, T.E. (1993) Partially folded conformation of the (30–51) intermediate in the disulphide folding pathway of bovine pancreatic trypsin inhibitor. ^{1}H and ^{15}N resonance assignments and determination of backbone dynamics from ^{15}N relaxation measurements. *J. Mol. Biol.* **229,** 1125––1146.

53. van Mierlo, C.P.M., Kemmink, J., Neuhaus, D., Darby, N.J. & Creighton, T.E. (1994) ^{1}H NMR analysis of the partly-folded non-native two-disulphide intermediates (30–51, 5–14) and (30–51, 5–38) in the folding pathway of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **235**, 1044–1061.

54. Berndt, K.D., Guntert, P., Orbons, L.P. & Wuthrich, K. (1992) Determination of a high quality nuclear magnetic resonance solution structure of the bovine pancreatic trypsin inhibitor and comparison with three crystal structures. *J. Mol. Biol.* **227,** 757–775.

55. Zweckstetter, M., Czisch, M., Mayer, U., Chu, M.-L., Zinth, W., Timpl, R. & Holak, T.A. (1995) Structure and multiple conformations of the Kunitz-type domain from human type VI collagen α3(VI) chain in solution. *Structure* **4,** 195–209.

56. Moses, E. & Hinz, H.-J. (1983) Basic pancreatic trypsin inhibitor has unusual thermodynamic stability parameters. *J. Mol. Biol.* **170,** 765–776.

57. Makhatadze, G.I., Kim, K.-S., Woodward, C. & Privalov, P.L. (1993) Thermodynamics of BPTI folding. *Protein Sci.* **2,** 2028–2036.

58. Richardson, J.S. (1985) Schematic drawings of protein structures. *Methods. Enzymol.* **115,** 359–380.

59. Creighton, T.E. & Goldenberg, D.P. (1984) Kinetic role of the meta-stable native like two disulfide species in the folding transition of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **179,** 497–526.

60. Creighton, T.E. (1978) Experimental studies of protein folding and unfolding. *Prog. Biophys. Mol. Biol.* **33,** 231–297.

61. Creighton, T.E. (1974) Renaturation of the reduced bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **87**, 563–577.

62. Creighton, T.E. (1977) Conformational restrictions on the pathway of folding and unfolding of BPTI. *J. Mol. Biol.* **113,** 275–293.

63. Weissman, J.S. & Kim, P.S. (1992) Kinetic role of nonnative species in the folding of bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 9900–9904.

64. Weissman, J.S. & Kim, P.S. (1992) The pro region of BPTI facilitates folding. *Cell* **71,** 841–851.

65. Weissman, J.S. & Kim, P.S. (1993) Efficient catalysis of disulfide rearrangements by protein disulfide isomerase. *Nature* **365,** 185–188.

66. Weissman, J.S. & Kim, P.S. (1995) A kinetic explanation for the rearrangement pathway of BPTI folding. *Nature Struct. Biol.* **2,** 1123–1130.

67. Dadlez, M. & Kim, P.S. (1995) A third native one-disulphide intermediate in the folding of bovine pancreatic trypsin inhibitor. *Nature Struct. Biol.* **2,** 674–679.

68. Dadlez, M. & Kim, P.S. (1996) Rapid formation of the native 14–38 disulfide bond in the early stages of BPTI folding. *Biochemistry* **35,** 16153–16164.

69. Dadlez, M. (1997) Hydrophobic interactions accelerate early stages of the folding of BPTI. *Biochemistry* **36,** 2788–2797.

70. Creighton, T.E. (1992) The disulfide folding pathway of BPTI. *Science* **256,** 111–112.

71. Weissman, J.S. & Kim, P.S. (1992) The disulfide folding pathway of BPTI. *Science* **256,** 112–114.

72. Goldenberg, D.P. (1992) Native and non-native intermediates in the BPTI folding pathway. *Trends Biochem. Sci.* **17,** 257–261.

73. Hoffman, M. (1991) Straightening out the protein folding puzzle. *Science* **253,** 1357–1358.

74. Creighton, T.E. (1979) Electrophoretic analysis of the unfolding of proteins by urea. *J. Mol. Biol.* **129,** 235–264.

75. Darby, N.J., Morin, P.E., Talbo, G. & Creighton, T.E. (1995) Refolding of bovine pancreatic trypsin inhibitor *via* non-native disulphide intermediates. *J. Mol. Biol.* **249,** 463–477.

76. Darby, N.J. & Creighton, T.E. (1993) Dissecting the disulphide-coupled folding pathway of bovine pancreatic trypsin inhibitor. Forming the first disulfide bonds in analogues of the reduced protein. *J. Mol. Biol.* **232,** 873–896.

77. Creighton, T.E. (1988) On the relevance of non-random polypeptide conformations for protein folding. *Biophys. Chem.* **31,** 155–162.

78. Goldenberg, D.P. & Zhang, J. (1993) Small effects of amino acid replecements on the reduced and unfolded state of pancreatic trypsin inhibitor. *Proteins: Struct. Funct. Genet.* **6,** 259–266.

79. Kosen, P.A., Creighton, T.E. & Blout, E.R. (1980) Ultraviolet difference spectroscopy of intermediates trapped in unfolding and refolding of bovine pancreatic trypsin inhibitor. *Biochemistry* **19,** 4936–4944.

80. Kosen, P.A., Creighton, T.E. & Blout, E.R. (1981) Circular dichroism spectroscopy of bovine pancreatic trypsin inhibitor and five altered conformational states. Relationship of conformation and the refolding pathway of the trypsin inhibitor. *Biochemistry* **20,** 5744–5754.

81. Creighton, T.E. (1975) Reactivities of the cysteine residues of the reduced pancreatic trypsin inhibitor. *J. Mol. Biol.* **96,** 777–782.

82. Roder, H. (1981) Ph.D. Thesis No. 6932, ETH Zurich.

83. Kemmink, J. & Creighton, T.E. (1993) Local conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor and their roles in folding. *J. Mol. Biol.* **234,** 861–878.

84. Kemmink, J., van Mierlo, C.P.M., Scheek, R.M. & Creighton, T.E. (1993) Local structure due to an aromatic-amide interaction observed by ^{1}H-nuclear magnetic resonance spectroscopy in peptide related to the N-terminus of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **230,** 312–322.

85. Lumb, K.J. & Kim, P.S. (1994) Formation of a hydrophobic cluster in denatured bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **236,** 412–420.

86. Pan, H., Barbar, E., Barany, G. & Woodward, C. (1995) Extensive non-random structure in reduced and unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* **34,** 13974–13981.

87. Amir, D. & Haas, E. (1988) Reduced bovine pancreatic trypsin inhibitor has a compact structure. *Biochemistry* **27,** 8889–8893.

88. Amir, D., Krausz, S. & Haas, E. (1992) Detection of local structures in reduced unfolded bovine pancreatic trypsin inhibitor. *Proteins: Struct. Funct. Genet.* **13,** 162–173.

89. Gottfried, D.S. & Haas, E. (1992) Nonlocal interactions stabilize compact folding intermediates in reduced unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* **31,** 12353–12362.

90. Ittah, V. & Haas, E. (1995) Nonlocal interactions stabilize long range loops in the initial folding intermediates of reduced bovine pancreatic trypsin inhibitor. *Biochemistry* **34,** 4493–4506.

91. Ferrer, M., Barany, G. & Woodward, C. (1995) Partially folded, molten globule and molten coil states of bovine pancreatic trypsin inhibitor. *Nature Struct. Biol.* **2,** 211–217.

92. Krokoszynska, I., Dadlez, M. & Otlewski, J. (1997) Early folding intermediates of BPTI folding bind to trypsin. *J. Mol. Biol.* (in press).

93. Barbar, E., Barany, G. & Woodward, C. (1995) Dynamic structure of a highly ordered

β-sheet molten globule: Multiple conformations with a stable core. *Biochemistry* **34**, 11423–11434.

94. States, D.J., Creighton, T.E., Dobson, C.M. & Karplus, M. (1987) Conformations of intermediates in the folding of the pancreatic trypsin inhibitor. *J. Mol. Biol.* **195,** 731–739.

95. Oas, T.G. & Kim, P.S. (1988) A peptide model of a protein folding intermediate. *Nature* **336**, 42–48.

96. Staley, J.P. & Kim, P.S. (1990) Role of a subdomain in the folding of bovine pancreatic trypsin inhibitor. *Nature* **344**, 685–688.

97. Staley, J.P. & Kim, P.S. (1992) Complete folding of bovine pancreatic trypsin inhibitor with only a single disulfide bond. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1519–1523.

98. Staley, J.P. & Kim, P.S. (1994) Formation of a native-like subdomain in a partially-folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Sci.* **10**, 1822–1832.

99. Darby, N.J., van Mierlo, C.P.M., Scott, G.H.E., Neuhaus, D. & Creighton, T.E. (1992) Kinetic roles and conformational properties of the non-native two-disulfide intermediates in the refolding of BPTI. *J. Mol. Biol.* **22,** 905–911.

100. States, D.J., Dobson, C.M., Karplus, M. & Creighton, T.E. (1984) A new two-disulfide intermediate in the refolding of reduced BPTI. *J. Mol. Biol.* **174**, 411–418.

101. Schulman, B.A. & Kim, P.S. (1994) Hydrogen exchange in BPTI variants that do not share a common disulfide bond. *Protein Sci.* **3,** 2226–2232.

102. Stassinopulou, C.I., Wagner, G. & Wuthrich, K. (1984) Two dimensional ^{1}H NMR of two chemically modified analogs of the basic pancreatic trypsin inhibitor. *Eur. J. Biochem.* **145**, 423–430.

103. Goldenberg, D.P. (1988) Kinetic analysis of the folding and unfolding of a mutant form of bovine pancreatic trypsin inhibitor lacking the cysteine 14 and cysteine 38 thiols. *Biochemistry* **27**, 2481–2489.

104. Zhang, J.-X. & Goldenberg, D.P. (1993) Amino acid replacement that eliminates kinetic traps in the folding pathway of BPTI. *Biochemistry* **32**, 14075–14081.

105. Ostermeier, M. & Georgiou, G. (1994) The folding of BPTI in the *Escherichia coli* periplasm. *J. Biol. Chem.* **269,** 21072–221077.

106. Heinemann, U. & Saenger, W. (1982) Specific protein-nucleic acid recognition in ribonuclease T1-2′-guanylic acid complex. *Nature* **299**, 27–31.

107. Oobatake, M., Takahashi, S. & Ooi, T. (1979) Conformational stability of ribonuclease T1. I. Thermal denaturation and effects of salts. *J. Biochem.* **86,** 55–63.

108. Kiefhaber, T., Quaas, R., Hahn, U. & Schmid, F.X. (1990) Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry* **29,** 3053–3060.

109. Kiefhaber, T., Quaas, R., Hahn, U. & Schmid, F.X. (1990) Folding of ribonuclease T1. 2. Kinetic models for the folding and unfolding reactions. *Biochemistry* **29,** 3061–3070.

110. Kiefhaber, T., Schmid, F.X., Willaert, K., Engelborghs, Y. & Chaffotte, A. (1992) Structure of rapidly formed intermediate in RNase T1 folding. *Protein Sci.* **1,** 1162–1172.

111. Mullins, L.S., Pace, C.N. & Rauschel, F.M. (1993) Investigation of ribonuclease T1 folding intermediates by H/D amide exchange — 2D NMR spectroscopy. *Biochemistry* **32**, 6152–6156.

112. Fischer, G., Wittman-Liebold, B., Lang, K., Kiefhaber, T. & Schmid, F.X. (1989) Cyclophilin and peptidyl-prolyl-*cis*/*trans*-isomerase are probably identical proteins. *Nature* **337,** 476–478.

113. Schoenbrunner, E.R., Mayer, S., Tropshung, M., Fischer, G., Takahashi, N. & Schmid, F.X. (1991) Catalysis of protein folding by cyclophilins from different species is highly conserved. *J. Biol. Chem.* **266**, 3630–3637.

114. Kiefhaber, T., Grunert, H.P., Hahn, U. & Schmid, F.X. (1992) Folding of RNase T1 is decelerated by a specific tertiary contact in a folding intermediate. *Proteins: Struct. Funct. Genet.* **12,** 171–179.

115. Kiefhaber, T., Grunert, H.-P., Hahn, U. & Schmid, F.X. (1990) Replacement of a cis proline simplifies the mechanism of ribonuclease T1 folding. *Biochemistry* **29**, 6475–6480.

116. Mayr, L. & Schmid, F.X. (1993) Kinetic models for unfolding and refolding of ribonuclease T1 with substitution of *cis*-proline 39 with alanine. *J. Mol. Biol.* **231**, 913–926.

117. Mayr, L., Willibold, D., Roesch, P. & Schmid, F.X. (1995) Generation of a non-prolyl *cis* peptide bond in ribonuclease T1. *J. Mol. Biol.* **240**, 288–293.

118. Pace, N. & Creighton, T.E. (1986) The disulphide folding pathway of ribonuclease T1. *J. Mol. Biol.* **188**, 477–486.

119. Mayr, L. & Schmid, F.X. (1994) Role of the Cys2–Cys10 disulfide bond for the structure, stability, and folding kinetics of ribonuclease T1. *Protein Sci.* **3**, 227–239.

120. Pace, N., Grimsley, G.R., Thomson, J.A. & Barnett, B.J. (1988) Conformational stability and activity of RNase T1 with zero, one and two intact disulfide bonds. *J. Biol. Chem.* **263**, 11820–11825.

121. Mucke, M. & Schmid, F.X. (1992) Enzymatic catalysis of prolyl isomerization in an unfolding protein. *Biochemistry* **31**, 7848–7854.

122. Mucke, M. & Schmid, F.X. (1994) Folding of ribonuclease T1 in the absence of the disulfide bonds. *Biochemistry* **33**, 14608–14619.

123. Ruoppolo, M. & Freedman, R.B. (1995) Refolding by disulfide isomerization: the mixed disulfide between ribonuclease T1 and glutathione as a model refolding substrate. *Biochemistry* **34**, 9380–9388.

124. Frech, C. & Schmid, F.X. (1995) Influence of protein conformation on disulfide bond formation in the oxidative folding of ribonuclease T1. *J. Mol. Biol.* **251**, 135–149.

125. Wlodawer, A., Svensson, L.A., Sjolin, L. & Gilliland, G.L. (1988) Structure of phosphate-free ribonuclease A refined at 1.26 Å. *Biochemistry* **27**, 2705–2713.

126. Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223–230.

127. Cook, K.H., Schmid, F.X. & Baldwin, R.L. (1979) Role of proline isomerization in folding of ribonuclease A. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6157–6161.

128. Schultz, D.A., Schmid, F.X. & Baldwin, R.L. (1982) *Cis* proline mutants of ribonuclease A.II.Elimination of the slow folding forms by mutation. *Protein Sci.* **1**, 917–924.

129. Dodge, R.W. & Scheraga, H.A. (1996) Folding and unfolding kinetics of the proline to alanine mutants of bovine pancreatic ribonuclease A. *Biochemistry* **35**, 1548–1559.

130. Hantgan, R.R., Hammes, G.G. & Scheraga, H.A. (1974) Pathways of folding of reduced bovine pancreatic ribonuclease. *Biochemistry* **13**, 3421–3431.

131. Rothwarf, D.M. & Scheraga, H.A. (1993) Regeneration of bovine pancreatic ribonuclease. *Biochemistry* **32**, 2671–2703.

132. Creighton, T.E. (1979) Intermediates in the refolding of reduced ribonuclease A. *J. Mol. Biol.* **129**, 411–431.

133. Wearne, S.J. & Creighton, T.E. (1988) Further experimental studies of the disulfide folding transition of ribonuclease A. *Proteins: Struct. Funct. Genet.* **4**, 251–261.

134. Konishi, Y., Ooi, T. & Scheraga, H.A. (1982) Regeneration of ribonuclease A from reduced protein. Rate-limiting steps. *Biochemistry* **21**, 4734–4740.

135. Konishi, Y., Ooi, T. & Scheraga, H.A. (1982) Regeneration of ribonuclease A from reduced protein. Energetic analysis. *Biochemistry* **21**, 4741–4748.

136. Konishi, Y., Ooi, T. & Scheraga, H.A. (1982) Regeneration of RNase A from the reduced protein: Models of regeneration pathways. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5734–5738.

137. Konishi, Y., Ooi, T. & Scheraga, H.A. (1981) Regeneration of ribonuclease A from reduced protein. Isolation and identification of intermediates and equilibrium treatment. *Biochemistry* **20**, 3945–3955.

138. Konishi, Y. & Scheraga, H.A. (1980) Regeneration of ribonuclease A from reduced protein. 2. Conformational analysis of the intermediates by NMR spectroscopy. *Biochemistry* **19**, 1316–1322.

139. Creighton, T.E. (1988) Toward a better understanding of protein folding pathways. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5082–5086.

140. Scheraga, H.A., Konishi, Y., Rothwarf, D.M. & Hui, P.W. (1987) Toward a better understanding of the folding of ribonuclease A. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5740–5744.

141. Rothwarf, D.M. & Scheraga, H.A. (1991) Regeneration and reduction of native bovine pancreatic ribonuclease A with oxidized and

reduced DTT. *J. Am. Chem. Soc.* **113,** 6293–6295.

142. Telluri, S., Rothwarf, D.M. & Scheraga, H.A. (1994) Structural characterization of a three disulfide intermediate of ribonuclease A involved both in the folding and unfolding pathway. *Biochemistry* **33,** 10437–10449.

143. Buckler, D.R., Haas, E. & Scheraga, H.A. (1995) Analysis of the structure of ribonuclease A in native and partially denatured states by time-resolved nonradiative dynamic excitation energy transfer between site-specific extrinsic probes. *Biochemistry* **34,** 15965–15978.

144. Beals, J.M., Haas, E., Krausz, S. & Scheraga, H.A. (1991) Conformational studies of a peptide corresponding to a region of the C-terminus of ribonuclease A. *Biochemistry* **30,** 7680–7692.

145. Sosnick, T.R. & Trewhella, J. (1992) Denatured states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry* **31,** 8329–8335.

146. Wright, P.E., Dyson, H.J. & Lerner, A.R. (1988) Conformation of peptide fragments of proteins in aqueous solution: Implications for initiation of protein folding. *Biochemistry* **27,** 7167–7175.

147. Acharya, K.R., Stuart, D.I., Walker, N.P.C., Lewis, M. & Philips, D.C. (1989) Refined structure of baboon α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol.* **208,** 99–127.

148. Acharya, K.R., Ren, J., Stuart, D.I., Philips, D.C. & Fenna, R.E. (1991) Crystal structure of human α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol.* **221,** 571–581.

149. Vanaman, T.C., Brew, K. & Hill, R.L. (1970) The disulfide bonds of bovine α-lactalbumin. *J. Biol. Chem.* **245,** 4583–4590.

150. Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S. (1976) Three-state denaturation of α-lactalbumin by guanidine hydrochloride. *J. Mol. Biol.* **106,** 359–373.

151. Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S.Y. & Ptitsyn, O.B. (1981) α-Lactoglobulin: compact state with fluctuating tertiary structure? *FEBS Lett.* **136,** 311–315.

152. Xie, D., Bhakuni, V. & Freire, E. (1991) Calorimetric determination of the energetics of the molten globule intermediate in protein folding: apo α-lactoglobulin. *Biochemistry* **30,** 10673–10678.

153. Ewbank, J.J. & Creighton, T.E. (1993) Pathway of disulphide-coupled unfolding and refolding of bovine α-lactalbumin. *Biochemistry* **32,** 3677–3693.

154. Ewbank, J.J. & Creighton, T.E. (1993) Structural characterization of the disulfide folding intermediate of bovine α-lactalbumin. *Biochemistry* **32,** 3694–3707.

155. Ewbank, J.J. & Creighton, T.E. (1991) The molten globule protein conformation probed by disulphide bonds. *Nature* **350,** 518–520.

156. Peng, Z.Y. & Kim, P.S. (1994) A protein dissection study of a molten globule. *Biochemistry* **33,** 2136–2141.

157. Wu, C.L., Peng, Z.Y. & Kim, P.S. (1995) Bipartite structure of the α-lactalbumin molten globule. *Nature Struct. Biol.* **2,** 281–286.

158. Peng, Z.Y., Wu, C.L. & Kim, P.S. (1995) Local structural preferences in the α-lactalbumin molten globule. *Biochemistry* **34,** 3248–3252.

159. Kuwajima, K. (1989) The molten globule as a clue for understanding the folding and cooperativity of globular protein structure. *Proteins: Struct. Funct. Genet.* **6,** 87–103.

160. Ptitsyn, O.B. (1992) in *Protein Folding* (Creighton, T.E., ed.) pp. 243–300, W.H. Freeman & Co., New York.

161. de Young, L.R., Burton, L.E., Liu, J., Powell, M.F., Schmelzer, C.H. & Skelton, N.J. (1996) RhNGF slow unfolding is not due to proline isomerization: Possibility of a cysteine knot loop-threading mechanism. *Protein Sci.* **5,** 1554–1566.

162. Owers Narhi, L., Hua, Q.-X., Arakawa, T., Fox, G.M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J.A. & Weiss, M.A. (1993) Role of native disulfide bonds in the structure and activity of insulin-like growth factor 1: Genetic models of protein folding intermediates. *Biochemistry* **32,** 5214–5221.

163. Price-Carter, M., Gray, W.R. & Goldenberg, D.P. (1996) Folding of ω-conotoxins. 2. Influence of precursor sequences and protein disulfide isomerase. *Biochemistry* **35,** 15547–15557.

164. Kim, J.S. & Raines, R.T. (1994) A misfolded but active dimer of bovine seminal ribonuclease. *Eur. J. Biochem.* **224,** 109–114.

165. O'Shea, E.K., Rutkowski, R., Stafford, W.F. & Kim, P.S. (1989) Preferential heterodimer formation by isolated leucine zippers from fos and jun. *Science* **245,** 646–648.