FORUM REVIEW ARTICLE

Protease Inhibitors as Models for the Study of Oxidative Folding

Joan L. Arolas* and Salvador Ventura

Abstract

The correct balance between proteases and their natural protein inhibitors is of great importance in living systems. Protease inhibitors usually comprise small folds that are crosslinked by a high number of disulfide bonds, making them perfect models for the study of oxidative folding. To date, the oxidative folding of numerous protease inhibitors has been analyzed, revealing a great diversity of folding pathways that differ mainly in the heterogeneity and native disulfide-bond content of their intermediates. The two extremes of this diversity are represented by bovine pancreatic trypsin inhibitor and hirudin, which fold, respectively, *via* few native intermediates and heterogeneous scrambled isomers. Other proteins, such as leech carboxypeptidase inhibitor, share characteristics of both models displaying mixed folding pathways. The study of the oxidative folding of two-domain inhibitors, such as secretory leukocyte protease inhibitor, tick carboxypeptidase inhibitor, and *Ascaris* carboxypeptidase inhibitor, has provided some clues about how two-domain protease inhibitors may fold, that is, either by folding each domain autonomously or with one domain assisting in the folding of the other. Finally, the recent determination of the structures of the major intermediates of protease inhibitors has shed light on the molecular mechanisms guiding the oxidative folding of small disulfide-rich proteins. *Antioxid. Redox Signal*. 14, 97–112.

Introduction

PROTEOLYTIC ENZYMES, commonly referred to as proteases, are essential in nature for the survival of all living organisms. Despite their important physiological functions, enzymes that cleave other proteins are potentially very damaging to their environment, so their activities must be kept strictly under control. Distinct mechanisms exist for the regulation of the catalytic activity of proteases, very important among them being the interactions of proteases with their protein inhibitors (17). Protein inhibitors of proteolytic enzymes comprise the largest and structurally most diverse group of naturally occurring enzyme inhibitors. A comprehensive list of the different inhibitor families is available on the web at http://merops.sanger.ac.uk/ (62). Protease inhibitors are usually classified either by their mechanisms of action, which are surprisingly diverse, or by the type of protease they inhibit, that is, serine, cysteine, aspartic, or metalloproteases—protein inhibitors of threonine and glutamyl proteases are yet to be discovered [for review, see (58, 63)]. Protease inhibitors are often small in size, with one or multiple tandem domains, and are crosslinked by several disulfide bonds. These characteristics provide protease inhibitors with high stability and a long half-life and make them very attractive models for the study of protein folding. Besides enzymes such as ribonuclease A (RNase A) and lysozyme, growth factors, and toxins [for review, see (2)], numerous serine and metalloprotease inhibitors have been investigated using the technique of oxidative folding.

Oxidative folding is the *in vitro* process by which a fully reduced and unfolded protein gains both its native disulfide bonds (disulfide-bond formation) and native structure (conformational folding). The folding processes of many protease inhibitors can be efficiently tracked owing to the small and disulfide-rich nature of their protein folds. Thus, their folding intermediates are chemically trapped in a time-course manner, either irreversibly by alkylation or reversibly by acidification, and further analyzed/purified by reversed-phase chromatography [for review, see (55)]. The isolation of these partially folded species allows their characterization in terms of disulfide-bond connectivity (number and type—native or nonnative) and structure. The derived information will help to define the studied protein's oxidative folding pathway on the basis of the following features: (i) the extent of the

Departament de Bioquímica i Biologia Molecular, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain.
*Current affiliation: Proteolysis Lab, Department of Structural Biology, Molecular Biology Institute of Barcelona, CSIC, Barcelona, Spain.
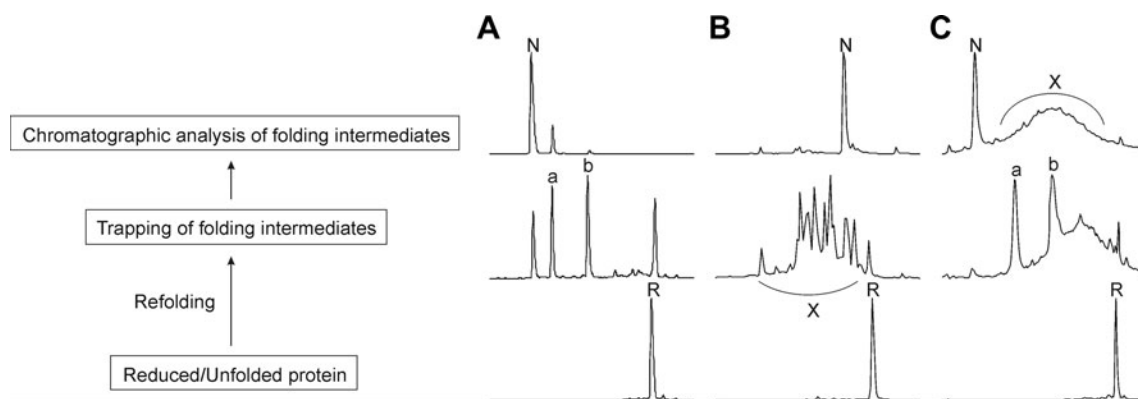
FIG. 1.   Oxidative folding pathways of divergent models. Folding intermediates of leech-derived tryptase inhibitor (LDTI) **(A)**, potato carboxypeptidase inhibitor (PCI) **(B)**, and leech carboxypeptidase inhibitor **(C)** trapped by acidification at selected time points and further analyzed by reversed-phase high-performance liquid chromatography. R and N indicate the elution positions of the fully reduced/unfolded and native forms, respectively. "a" and "b" stand for intermediates containing exclusively native disulfide bonds. "X" represents a highly heterogeneous population of scrambled isomers, that is, species containing nonnative disulfide bonds.

heterogeneity of folding intermediates, (ii) the predominance of intermediates containing native disulfide bonds and native-like structures, and (iii) the presence of scrambled isomers (fully oxidized species containing at least two nonnative disulfides) as intermediates along the folding reaction (32, 33). The two opposite extremes of the oxidative folding pathways are exemplified by proteins that fold through the formation of a few intermediates adopting native disulfide bonds, and proteins that fold *via* highly heterogeneous intermediates containing nonnative disulfides, that is, their folding pathways are populated by scrambled isomers (Fig. 1). On the other hand, other proteins share similarity and dissimilarity with both models, displaying mixed folding pathways.

The present work reviews the pathways of oxidative folding reported thus far for single- and two-domain protease inhibitors. The determinants that account for the large diversity of folding landscapes in these proteins are discussed, together with the recent structural determination of disulfide intermediates that provide fundamental insights into a better understanding of oxidative folding.

Oxidative Folding *via* Few Native Intermediates

Bovine pancreatic trypsin inhibitor

Bovine pancreatic trypsin inhibitor (BPTI) is a 58-residue serine protease inhibitor linked by three disulfide bonds: Cys5-Cys55, Cys14-Cys38, and Cys30-Cys51 (Fig. 2). Crystal and solution structures of BPTI and a number of engineered variants have shown that it is a compact molecule with a double-stranded antiparallel β-sheet, two short segments of α-helix, and a small hydrophobic core (16, 50, 51). BPTI was, together with RNase A, one of the first models for which the characterization of folding pathways was addressed. It has been thoroughly studied for more than three decades, and the conformational features and kinetic roles of all its relevant folding intermediates are already known, making this molecule an outstanding model to understand the structural, sequential, and kinetic determinants that drive oxidative folding (23, 24, 29, 39, 78, 79).

Disulfide bonds are essential for BPTI folding, function, and stability. Although the reduced inhibitor is mostly un-

folded (40), several studies converge to indicate that this state is not a random coil but exhibits some native-like conformational preferences and local interactions between certain residues close in sequence (44, 46). These contacts appear to be strong enough to compensate for the loss of conformational entropy in the unfolded state, although they display reduced stability and are transient, which means that they may be easily broken during subsequent conformational folding (28).

Pioneering studies on the oxidative folding of BPTI were conducted by Creighton and coworkers. They exploited the high reactivity of free cysteines at slightly basic pH to trap by alkylation the different populations of folding intermediates that accumulate along the folding pathway (23, 25, 26). This work revealed that reduced and unfolded BPTI forms its initial disulfide bonds almost randomly, rendering a heterogeneous population of 1-disulfide (1S) molecules. The 1S population subsequently rearranges by intramolecular disulfide exchange (reshuffling) into predominant 1S intermediates, reaching equilibrium before the formation of the second disulfide bond. Although BPTI can form 15 different 1S species, the ensemble is dominated by a species containing a native disulfide between Cys30 and Cys51, designated as [30–51]. The predominance of [30–51] is due to the fact that, in contrast to most of the other 1S species, it adopts a partly folded structure that endows it with the lowest free energy in the ensemble. The conformation of [30–51] is thought to resemble that of native BPTI, with the disulfide bond linking the two major secondary structure elements. The interaction between the β-sheet and the α-helix stabilizes the protein and at the same time protects the disulfide bond. The rest of the polypeptide chain remains very flexible (75). In kinetic terms, the shielding of the Cys30–Cys51 disulfide bond results in a slower breakdown ratio compared with the rest of the bonds in other 1S species, and accordingly, all subsequently formed intermediates in the folding pathway maintain this formed bond. Under the conditions used by Creighton and coworkers, a strong conformational bias is observed, since the other natively bonded species represented only 3% of the ensemble, as in the case of [5–55], or simply were not detected, as in the case of [14–38]. The [30–51] intermediate readily oxidizes two of its free cysteines, producing three major 2S species, [30–51, 5–38],
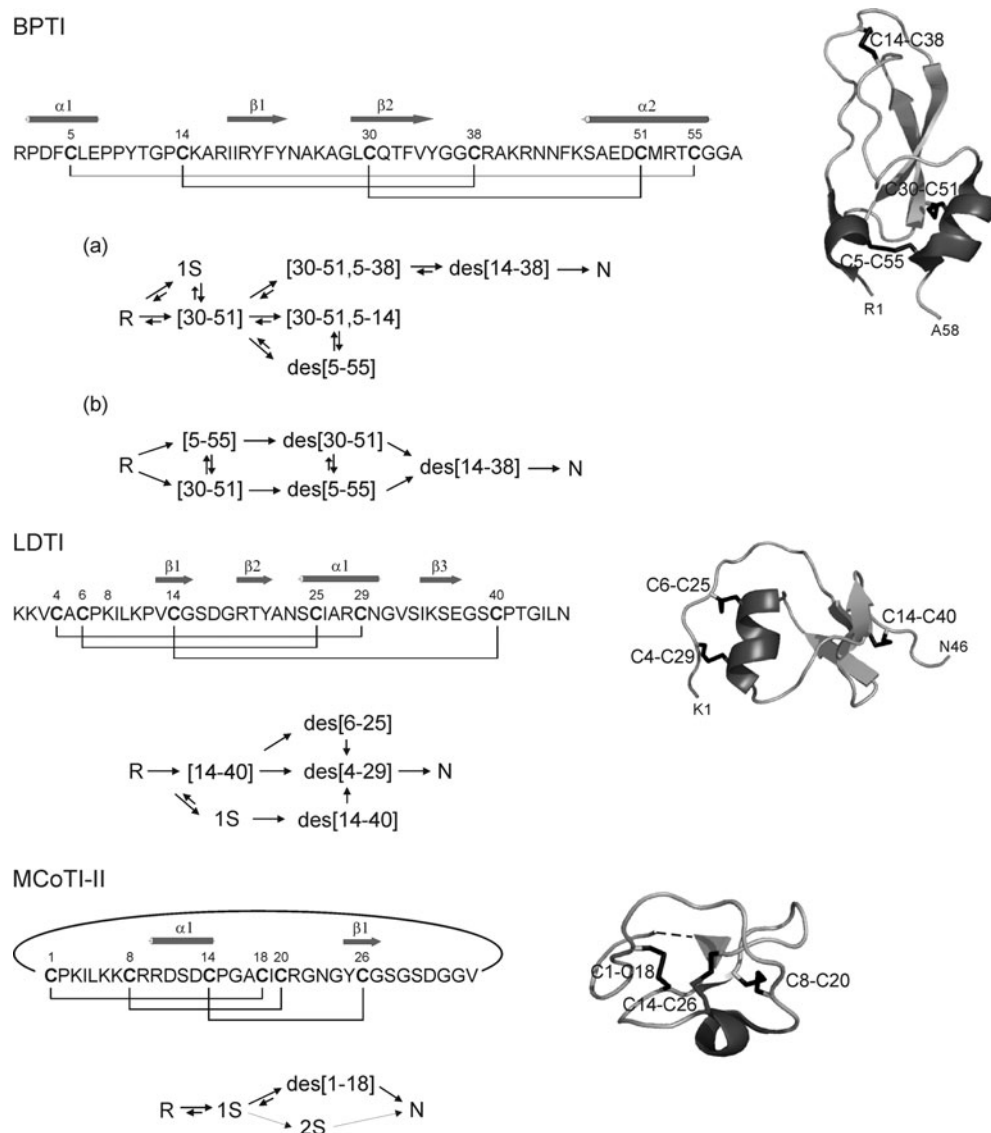
FIG. 2.   Oxidative folding of bovine pancreatic trypsin inhibitor (BPTI)-like proteins. The amino acid sequence, secondary structure elements, and disulfide pairings of BPTI (3 SS and 58 residues), LDTI (3 SS and 46 residues), and *Momordica cochinchinensis* trypsin inhibitor II (MCoTI-II; 3 SS and 34 residues) are schematically shown. The Protein Data Bank ID codes for the structures of BPTI, LDTI, and MCoTI-II are, respectively, 1PIT, 1LDT, and 1IB9 (ribbon plots prepared with program PyMOL). The disulfide bonds are shown as black sticks. An overview of the oxidative folding pathway of BPTI [models proposed by Creighton and coworkers (a) and Weissman and Kim (b)], LDTI, and MCoTI-II is shown below each sequence. R and N indicate the fully reduced/unfolded and native forms, respectively. The disulfide pairings of major folding intermediates are shown between parentheses, and the missing disulfides between parentheses with the prefix "des" (des species). 1S and 2S are ensembles of molecules with the corresponding number of disulfide bonds.

[30–51, 5–14], and [30–51, 14–38], among which only the last one contains two native disulfide bonds; that is, it may be termed des[5–55]. Although the formation of the Cys5–Cys55 disulfide bond in des[5–55] would lead to native BPTI, this step finds a high energy barrier for its completion due to the highly folded structure of this intermediate. In contrast, this intermediate can rearrange its second disulfide bond to form either of the nonnative Cys5–Cys14 or Cys5–Cys38 bonds and therefore is in rapid equilibrium with the other two 2S intermediates. Subsequently, in the slowest intramolecular reaction of the pathway, these two intermediates reshuffle their nonnative disulfide bonds to render [30–51, 5–55], which

contains two of the three native BPTI disulfides (des[14–38]). In this intermediate the free thiols of Cys14 and Cys38 are on the surface and spatially close, and may oxidize rapidly to form the native inhibitor.

Weissman and Kim investigated BPTI folding with a different experimental approach using acid-trapping and reversed-phase high-performance liquid chromatography analysis of the intermediates. They proposed a revised pathway for the oxidative folding of this inhibitor, which, in contrast to that proposed by Creighton and coworkers, is characterized by the absence of stable intermediates containing nonnative disulfide bonds (Fig. 2, model a *vs.* b) (78, 79).

This results in a less heterogeneous pathway with only six significantly populous intermediates. The 1S population initially forms its disulfide bonds almost randomly, but then reshuffles rapidly to render two natively bonded species: the already-described [30–51] intermediate and a new intermediate [5–55], which also likely possesses a folded conformation. In both forms, the free Cys14 and Cys38 would be located in flexible regions preferentially exposed to solvent when compared with the accessibility of the other two free thiols. Accordingly, the formation of the Cys14–Cys38 disulfide in this 1S species would render only two 2S intermediates: [5–55, 14–38] (named des[30–51] or N*) and [30–51, 14–38] (named des[5–55] or N′). These two forms act as strong kinetic traps, freezing the folding reaction. Importantly, only half of the unfolded and reduced BPTI molecules are able to fold productively into the native form. The rest of the protein remains kinetically trapped as N*. Different structural studies have shown that N* has a native-like conformation in which the free thiols of Cys30 and Cys50 are buried and protected from the solvent, which precludes their oxidation. In this way, N* is defined as an off-pathway intermediate that is stable for weeks in this conformation. Confirming this view, it has been recently reported that redox conditions that disfavor the accumulation of N* tend to accelerate and increase the productivity of the BPTI folding reaction (52). As observed by Creighton and coworkers, N′ is a well-folded intermediate in which the missing disulfide bond between Cys5 and Cys55 cannot be directly formed to render the native BPTI structure. According to Weissman and Kim, the burial of the reactive cysteine residues inside the N′ structure forces this intermediate to unfold, to gain its native conformation. However, fluctuations inside the N′ structure promote a rearrangement reaction in which the free thiols attack the already formed disulfides in the intermediate and promote the conversion of N′ into N* and the [5–55, 30–51] intermediate (named des[14–38] or N^{SH-SH}). The stability of the N′ structure makes unfolding reactions energetically expensive so that the rate of this rearrangement is slow. In addition, the formation of N* and N^{SH-SH} from N′ is not a direct reaction but implies the creation of a transient population of two intermediates containing nonnative disulfide bonds, [30–51, 5–14] and [30–51, 5–38], which correspond to the productive nonnative intermediates in Creighton's scheme. Once N^{SH-SH} is formed, as described above, the flexibility and accessibility of the free Cys14 and Cys34 thiols allow rapid oxidation to form the native protein.

The disparity in the BPTI folding pathways proposed by both laboratories conforms to the different solution conditions in which the refolding reactions were performed and the different strategies used to trap and analyze the intermediates. With time, Weissman and Kim's view has prevailed and BPTI is considered the archetype of a disulfide-rich protein in which the folding pathway proceeds rapidly through the formation of a limited number of stable, native-like intermediates in the absence of scrambled isomers.

An important point raised by Creighton and coworkers is that inside the cell BPTI is synthesized as a pre-pro-inhibitor, containing a 22-residue pre N-terminal extension that leads to its translocation into the endoplasmic reticulum and two pro-segments of 13 and 7 residues at the N- and C-terminal sides of the mature sequence, respectively. It is not the mature form that folds in the endoplasmic reticulum but the pro-inhibitor. The pro-segment contains a cysteine residue that has been shown to react with the mature BPTI cysteines (80). However, no significant differences were observed between the folding rates of mature and pro-BPTI in microsomes (27) and the truncation of the side chain of the pro-segment's cysteine by mutation to serine did not affect the folding reaction either. Unfortunately, a detailed picture of the folding pathway of the physiologically relevant pro-BPTI form is still missing.

Leech-derived tryptase inhibitor

Human tryptase β is a trypsin-like serine protease involved in the pathogenesis of disorders such as asthma and arthritis (71). Leech-derived tryptase inhibitor (LDTI) is a nonclassical Kazal-type inhibitor isolated from the medicinal leech *Hirudo medicinalis* (14, 72). It is one of two proteins shown to bind tightly to human tryptase β, making it a molecule of considerable biomedical interest. LDTI comprises 46 residues and three disulfide bonds (Cys4-Cys29, Cys6-Cys25, and Cys14-Cys40) (Fig. 2). Its structure, solved both in solution by nuclear magnetic resonance (NMR) and in a crystalline complex with trypsin, shows a short central α-helix and a small triple-stranded antiparallel β-sheet (41, 56, 73). The first two disulfides stabilize the α-helix, joining it to the N-terminus of the protein in a cysteine-stabilized α-helical motif that is widely found in other bioactive peptides such as endothelins and toxins from insects and snakes. In the third disulfide, Cys14 in the β1-strand is linked to Cys40, restricting the mobility of the β3-strand, and thereby stabilizing the small β-sheet. This particular topology and disulfide connectivity makes LDTI an interesting model protein for oxidative folding studies.

The oxidative folding pathway of LDTI has been recently analyzed using acid-trapping, reversed-phase high-performance liquid chromatography and disulfide-pairing analysis of its intermediates (3, 59). The folding of LDTI from the reduced and unfolded state proceeds through a sequential oxidation of its cysteine residues toward the native conformation. As in the case of BPTI, although 75 different disulfide-bonded intermediates can be theoretically formed during the refolding reaction, only 5 of these species accumulate to a significant extent. First, LDTI oxidizes one of its disulfides to render 1S intermediates that rapidly reach equilibrium. The distribution of disulfide bonds in this ensemble significantly deviates from that expected from simple loop entropy contributions, with a preferential accumulation of the natively bonded [14–40] species. NMR studies of [14–40] strongly suggest that the characteristic small triple-stranded antiparallel β-sheet of LDTI is already present in this intermediate. It is likely that secondary structure propensities promote the transient formation of the β-sheet approaching Cys14 and Cys40. The formation of the disulfide bond linking the two external β-strands would further stabilize the β-sheet. The structural properties of [14–40] resemble those of [30–51] BPTI in the sense that the presence of a folded secondary structure preferentially protects the disulfide bond from reduction and kinetically favors the prevalence of this intermediate in the 1S population. Two of the three 2S folding intermediates, des[6–25] and des[4–29], are generated directly from the oxidation of two of its free cysteines. The rest of the 1S population rearranges and oxidizes its disulfide and free cysteines, respectively, to form the third 2S intermediate: des[14–40]. Therefore, the three 2S intermediates contain exclusively native disulfide

bonds. In principle, the oxidation of the two free cysteines in any of these species would result in the formation of a native and functional LDTI. However, when the intermediates were trapped and isolated and their folding was re-initiated at neutral pH, it became clear that neither des[14–40] nor des[6–25] can form directly native LDTI but rather reshuffle their disulfides to form des[4–29], which turns out to be the only productive intermediate in the folding pathway. The convergence of all 2S intermediates into des[4–29] is due to both the relative conformational properties of these species and the different reactivity of their free cysteines. The recently determined structures of these three 2S intermediates show that they are highly native (see structural discussion later) (59). In the des[4–29] structure, Cys29 at the C-terminus of the α-helix is fully exposed and can react with Cys4 at the highly flexible N-terminus to produce the native inhibitor. The des[6–25] intermediate has a very short half-life and rapidly reshuffles into des[4–29]. Cys6 has a highly reactive thiolate group. However, in the structure of the intermediate it is far away from Cys25, which precludes the direct oxidation reaction. Instead, Cys6 attacks the closer Cys4-Cys29 disulfide bond promoting a fast reshuffling to form des[4–29]. The oxidative folding of BPTI illustrated how the formation of a disulfide bond is kinetically impeded, when, once formed, becomes protected in a stable and structured conformation. This is likely the reason the des[14–40] intermediate, similar to des[5–55] and des[30–51] in BPTI, cannot oxidize directly into the native form and instead rearranges to form des[4–29]. In contrast to what is observed for des[6–25], the rearrangement of des[14–40] into the productive intermediate is a slow reaction, which promotes the accumulation of this intermediate during the folding reaction. Overall, the folding pathway of LDTI is dominated by the presence of four natively bonded intermediates without significant accumulation of scrambled forms. This simplicity in the conformational space that the protein has to sample to attain the native structure results in a fast and efficient folding pathway.

Momordica cochinchinensis trypsin inhibitor II

Momordica cochinchinensis trypsin inhibitor II (MCoTI-II) is a serine protease inhibitor belonging to the cyclotide subfamily (22). It comprises 34 residues and three disulfide bonds, with a structure similar to that of squash protease inhibitors, with their cysteine residues displaying a I–IV, II–V, and III–VI connectivity forming a classical cystine knot motif (Fig. 2) (42, 49). However, in the case of MCoTI-II the N- and C-terminus of the molecule are covalently linked resulting in a cyclic molecule (20). The combination of these two structural features endows this inhibitor with an exceptionally high stability. The oxidative folding pathway of MCoTI-II and other cyclotides has been studied and reviewed by Craik and coworkers (19, 20). The folding pathway of MCoTI-II is extremely simple, with a predominant 2S intermediate (des[1–18]; see structural discussion later) arising from the oxidation of an initially heterogeneous 1S population. This intermediate has the two native Cys8-Cys20 and Cys14-Cys26 disulfide bonds and lacks the Cys1-Cys18 disulfide that is formed by direct oxidation leading to the native protein. The formation of the third disulfide bond is the rate-limiting step in the oxidative folding reaction. Interestingly, apart from this main route there is a minor formation of other 2S intermediates that

can also attain the native structure. The energy barrier precluding the rapid oxidation of the des[1–18] intermediate into native MCoTI-II is thought to depend on a structural rearrangement that would allow orienting the reactive thiols in the free cysteines in proximity (21). The circular structure of MCoTI-II strongly reduces the entropy associated with the unfolded and reduced state of the inhibitor relative to that of other linear molecules and is likely responsible for the limited number of conformational states the protein has to sample before attaining the native structure. Again, as in the cases of BPTI and LDTI, the presence of a reduced number of natively bonded intermediates and the absence of scrambled isomers results in a fast and efficient folding process.

Oxidative Folding *via* Heterogeneous Scrambled Isomers

Hirudin

Hirudin is a thrombin-specific protease inhibitor isolated from the medicinal leech *H. medicinalis* (30). It is a 65-residue protein consisting of two functional domains: a globular N-terminal domain (NTD) formed by four β-sheets crosslinked by three disulfide bonds (Cys6-Cys14, Cys16-Cys28, and Cys22-Cys39) and a disordered negatively charged C-terminal extension (Fig. 3). The isolated functional domains possess anticoagulant activity and bind to different regions of the enzyme: the globular domain docks into the catalytic site of thrombin, whereas the C-terminal tail interacts with its fibrinogen binding site (43, 48, 68).

Chatrenet and Chang studied the oxidative folding of hirudin in the absence and presence of a thiol agent (37). In contrast to the proteins discussed in the previous section, no predominant folding pathways or preferentially populated intermediates were detected. Hirudin appears to fold following a trial-and-error mechanism in which the six cysteines are equally involved in the disulfide reshuffling reactions that lead to the native state. The main intermediates that populate the folding process of hirudin are non-natively bonded 3S species, that is, scrambled isomers. In a further study the same authors investigated the folding reaction of the globular domain of hirudin (38). The folding of this protein can be dissected into two well-differentiated stages: an initial stage of nonspecific packing and a final step of consolidation that leads to the native structure. In the first stage, reduced and unfolded hirudin oxidizes its free cysteines to sequentially form equilibrated populations of 1S, 2S, and 3S (scrambled) intermediates. All these ensembles are highly heterogeneous and there is no apparent contribution of noncovalent interactions to the intermediate distribution. Of a total of 60 potential 1S and 2S species, at least 30 accumulate in the reaction, and 11 of the 14 theoretically possible scrambled isomers could be detected (36). This finding suggests that the driving force at this initial folding stage is an unspecific hydrophobic collapse to attain a compact conformation through a nearly random disulfide pairing. Scrambled hirudin forms are more compact than 1S or 2S intermediates, but they do not possess a unique restricted conformation and accordingly, in contrast to native hirudin, are highly sensitive to proteolytic attack. During the last stage of folding this heterogeneous population of scrambled intermediates reshuffles its disulfide bonds to attain native disulfide connectivity. The consolidation of scrambled
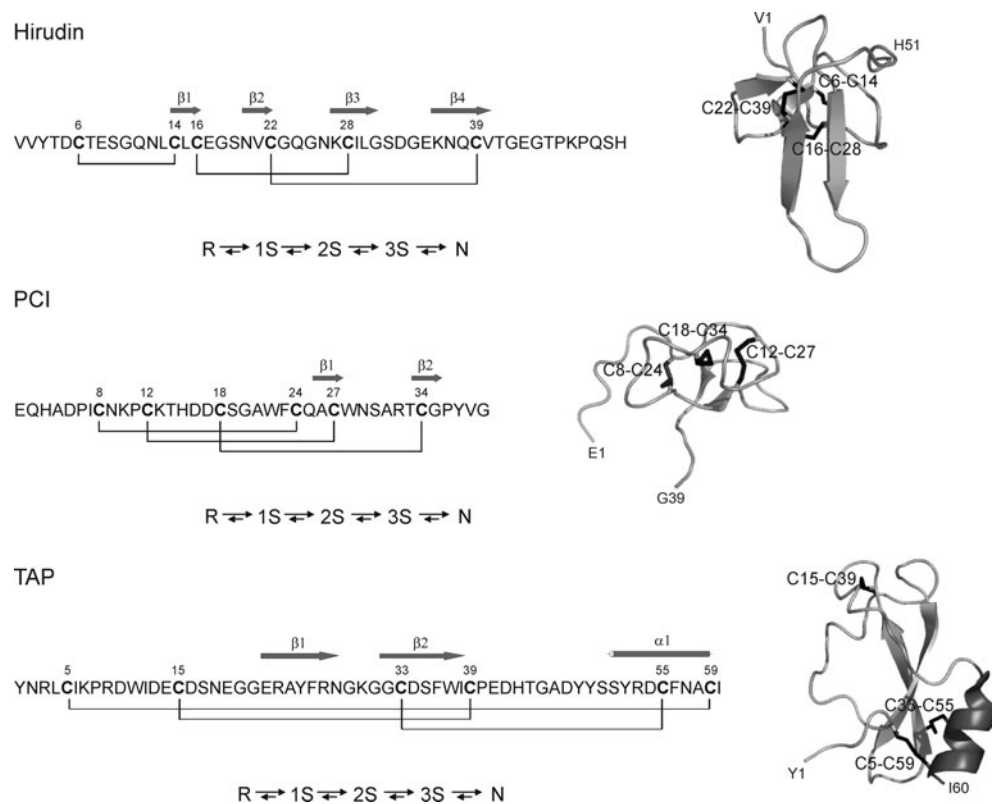
FIG. 3.   Oxidative folding of hirudin-like proteins. The amino acid sequence, secondary structure elements, and disulfide pairings of hirudin (3 SS and 51 residues), PCI (3 SS and 39 residues), and tick anticoagulant peptide (TAP; 3 SS and 60 residues) are schematically shown. The Protein Data Bank ID codes for the structures of hirudin, PCI, and TAP are, respectively, 3HTC, 1TAP, and 1H20 (ribbon plots prepared with program PyMOL). The C-terminal domain (CTD) of hirudin (residues 50–65) has a disordered structure and is not shown here. The disulfide bonds are shown as black sticks. An overview of the oxidative folding pathway of hirudin, PCI, and TAP is shown below each sequence. The nomenclature of the different species is described in the legend of Figure 2.

species is highly enhanced by the presence of a thiol catalyst; in the scrambled isomers all cysteines are involved in the formation of disulfide bonds, so there are no free thiolates to behave as catalysts in the reshuffling reaction. Therefore, scrambled isomers act as kinetic traps, their consolidation being the rate-limiting step in the folding reaction. In contrast to what was previously assumed, in the case of hirudin scrambled isomers are on-pathway intermediates and not dead-end species. The establishment of specific noncovalent interactions guides the attainment of the native structure from scrambled isomers, but these native-like contacts do not occur through a preferential pathway, which makes the overall process extremely slow and inefficient in comparison with that of BPTI-like proteins.

Potato carboxypeptidase inhibitor

Potato carboxypeptidase inhibitor (PCI) is a 39-residue globular protein that strongly inhibits metallocarboxypeptidases (64, 67). This small protein is part of the defense system of the potato plant against insect attack. Its structure is known both in solution (45) and in a crystalline complex with bovine carboxypeptidase A (64). It consists of a 27-residue core, stabilized by the presence of three disulfide bonds (Fig. 3). Cys8-Cys24 and Cys12-Cys17 form a nine-residue ring that Cys18-Cys34 is positioned through, resulting in a classical

T-knot motif. From this central core protrude a seven-residue N-terminal and a five-residue C-terminal tail, the latter being the primary binding site of PCI to the enzyme.

In contrast to hirudin, PCI lacks secondary structure elements, except for a short 3_{10} helix and a very small β-sheet. However, mechanistically, the oxidative folding pathways of these two proteins are nearly indistinguishable (9, 34, 76). As for hirudin, the folding of PCI shows an initial step of nonspecific disulfide pairing characterized by a flow from the reduced/unfolded form to scrambled isomers through the sequential formation of highly heterogeneous 1S and 2S populations. This process is not guided by noncovalent interactions since the presence of denaturants does not alter the distribution and/or flow of the different species that accumulate. Subsequently, scrambled isomers reshuffle their disulfide bonds to produce the native inhibitor. Again, from a kinetic point of view, the process of consolidation represents the rate-limiting step in the folding reaction. The oxidative folding of PCI is an extremely inefficient process, being about 10 times slower than that of hirudin. The absence of a predominant secondary structure that could efficiently drive the reaction toward a global minimum and/or the high energy barrier associated with the exposure of hydrophobic residues (partially hidden in the core of scrambled PCI forms) necessary to attain the native fold might account for this behavior.

Similar to BPTI, PCI is synthesized *in vivo* as a precursor protein that includes a 27-residue N-terminal and a seven-residue C-terminal pro-region. Comparison of the oxidative folding of mature PCI and PCI containing the N-terminal pro-sequence indicated that this molecule does not depend on this pro-region for its folding *in vitro*, or *in vivo* in *Escherichia coli* (18).

Tick anticoagulant peptide

Tick anticoagulant peptide (TAP) is a small serine protease inhibitor purified from extracts of the soft tick *Ornithodoros moubata* (77). This 60-residue protein is a slow tight-binding inhibitor specific for the blood coagulation factor Xa. Its structure was solved in solution by NMR (1, 53). Importantly, TAP was shown to be a Kunitz-type inhibitor, with an overall conformation closely related to that of BPTI, that is, consisting of a double-stranded antiparallel β-sheet, an α-helix, and a small hydrophobic core. In addition, BPTI and TAP share the same disulfide connectivity, with TAP's Cys5-Cys59, Cys15-Cys39, and Cys33-Cys55 disulfides being structurally homologous to BPTI's Cys5-Cys55, Cys14-Cys38, and Cys30-Cys51 disulfides, respectively. Therefore, in principle, one should expect that the oxidative folding of TAP would proceed efficiently through a preferential route displaying a limited number of stable, native-like intermediates. Instead, reduced and unfolded TAP acquires its native structure following a pathway that is, in many aspects, indistinguishable from that of hirudin and PCI (Fig. 3). Ten of 15 theoretically possible 1S isomers were detected by Chang and coworkers at the initial folding steps (31), indicating a quasi-random formation of the first disulfide and, in contrast to what happens with BPTI, the absence of selective noncovalent contacts that can structurally and kinetically promote the accumulation of a preferential 1S intermediate. The species in the equilibrated 1S population oxidize two of their free cysteines to render a heterogeneous 2S ensemble in which, again, no structural bias is observed. Upon equilibration the 2S intermediates form the last disulfide to produce scrambled isomers that become trapped in the absence of thiol catalysts. Although the scrambled ensemble is still heterogeneous, 4 of the 14 possible 3S isomers represent about 70% of the population. This reduction in the heterogeneity of the conformational space is not the result of the establishment of selective contacts between side chains, but rather of an unspecific hydrophobic collapse, since a common feature of these major scrambled isomers is that none contain native disulfides, and in all cases the detected disulfide connectivity is incompatible with the presence of a native double-stranded β-sheet. As for hirudin and PCI, the last stage in the oxidative folding of TAP is the consolidation of the scrambled isomers to acquire the native conformation. For TAP, all major scrambled forms are able to convert into the native protein at the end of the reaction, but before attaining this conformation they equilibrate with the other scrambled forms. Isolated scrambled forms equilibrate faster with other scrambled species sharing a common disulfide bond, because such conversions imply reorganization of only two pairs of disulfides.

The heterogeneity of folding intermediates, the sequential flow from the reduced and unfolded protein through oxidation of disulfides, and the inefficiency of the last stage of folding in the absence of thiol catalysts due to the large ac-cumulation of scrambled isomers are characteristics reminiscent of those exhibited by hirudin and PCI. However, there is one important difference concerning the stability and conformation of the intermediates in the folding pathway. In the case of TAP, all three classes of intermediates, that is, 1S, 2S, and 3S, contain species that are susceptible to denaturant, suggesting that their structures are locally stabilized by noncovalent interactions. Although the free energy associated with these species is likely to be marginal, especially for 1S and 2S, overall they act as kinetic traps that slow down the transition toward the subsequent equilibrated population. In the presence of a denaturing agent that disrupts these contacts, the speed and efficiency of the folding reaction are significantly increased. These results suggest that the presence of stable intermediates does not necessarily facilitate folding; they can impede kinetically the flow of intermediates and delay the attainment of the functional fold (35).

Oxidative Folding *via* Native and Scrambled Isomers

BPTI- and hirudin-like proteins exemplify two opposite oxidative folding landscapes. As illustrated by the cases of BPTI and TAP, the oxidative folding of a protein is not determined exclusively by its global protein conformation and/or disulfide connectivity. Although the folding behavior of both types of proteins should be somehow encoded in the primary polypeptide sequence, it is still not possible to decipher how this linear information determines the complex route that a small disulfide-rich protein has to follow to gain its unique functional structure. Thus, the study of protein models displaying mixed pathways in which both native and scrambled intermediates accumulate is of much interest to provide insight into the mechanisms of oxidative folding.

Leech carboxypeptidase inhibitor

Leech carboxypeptidase inhibitor (LCI) is a 67-residue metallocarboxypeptidase inhibitor found in the medicinal leech *H. medicinalis* (66). Interestingly, it is a potent inhibitor of plasma carboxypeptidase B, also known as thrombin-activatable fibrinolysis inhibitor, which is an emerging biomedical target for thrombolysis and thrombolytic therapies. The structure of LCI was solved both in solution and in a crystalline complex with human carboxypeptidase A2 (65). LCI folds in a compact domain comprising a five-stranded antiparallel β-sheet with a $\beta3$-$\beta1$-$\beta2$-$\beta5$-$\beta4$ topology and a short α-helix that packs onto the most compact part of the β-structure (Fig. 4). The native conformation of the inhibitor is highly stabilized by the occurrence of four disulfide bonds, Cys11-Cys34, Cys18-Cys62, Cys19-Cys43, and Cys22-Cys58, all of them located within regular secondary structure elements. The disordered C-terminal tail of LCI is the primary binding site of the enzyme, as previously described for PCI.

LCI displays a higher extent of secondary structure than its three-disulfide protease inhibitor counterparts, which affects the accessibility, proximity, and reactivity of the thiols and disulfide bonds. In addition, the presence of two additional cysteine residues increases almost 10-fold the number of intermediates that could potentially arise during the oxidative folding reaction. The combination of these two factors results in a rather complex folding landscape (4, 69). The folding of LCI proceeds initially through a heterogeneous formation of 1S and 2S intermediates. The presence of moderate amounts
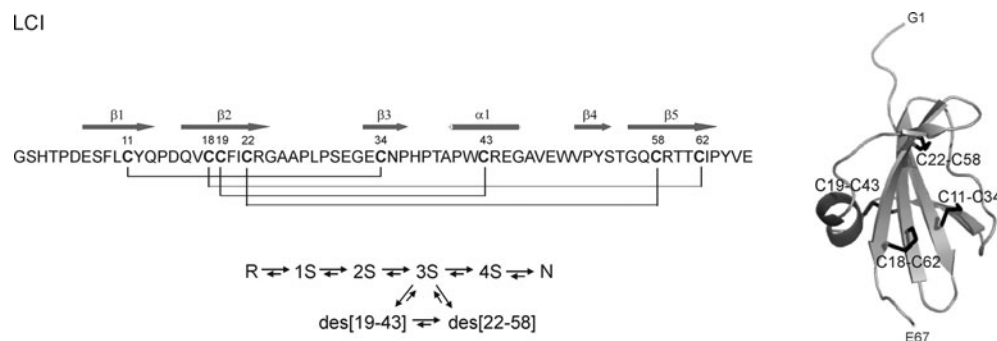
FIG. 4. Oxidative folding of leech carboxypeptidase inhibitor (LCI). The amino acid sequence of LCI (4 SS and 67 residues), as well as its secondary structure elements and disulfide pairings, is schematically shown. The Protein Data Bank ID code for the structure is 1DTV (ribbon plot prepared with program PyMOL). The disulfide bonds are shown as black sticks. An overview of the oxidative folding pathway of LCI is shown at the bottom. The nomenclature of the different species is described in the legend of Figure 2.

of denaturing agents does not alter the number or relative abundance of 1S and 2S species, suggesting that their distribution at equilibrium does not respond to the establishment of specific side chain contacts, but rather to entropy constraints. This is in good agreement with computational predictions, suggesting that 1S and 2S LCI species do not contain an extensive regular structure (6). This heterogeneity is strongly reduced in the 3S population, which was initially thought to consist exclusively of two intermediates displaying native disulfide connectivity: des[19–43] and des[22–58]. Because no other 3S forms were detected it was assumed that these two native-like species could originate directly from the oxidation of the third disulfide bond in the 2S ensemble. However, the lack of an enthalpic bias toward the formation of native disulfides in 2S implies that only a minor fraction of the isomers in this ensemble would display a native combination of disulfide bonds, making the direct and preferential formation of natively bonded 3S species a very improbable reaction. In a recent study, the contribution of each particular disulfide to the folding of LCI was investigated by constructing a series of single and multiple mutants lacking one to four disulfide bonds (6). This work allowed reformulating the folding pathway of this protein, specifically the formation, distribution, and role of 3S isomers in the folding reaction. The results demonstrated that in LCI 3S isomers can be effectively formed in the absence of the fourth disulfide bond and, more importantly, that these wrongly bonded species can rearrange their disulfides to render 3S forms with native disulfides and native-like conformations. Therefore, a more plausible mechanism for the formation of des[19–43] and des[22–58] intermediates would be the nonspecific oxidation of the unfolded 2S ensemble to 3S species and the subsequent reshuffling of these forms into these two predominant intermediates. Thus, the initial stages of LCI folding comprise a quasi-stochastic consolidation of the protein, and resemble those of hirudin. Nevertheless, des[19–43] and des[22–58] intermediates are reminiscent of the intermediates present in the oxidative folding of BPTI, in the sense that in contrast to the other 3S species, they have a highly folded structure. This endows them with the highest stability in the ensemble, and at the same time protects their native disulfide bonds from reduction and/or reshuffling, which kinetically favors the large accumulation of these species in the 3S ensemble (7, 11). The high stability and low reactivity of des[19–43] and des[22–58]

cause them to act as strong kinetic traps that freeze the reaction at the 3S stage. Consequently, the destabilization of such intermediates by denaturing agents allows the progress of the pathway and accelerates the recovery of native protein. In principle, the direct oxidation of the free thiolates in any of these intermediates would promote the formation of the native inhibitor. The formation of the native conformation is thermodynamically favored, since it is at least two times more stable than any of the intermediates. However, in contrast to what happens with BPTI and LDTI, the free thiolates are not in flexible and exposed loops but shielded by the secondary structure elements (see structural discussion later), as demonstrated by their inability to react with external thiols. This hidden location of the free cysteines in the intermediates has a dramatic effect on the progress of the folding reaction because structural fluctuations expose both the thiols and the disulfide bonds, thus promoting their reshuffling instead of the oxidation of free thiols into the native disulfide pairing. Therefore, the gained native structure becomes lost and a population of compact but unstructured 4S scrambled isomers subsequently accumulates. As for hirudin and PCI, the reshuffling of these disulfide-blocked species is very slow in the absence of an external thiol catalyst. The reorganization of the misconnected disulfide bonds to finally attain the native form becomes the major rate-limiting step in the LCI folding reaction. In the folding pathway of LCI both scrambled and native intermediates are necessary to attain a unique stable conformation. In addition, in contrast to general assumptions, scrambled but not native-like intermediates are the productive species. LCI illustrates how the search for the conformational stability and reduced flexibility that provide the native disulfide connectivity complicates the folding reaction and makes it extremely inefficient, suggesting that the conformational space that small disulfide-rich proteins have to explore to attain the functional form is shaped primarily in thermodynamic terms.

Oxidative Folding of Two-Domain Protease Inhibitors

Approximately 75% of eukaryotic protein sequences fold into conformations consisting of more than one domain (15). However, information about the mechanisms driving the folding of two-domain proteins is still scarce. This is mainly due to the limitation of indirect structural probes such as intrinsic

fluorescence or ellipticity to monitor folding/unfolding transitions in two-domain proteins where the two protein moieties can fold/unfold autonomously or in a cooperative way. During oxidative folding the formation of disulfide bonds acts as a covalent structural probe, providing unbiased information about the proximity of the different regions of the polypeptide chain. Therefore, studies of multidomain proteins in which the different moieties are stabilized by disulfide bonds might allow characterization of the extent to which the folding reactions of this kind of polypeptide relies on the establishment of inter-domain interactions. Recent studies on the oxidative folding of two-domain protease inhibitors have provided valuable insights into the complicated folding landscape that multidomain proteins must face before attaining their native conformation.
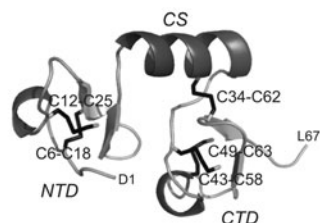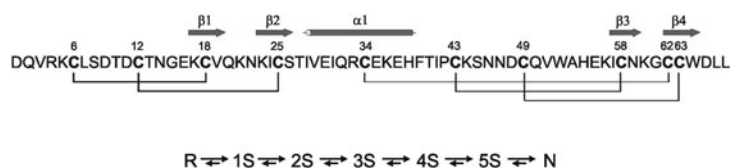
Ascaris carboxypeptidase inhibitor

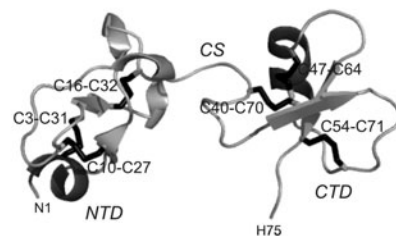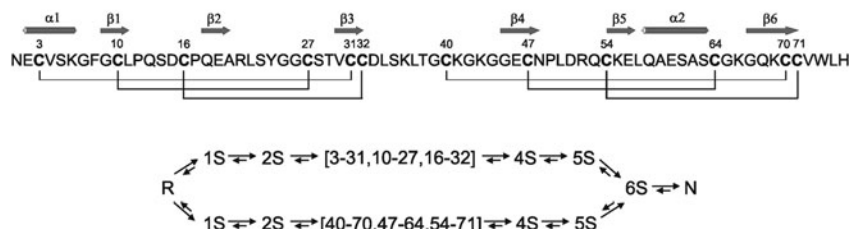The metallocarboxypeptidase inhibitor *Ascaris* carboxy-peptidase inhibitor (ACI) isolated from *Ascaris*, the most common human parasite of the gastrointestinal tract that causes the pandemic disease ascariasis, is 67 residues long and contains five disulfide bonds (70). The crystal structure of ACI in complex with human carboxypeptidase A1 reveals two structurally similar tandem modules (Fig. 5), an NTD and a C-terminal domain (CTD), which are linked by a connecting segment (CS). Each domain comprises a small helix and a C-terminal β-sheet that are stabilized by two disulfide bonds (Cys6-Cys18, Cys12-Cys25; and Cys43-Cys58, Cys49-Cys63). The CS includes an α-helix, which is linked to the β-sheet of the CTD by the Cys34-Cys62 disulfide bond.

The oxidative folding of ACI has been recently examined, showing that this protein folds through the sequential formation of 1S, 2S, 3S, and 4S intermediates that subsequently evolve into a mixture of 5S (scrambled) isomers (13). Similar to the cases of hirudin, PCI, TAP, and LCI, the reshuffling of this heterogeneous population of scrambled isomers into the native protein represents the rate-limiting step of the folding reaction and is efficiently promoted by the presence of thiol
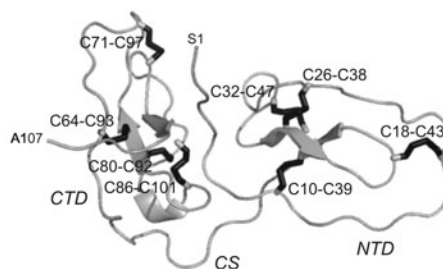

FIG. 5. Oxidative folding of two-domain protease inhibitors. The amino acid sequence, secondary structure elements, and disulfide pairings of *Ascaris* carboxypeptidase inhibitor (ACI; 5 SS and 67 residues), tick carboxypeptidase inhibitor (TCI; 6 SS and 75 residues), and secretory leukocyte protease inhibitor (SLPI; 8 SS and 107 residues) are schematically shown. The Protein Data Bank ID codes for the structures of ACI and TCI are, respectively, 3FJU and 2JTO (ribbon plots prepared with program PyMOL). The disulfide bonds are shown as black sticks. The N-terminal domain (NTD) and CTD and the connecting segment (CS) are indicated. An overview of the oxidative folding pathway of ACI, TCI, and SLPI is shown below each sequence. The nomenclature of the different species is described in the legend of Figure 2.

catalysts. However, the presence of redox agents does not prevent the final accumulation of a major scrambled isomer, which elutes very close to native ACI in the reversed-phase chromatogram. This scrambled species could arise from the presence of two consecutive cysteines located at the C-terminus of ACI, which could easily lead to divergent disulfide pairings (see also *Tick carboxypeptidase inhibitor* section). This nonproductive species is not expected to accumulate *in vivo* since the presence of protein disulfide isomerase, which catalyzes the formation and isomerization of disulfide bonds in the eukaryotic endoplasmic reticulum, efficiently prevents its formation. The folding analysis of truncated ACI variants lacking the NTD or both the NTD and CS revealed that the NTD has an influence on the folding of the CTD, so that the two domains of ACI do not fold autonomously, but rather cooperatively. It appears that the NTD assists in the folding of the CTD, most likely through transient interdomain disulfide interactions, acting as a kind of intramolecular chaperone. This result provides a plausible explanation for the two-domain architecture of this molecule, since the NTD does not contribute significantly to the inhibition of target enzymes.

Tick carboxypeptidase inhibitor

Tick carboxypeptidase inhibitor (TCI) is a protein of 75 residues and six disulfide bonds isolated from the hard tick *Riphicephalus bursa* (8). Like LCI, it is a tight binding inhibitor of thrombin-activatable fibrinolysis inhibitor able to stimulate the fibrinolysis of blood clots and thus constituting a promising molecule for thrombolytic therapies (8). The crystal structure of TCI in complex with different metallocarboxypeptidases revealed that it consists of two domains that are structurally similar and contribute to the inhibition of the enzyme in a double-headed manner (12). Each domain consists of a short α-helix followed by a small twisted antiparallel β-sheet, a motif structurally homologous to that of β-defensin-like proteins (Fig. 5). The solution structure of TCI has recently shown that these two rigid globular domains are joined by a flexible linker (60). Each domain is strongly constrained by the presence of three disulfide bonds with identical I–V, II–IV, and III–VI connectivity.

The oxidative folding of TCI from the reduced and unfolded protein to the native state proceeds through a sequential oxidation of cysteine residues, which is characterized by the presence of intermediates displaying one to six disulfide bonds (5). The overall mechanism recalls that of the single-domain LCI, because both scrambled and natively bonded species are obligatory intermediates in the folding pathway. The presence of 12 free cysteines implies a large number of possible intermediates. Indeed, numerous species are found in the 1S and 2S population, without any detectable enthalpic bias. However, the equilibrated 3S ensemble is clearly dominated by the accumulation of only two folding intermediates that comprise three native disulfide bonds in one domain and no disulfide bonds in the other: [3–31, 10–27, 16–32] and [40–70, 47–64, 54–71]. Importantly, nonnative 3S species could not be detected, indicating that they either do not form or, more likely, are kinetically unstable relative to the natively connected folding intermediates. The structures of the 3S isolated intermediates support this view (see structural discussion later), because in both cases the disulfide-linked domain displays a stable native-like conformation, which

protects disulfides from reduction and reshuffling (10). According to the TCI structure any other combination of disulfides would result in at least one of the bonds being accessible to solvent and therefore potentially reactive. The folding reaction proceeds through a sequential oxidation of the free thiols in the intermediates to finally render an equilibrated population of 6S scrambled isomers. The folding of TCI finds difficulty reaching completion in the absence of thiol catalysts, and only 15% of the protein attains the native conformation, even after 2 days of refolding. The reshuffling reaction is exceedingly slow and therefore constitutes the major rate-limiting step of the folding pathway. Similar to ACI, even in the presence of redox agents there is a final accumulation of a stable scrambled isomer that elutes closely to native TCI. Again, this nonproductive species is avoided by the addition of protein disulfide isomerase to the folding reaction mixture.

The oxidative folding of TCI comprises the accumulation of many disulfide intermediates. In principle, this complexity could arise from the formation of nonnative interdomain disulfide bonds. Nevertheless, this type of disulfide could not be detected, suggesting that the two domains fold autonomously without any extensive interaction between both protein moieties. This lack of cooperation between the two domains occurs already at the early stages of folding that lead to the formation of the two major 3S intermediates. This fact is evidenced by the ability of individual TCI domains to acquire a native disulfide pairing from their respective reduced and unfolded states. Importantly, despite their homologous structures and identical disulfide connectivity, the two domains of TCI fold at very different rates and through distinct mechanisms. The CTD folds rapidly through a limited number of intermediates, whereas in the case of the NTD the reaction is slower and the intermediates are far more heterogeneous. In both TCI domains the formation of scrambled isomers precedes and hinders the acquisition of native disulfide connectivity. The formation of two different 3S intermediates in which only one domain is folded is an exceptional property of TCI. Usually, when the different moieties of a multidomain protein fold autonomously, only the thermodynamically most stable or kinetically most favored domain is able to acquire a native structure before the other domains can fold. In TCI the two major 3S intermediates can independently evolve through sequential oxidation to produce the native protein. Overall, the folding of TCI results from the sum of the folding pathways of its individual domains and exhibits a previously unreported symmetry in which, independent of the domain that folds first, the second can fold in a noncooperative manner to render the functional inhibitor. The sequential and autonomous folding of two 3S domains clearly implies a much less complex conformational search than that associated with the cooperative formation of six disulfide bonds, in which case >10,000 possible 6S isomers are possible.

Secretory leukocyte protease inhibitor

Secretory leukocyte protease inhibitor (SLPI), also named mucous proteinase inhibitor, is a 107-residue serine protease inhibitor found in human mucous fluids (74). Its physiological function seems to be the protection of the mucosal epidermis against degradation by proteases, especially those contained in granulocytic lysosomes, which suggests a possible

therapeutic use against local inflammatory processes. The crystal structure of SLPI in complex with α-chymotrypsin shows that the inhibitor folds in two independent domains in a boomerang-like structure (Fig. 5) (47). The C-terminal sequence of each domain organizes into two β-strands linked by a hairpin turn, whereas the N-terminal part consists of a loop covalently linked to the β-sheet, which constitutes the primary binding site of the protease. Each domain is stabilized by the presence of four disulfide bonds with I–VI, II–VII, III–V, and IV–VIII connectivity.

The presence of 16 free cysteine residues in the reduced and unfolded state of SLPI could result in an extremely complex oxidative folding reaction for this inhibitor. More than 40 million different species could accumulate along the pathway, with 2 million corresponding to 8S scrambled isomers; that is, the folding reaction would be exceedingly inefficient. Lin and Chang analyzed the oxidative folding of SLPI, finding that, surprisingly, despite the complicated theoretical folding landscape of SLPI, this protein folds very efficiently into the two-domain native architecture without requiring a thiol catalyst (54). This property is unique and intriguing in light of the SLPI native structure, which lacks an extensive regular structure that could drive a coupled conformational and oxidative folding reaction. Even in the case that the protein adopts a folding pathway in which only native disulfide bonds are formed, 254 different species would populate the pathway. Moreover, the folding pathway of SLPI comprises a highly heterogeneous population of folding intermediates that includes scrambled isomers, which discredits the idea that folding occurs through a BPTI-like mechanism. In fact, the folding of SLPI follows a two-stage mechanism that combines features of both BPTI- and hirudin-like proteins. During the first stage of folding, reduced and unfolded SLPI sequentially oxidizes its disulfide bonds to form a highly heterogeneous mixture of intermediates that finally converges in the formation of 6S and 7S isomers comprising nonnatively and natively bonded intermediates. The main 6S intermediate is des[18–43, 71–97], which contains the three cysteines located in the β-strands of each domain correctly bonded, forming I–VI, III–V, and IV–VIII disulfide bonds, and lacks the II–VII bond joining the protease binding loops in both domains. The main 7S intermediate corresponds to des[18–43], which lacks the II–VII bond in the first domain. Presumably both intermediates display a partially folded conformation, which protects preferentially the linked disulfides promoting its accumulation and its role as a kinetic trap in the flow of the reaction. In the second stage of folding, des[18–43, 71–97] converts into des[18–43] through direct oxidation of its Cys71 and Cys97 residues, and the newly formed and previously accumulated des[18–43] oxidizes into the native SLPI. Similar to BPTI, these reactions are favored by the presence of free cysteines in flexible and exposed loops. This is likely the reason that they can act as thiol catalysts, facilitating the reshuffling of the remaining nonnatively bonded isomers toward native pairings that further evolve into native intermediates and finally form the functional inhibitor. It also explains why SLPI folds efficiently in the absence of external redox agents. Therefore, in contrast to TCI and more similar to ACI, the folding of SLPI comprises interdomain interactions and is a cooperative reaction.

Structures of Protease Inhibitors' Folding Intermediates

The paradigmatic cases of BPTI and TAP illustrate how the native nature and heterogeneity of the different intermediates that populate the folding pathway of a given protein cannot be directly inferred from its native structure and/or disulfide-bond connectivity. This situation is in contrast to the case of proteins folding *via* a two-state mechanism, which constitute the other major benchmark for the characterization of folding reactions. In two-state folders, the mere topology of the native state determines the overall protein folding rate and the placement of transition states in the pathway (61), in such a way that proteins featuring primarily local contacts tend to fold more rapidly and exhibit less compact folding transition states than those displaying more nonlocal interactions. This rule does not apply to oxidative folding because conformational folding is obligatorily coupled to disulfide-bond formation. This means that the establishment of nonnative local contacts can be forced by nonnative disulfide pairings or, on the other hand, that native local contacts and secondary/tertiary structure formation can be strongly accelerated by the formation of native covalent links. At the same time, the progress of conformational folding influences the proximity, accessibility, and reactivity of free thiols and disulfide bonds. This makes the prediction of folding pathways in small disulfide-rich proteins a daunting task. Overall, the complex process of oxidative folding appears to depend on the specific conformational properties of the intermediates that accumulate in the reaction, rather than in the entropically and enthalpically stabilized final conformation. Therefore, the structural characterization of disulfide folding intermediates is becoming a central issue in oxidative folding. So far such structural studies have largely focused on analogs where one or more disulfide bonds were chemically blocked or deleted by mutation of cysteines to alanines or serines, for example, in the special cases of BPTI or RNase A (2). However, chemically inert analogs suffer from steric strain and repulsion due to the volume and net charge of the blocking reagents, and genetically engineered analogs have been shown to differ frequently in stability and conformation from the corresponding genuine intermediates (6). In the last five years, seven genuine intermediates isolated directly from folding/unfolding reactions have been structurally characterized in acidic solution (*i.e.*, conditions where they are stable) by NMR spectrometry (Fig. 6). Notably, all of them correspond to protease inhibitor models: one intermediate from LCI, one from MCoTI-II, two from TCI, and three from LDTI (7, 10, 21, 59). These intermediates are des species lacking from one to three disulfide bonds, but in all cases the already-formed disulfides display native connectivity.

From a general point of view, oxidative folding intermediates lacking a single disulfide bond can be classified as productive or nonproductive species, where productive intermediates are those that can directly oxidize their free thiols to render the native and functional protein. As proposed by Scheraga and coworkers, they can be also classified as "disulfide-secure" or "disulfide-insecure" intermediates according to the reactivity of the already formed native disulfide bonds (57, 81). Nonproductive intermediates tend to be disulfide-insecure; they are normally long-lived, native-like, stable forms that act as strong kinetic traps during oxidative folding reactions. The solution structure of des[22–58] LCI
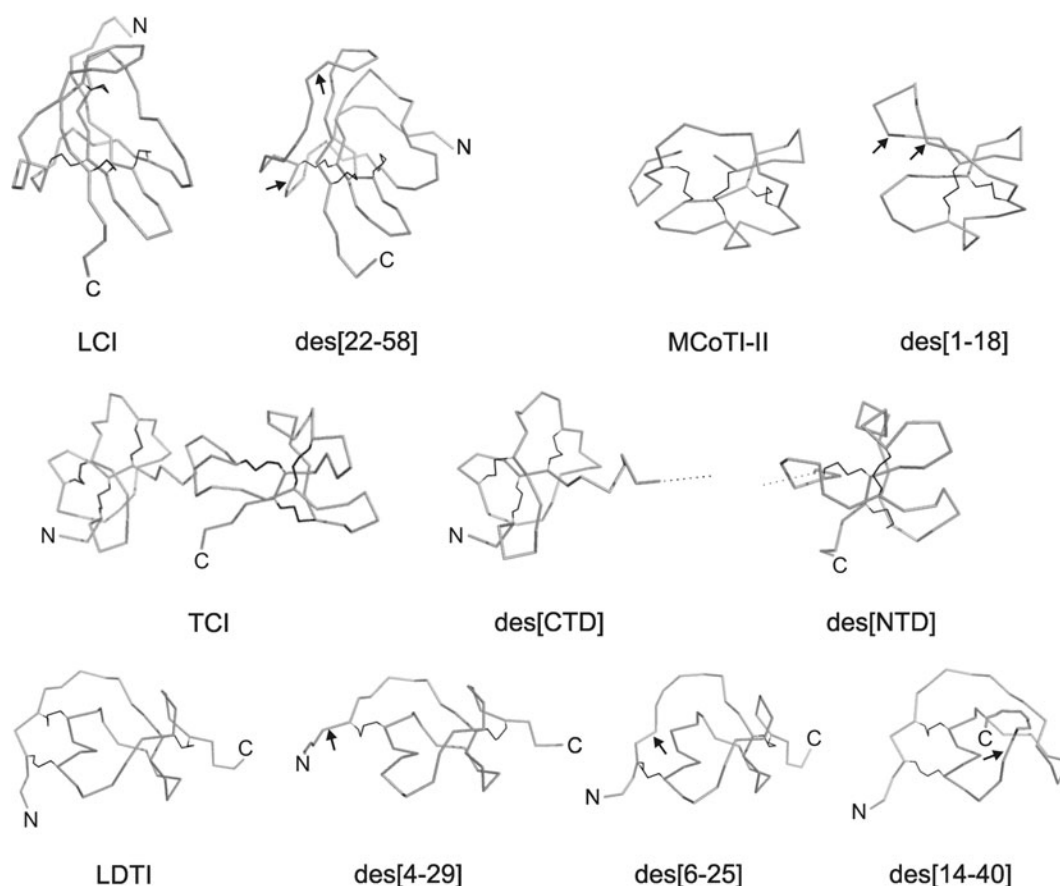
FIG. 6.   Solution structures of major folding intermediates. Cα-structure representation of the average structures of LCI, MCoTI-II, TCI, and LDTI (native forms), together with those of their major folding intermediates—missing the disulfide bonds shown between parentheses (figure prepared with program PyMOL). The Protein Data Bank ID codes are as follows: LCI (1ZFI), des[22–58] (1ZFL), MCoTI-II (1IB9), des[1–18] (2PO8), TCI (2K2X), des[CTD] (2K2Y), des[NTD] (2K2Z), LDTI (2KMO), des[4–29] (2KMP), des[6–25] (2KMQ), and des[14–40] (2KMR). For clarity, the molecules are in the same orientation as their ribbon plot structures in Figures 2, 4, and 5. Regions with increased flexibility are indicated by *arrows*. Dashed lines represent the unfolded and unpaired domains of TCI intermediates.

shows the molecular determinants responsible for such behavior (7). This intermediate possesses a well-defined globular conformation that closely resembles that of the native inhibitor, including a four-stranded antiparallel (β3-β1-β2-β5) β-sheet, although certain regions are more disordered than in the native state and the α-helix is not properly formed. Despite the increased flexibility and lower compactness of the intermediate, proton/deuterium exchange experiments demonstrated that both free cysteines (Cys22 and Cys58) and the three native disulfide bonds are located in protected regions and have limited accessibility. This restricted exposure to solvent is also evident from the fact that both free thiols and disulfide bonds are insensitive to the presence of external thiol agents. The burial of the free thiol groups inside a native-like structure would hinder any further progress in the folding reaction. However, the increased overall flexibility of the backbone promotes structural fluctuations that allow the intermediate to escape from this trap. Because of their similar degrees of protection, fluctuations that expose the free thiols also expose the disulfide bonds, resulting in disulfide reshuffling into 4S scrambled forms instead of direct oxidation of the missing disulfide bond into the native structure. In

des[22–58] LCI reshuffling can compete effectively with oxidation due to the conformational similarity between a reshuffling reaction and the second step of an oxidation reaction (81).

In contrast to des[22–58] LCI, the des[1–18] intermediate of MCoTI-II is a productive species able to directly render the native inhibitor by direct oxidation of its free thiols. The solution structure of this intermediate shows that it displays native secondary and tertiary structure, the only exception being a flexible loop, the movement of which in the native state is restricted by the Cys1–Cys18 disulfide bond, forcing its collapse against the protein core. The absence of the covalent link in the intermediate results in a loop more loosely associated with the rest of the protein (21). However, proton/ deuterium exchange experiments demonstrated that the number of slowly exchanging amide protons in the intermediate is comparable to that of the native protein, indicating that both forms display similar rigidity. The rigidity in the overall scaffold of des[1–18] MCoTI-II turns it into a disulfide-secure intermediate in which structural fluctuations preferentially change the environment of their thiol groups while keeping their disulfide bonds in a protected native location.

This is reflected by the surface accessibilities of the two free cysteines in the intermediate, which are much larger than those in the native state, where they become part of the hydrophobic core. Molecular dynamic simulations suggest that in the absence of the disulfide bond the rigid native conformation is thermodynamically destabilized with respect to the more flexible intermediate structure. Therefore, although the free cysteines are ready for oxidation and do not compete with the disulfide bonds, the fact that the protein backbone must adopt an energetically unfavorable conformation before both cysteines can oxidize slows down significantly the productive formation of the native form.

The two main 3S folding intermediates of the two-domain inhibitor TCI illustrate how the preferential burial of native disulfide bonds into a compact native structure protects them from reshuffling reactions that could be initiated by reactive free cysteines. As discussed above, each intermediate comprises three native disulfide bonds in one domain and no disulfide bonds in the other. This means that each intermediate has six free thiols that can potentially initiate a disulfide exchange reaction with the preformed disulfide bonds. Moreover, the solution structures of these two intermediates indicate that these free cysteines are found in a disordered and highly flexible context, which in principle would strongly favor the disulfide attack (10). However, this type of reaction does not occur and the already formed covalent bonds remain intact in both intermediates during the rest of the folding reaction. The reason for this protection is that the structures of the folded NTD and CTD in the intermediates are exactly the same as the corresponding ones in the native inhibitor, with all the disulfide bonds highly buried and protected from solvent in a globular and highly compact conformation. Therefore, the folding pathway proceeds with the oxidation of the free cysteines in the unfolded domain.

LDTI is the first protein model for which all the genuine des species preceding the formation of the native structure have been structurally characterized (59). As discussed above, they consist of three intermediates each containing two native disulfide bonds and a pair of free thiols: des[4–29], des[6–25], and des[14–40]. Among these species only des[4–29] is a productive intermediate that yields the native inhibitor by direct oxidation of its free thiols. The other two intermediates reorganize their disulfide bonds to converge in the productive intermediate. LDTI shows a reduced amount of secondary structure, that is, a short α-helix and a small triple-stranded antiparallel β-sheet in a β2-β1-β3 topology. Despite the overall packing of LDTI, which was thought to rely on the special Kazal-type connectivity of its disulfide bonds, the three des intermediates display essentially the same fold and secondary structure motifs as native LDTI, except des[14–40], which lacks the β3-strand due to the missing disulfide. Importantly, the nonproductive des[6–25] and the productive des[4–29] intermediates posses a well-packed and structurally identical core, with a rms-deviation of only 1.3Å, the main difference between them being a more flexible N-terminus in des[4–29]. In contrast to LCI, in LDTI these productive and nonproductive intermediates display almost identically accessible free thiols and disulfide bonds. In both intermediates the Cys14–Cys40 disulfide is buried in the native β-sheet, whereas the other disulfide connecting the α-helix to the N-terminal loop and the free thiols are fully accessible. This indicates that the productivity of a folding intermediate depends on more subtle structural determinants than the mere exposure of its free or bonded cysteine residues. In fact, the balance between reshuffling and oxidation depends on the accessibility, but also on the proximity and reactivity, of the groups involved. In des[6–25] the two latter factors shift the balance toward reshuffling into des[4–29] rather than toward the formation of the native inhibitor *via* oxidation. Cys25 is located at the N-terminus of the α-helix where a partial positive charge of the dipole stabilizes the thiolate, lowering the effective pK_a of this group and making it highly reactive. The flexibility of the backbone in the intermediate brings this reactive cysteine close to the accessible Cys4–Cys29 disulfide bond instead of to the free Cys6, which results in disulfide exchange instead of oxidation. Together, the structures of the LDTI folding intermediates suggest that the accessibility of the free thiols ensures that an intermediate is a disulfide-secure, productive species only if all the native disulfide bonds are fully buried and cannot be attacked by the free thiols.

Concluding Remarks

Protease inhibitors comprise an extremely large number of molecules with different sequential and structural characteristics. They are usually small proteins crosslinked by several disulfides, which have turned them into outstanding models for the study of the complex process of oxidative folding. Although only two classes of protease inhibitors have been studied to date, namely, serine and metalloprotease inhibitors, a surprisingly high diversity of folding scenarios have been detected. Expanding folding studies toward the unexplored cysteine and aspartic protease inhibitors would likely uncover new folding processes. Unfortunately, it is still not possible to forecast the folding landscapes that small disulfide-rich proteins have to face by using only the conformation and covalent connectivity of their native structures. The aforementioned structures of genuine folding intermediates are providing the first clues to the molecular determinants that guide the last stages of oxidative folding. However, structural information is still scarce regarding the initial folding steps preceding the formation of stable intermediates. In the forthcoming years research should be focused on these fast and intricate but crucial events. The use of well-established approaches for the study of fast-folding proteins will be of much help. Understanding the general rules that govern oxidative folding at different stages would provide a unique opportunity to attract theoretical groups to this research area and guide the development of bioinformatics approaches to assist in the prediction and design of the folding pathways of small disulfide-rich proteins. What becomes obvious is that protease inhibitors will, in the years to come, provide one of the main bedrocks from which to understand the complexity of oxidative folding.

Acknowledgments

References

1. Antuch W, Guntert P, Billeter M, Hawthorne T, Grossenbacher H, and Wuthrich K. NMR solution structure of the recombinant tick anticoagulant protein (rTAP), a factor Xa inhibitor from the tick *Ornithodoros moubata*. *FEBS Lett* 352: 251–257, 1994.

2. Arolas JL, Aviles FX, Chang JY, and Ventura S. Folding of small disulfide-rich proteins: clarifying the puzzle. *Trends Biochem Sci* 31: 292–301, 2006.

3. Arolas JL, Bronsoms S, Aviles FX, Ventura S, and Sommerhoff CP. Oxidative folding of leech-derived tryptase inhibitor via native disulfide-bonded intermediates. *Antioxid Redox Signal* 10: 77–85, 2008.

4. Arolas JL, Bronsoms S, Lorenzo J, Aviles FX, Chang JY, and Ventura S. Role of kinetic intermediates in the folding of leech carboxypeptidase inhibitor. *J Biol Chem* 279: 37261–37270, 2004.

5. Arolas JL, Bronsoms S, Ventura S, Aviles FX, and Calvete JJ. Characterizing the tick carboxypeptidase inhibitor: molecular basis for its two-domain nature. *J Biol Chem* 281: 22906–22916, 2006.

6. Arolas JL, Castillo V, Bronsoms S, Aviles FX, and Ventura S. Designing out disulfide bonds of leech carboxypeptidase inhibitor: implications for its folding, stability and function. *J Mol Biol* 392: 529–546, 2009.

7. Arolas JL, D'Silva L, Popowicz GM, Aviles FX, Holak TA, and Ventura S. NMR structural characterization and computational predictions of the major intermediate in oxidative folding of leech carboxypeptidase inhibitor. *Structure* 13: 1193–1202, 2005.

8. Arolas JL, Lorenzo J, Rovira A, Castella J, Aviles FX, and Sommerhoff CP. A carboxypeptidase inhibitor from the tick *Rhipicephalus bursa*: isolation, cDNA cloning, recombinant expression, and characterization. *J Biol Chem* 280: 3441–3448, 2005.

9. Arolas JL, Lorenzo J, Rovira A, Vendrell J, Aviles FX, and Ventura S. Secondary binding site of the potato carboxypeptidase inhibitor. Contribution to its structure, folding, and biological properties. *Biochemistry* 43: 7973–7982, 2004.

10. Arolas JL, Pantoja-Uceda D, Ventura S, Blanco FJ, and Aviles FX. The NMR structures of the major intermediates of the two-domain tick carboxypeptidase inhibitor reveal symmetry in its folding and unfolding pathways. *J Biol Chem* 283: 27110–27120, 2008.

11. Arolas JL, Popowicz GM, Bronsoms S, Aviles FX, Huber R, Holak TA, and Ventura S. Study of a major intermediate in the oxidative folding of leech carboxypeptidase inhibitor: contribution of the fourth disulfide bond. *J Mol Biol* 352: 961–975, 2005.

12. Arolas JL, Popowicz GM, Lorenzo J, Sommerhoff CP, Huber R, Aviles FX, and Holak TA. The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode. *J Mol Biol* 350: 489–498, 2005.

13. Arolas JL, Sanglas L, Lorenzo J, Bronsoms S, and Aviles FX. Insights into the two-domain architecture of the metallocarboxypeptidase inhibitor from the *Ascaris* parasite inferred from the mechanism of its oxidative folding. *Biochemistry* 48: 8225–8232, 2009.

14. Auerswald EA, Morenweiser R, Sommerhoff CP, Piechottka GP, Eckerskorn C, Gurtler LG, and Fritz H. Recombinant leech-derived tryptase inhibitor: construction, production, protein chemical characterization and inhibition of HIV-1 replication. *Biol Chem Hoppe Seyler* 375: 695–703, 1994.

15. Batey S, Scott KA, and Clarke J. Complex folding kinetics of a multidomain protein. *Biophys J* 90: 2120–2130, 2006.

16. Berndt KD, Guntert P, Orbons LP, and Wuthrich K. 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, 1992.

17. Bode W and Huber R. Structural basis of the endoproteinase-protein inhibitor interaction. *Biochim Biophys Acta* 1477: 241–252, 2000.

18. Bronsoms S, Villanueva J, Canals F, Querol E, and Aviles FX. Analysis of the effect of potato carboxypeptidase inhibitor pro-sequence on the folding of the mature protein. *Eur J Biochem* 270: 3641–3650, 2003.

19. Cemazar M, Daly NL, Haggblad S, Lo KP, Yulyaningsih E, and Craik DJ. Knots in rings. The circular knotted protein *Momordica cochinchinensis* trypsin inhibitor-II folds via a stable two-disulfide intermediate. *J Biol Chem* 281: 8224–8232, 2006.

20. Cemazar M, Gruber CW, and Craik DJ. Oxidative folding of cyclic cystine knot proteins. *Antioxid Redox Signal* 10: 103–111, 2008.

21. Cemazar M, Joshi A, Daly NL, Mark AE, and Craik DJ. The structure of a two-disulfide intermediate assists in elucidating the oxidative folding pathway of a cyclic cystine knot protein. *Structure* 16: 842–851, 2008.

22. Craik DJ, Cemazar M, Wang CK, and Daly NL. The cyclotide family of circular miniproteins: nature's combinatorial peptide template. *Biopolymers* 84: 250–266, 2006.

23. Creighton TE. Renaturation of the reduced bovine pancreatic trypsin inhibitor. *J Mol Biol* 87: 563–577, 1974.

24. Creighton TE. Electrophoretic analysis of the unfolding of proteins by urea. *J Mol Biol* 129: 235–264, 1979.

25. Creighton TE. Toward a better understanding of protein folding pathways. *Proc Natl Acad Sci U S A* 85: 5082–5086, 1988.

26. Creighton TE. Protein folding coupled to disulfide bond formation. *Biol Chem* 378: 731–744, 1997.

27. Creighton TE, Bagley CJ, Cooper L, Darby NJ, Freedman RB, Kemmink J, and Sheikh A. On the biosynthesis of bovine pancreatic trypsin inhibitor (BPTI). Structure, processing, folding and disulphide bond formation of the precursor *in vitro* and in microsomes. *J Mol Biol* 232: 1176–1196, 1993.

28. Creighton TE, Darby NJ, and Kemmink J. The roles of partly folded intermediates in protein folding. *FASEB J* 10: 110–118, 1996.

29. Creighton TE and Goldenberg DP. Kinetic role of a metastable native-like two-disulphide species in the folding transition of bovine pancreatic trypsin inhibitor. *J Mol Biol* 179: 497–526, 1984.

30. Chang JY. The functional domain of hirudin, a thrombin-specific inhibitor. *FEBS Lett* 164: 307–313, 1983.

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

32. Chang JY. Evidence for the underlying cause of diversity of the disulfide folding pathway. *Biochemistry* 43: 4522–4529, 2004.

33. Chang JY. Diversity of folding pathways and folding models of disulfide proteins. *Antioxid Redox Signal* 10: 171–177, 2008.

34. Chang JY, Canals F, Schindler P, Querol E, and Aviles FX. The disulfide folding pathway of potato carboxypeptidase inhibitor. *J Biol Chem* 269: 22087–22094, 1994.

35. Chang JY and Li L. Divergent folding pathways of two homologous proteins, BPTI and tick anticoagulant peptide: compartmentalization of folding intermediates and identification of kinetic traps. *Arch Biochem Biophys* 437: 85–95, 2005.

36. Chang JY, Schindler P, and Chatrenet B. The disulfide structures of scrambled hirudins. *J Biol Chem* 270: 11992–11997, 1995.

37. Chatrenet B and Chang JY. The folding of hirudin adopts a mechanism of trial and error. *J Biol Chem* 267: 3038–3043, 1992.

38. Chatrenet B and Chang JY. The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. *J Biol Chem* 268: 20988–20996, 1993.

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

40. Darby NJ and Creighton TE. Dissecting the disulphide-coupled folding pathway of bovine pancreatic trypsin inhibitor. Forming the first disulphide bonds in analogues of the reduced protein. *J Mol Biol* 232: 873–896, 1993.

41. Di Marco S and Priestle JP. Structure of the complex of leech-derived tryptase inhibitor (LDTI) with trypsin and modeling of the LDTI-tryptase system. *Structure* 5: 1465–1474, 1997.

42. Felizmenio-Quimio ME, Daly NL, and Craik DJ. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J Biol Chem* 276: 22875–22882, 2001.

43. Folkers PJ, Clore GM, Driscoll PC, Dodt J, Kohler S, and Gronenborn AM. Solution structure of recombinant hirudin and the Lys-47—-Glu mutant: a nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing study. *Biochemistry* 28: 2601–2617, 1989.

44. Goldenberg DP and Zhang JX. Small effects of amino acid replacements on the reduced and unfolded state of pancreatic trypsin inhibitor. *Proteins* 15: 322–329, 1993.

45. Gonzalez C, Neira JL, Ventura S, Bronsoms S, Rico M, and Aviles FX. Structure and dynamics of the potato carboxypeptidase inhibitor by 1H and 15N NMR. *Proteins* 50: 410–422, 2003.

46. Gottfried DS and Haas E. Nonlocal interactions stabilize compact folding intermediates in reduced unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* 31: 12353–12362, 1992.

47. Grutter MG, Fendrich G, Huber R, and Bode W. The 2.5 A X-ray crystal structure of the acid-stable proteinase inhibitor from human mucous secretions analysed in its complex with bovine alpha-chymotrypsin. *EMBO J* 7: 345–351, 1988.

48. Grutter MG, Priestle JP, Rahuel J, Grossenbacher H, Bode W, Hofsteenge J, and Stone SR. Crystal structure of the thrombin-hirudin complex: a novel mode of serine protease inhibition. *EMBO J* 9: 2361–2365, 1990.

49. Heitz A, Hernandez JF, Gagnon J, Hong TT, Pham TT, Nguyen TM, Le-Nguyen D, and Chiche L. Solution structure of the squash trypsin inhibitor MCoTI-II. A new family for cyclic knottins. *Biochemistry* 40: 7973–7983, 2001.

50. Housset D, Kim KS, Fuchs J, Woodward C, and Wlodawer A. Crystal structure of a Y35G mutant of bovine pancreatic trypsin inhibitor. *J Mol Biol* 220: 757–770, 1991.

51. Huber R, Kukla D, Bode W, Schwager P, Bartels K, Deisenhofer J, and Steigemann W. Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II. Crystallographic refinement at 1.9 A resolution. *J Mol Biol* 89: 73–101, 1974.

52. Kibria FM and Lees WJ. Balancing conformational and oxidative kinetic traps during the folding of bovine pancreatic trypsin inhibitor (BPTI) with glutathione and glutathione disulfide. *J Am Chem Soc* 130: 796–797, 2008.

53. Lim-Wilby MS, Hallenga K, de Maeyer M, Lasters I, Vlasuk GP, and Brunck TK. NMR structure determination of tick anticoagulant peptide (TAP). *Protein Sci* 4: 178–186, 1995.

54. Lin CC and Chang JY. Pathway of oxidative folding of secretory leucocyte protease inhibitor: an 8-disulfide protein exhibits a unique mechanism of folding. *Biochemistry* 45: 6231–6240, 2006.

55. Mamathambika BS and Bardwell JC. Disulfide-linked protein folding pathways. *Annu Rev Cell Dev Biol* 24: 211–235, 2008.

56. Muhlhahn P, Czisch M, Morenweiser R, Habermann B, Engh RA, Sommerhoff CP, Auerswald EA, and Holak TA. Structure of leech derived tryptase inhibitor (LDTI-C) in solution. *FEBS Lett* 355: 290–296, 1994.

57. Narayan M, Welker E, Wedemeyer WJ, and Scheraga HA. Oxidative folding of proteins. *Acc Chem Res* 33: 805–812, 2000.

58. Otlewski J, Jelen F, Zakrzewska M, and Oleksy A. The many faces of protease-protein inhibitor interaction. *EMBO J* 24: 1303–1310, 2005.

59. Pantoja-Uceda D, Arolas JL, Aviles FX, Santoro J, Ventura S, and Sommerhoff CP. Deciphering the structural basis that guides the oxidative folding of leech-derived tryptase inhibitor. *J Biol Chem* 284: 35612–35620, 2009.

60. Pantoja-Uceda D, Arolas JL, Garcia P, Lopez-Hernandez E, Padro D, Aviles FX, and Blanco FJ. The NMR structure and dynamics of the two-domain tick carboxypeptidase inhibitor reveal flexibility in its free form and stiffness upon binding to human carboxypeptidase B. *Biochemistry* 47: 7066–7078, 2008.

61. Plaxco KW, Simons KT, and Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277: 985–994, 1998.

62. Rawlings ND, Morton FR, Kok CY, Kong J, and Barrett AJ. MEROPS: the peptidase database. *Nucleic Acids Res* 36: D320–D325, 2008.

63. Rawlings ND, Tolle DP, and Barrett AJ. Evolutionary families of peptidase inhibitors. *Biochem J* 378: 705–716, 2004.

64. Rees DC and Lipscomb WN. Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 A resolution. *J Mol Biol* 160: 475–498, 1982.

65. Reverter D, Fernandez-Catalan C, Baumgartner R, Pfander R, Huber R, Bode W, Vendrell J, Holak TA, and Aviles FX. Structure of a novel leech carboxypeptidase inhibitor determined free in solution and in complex with human carboxypeptidase A2. *Nat Struct Biol* 7: 322–328, 2000.

66. Reverter D, Vendrell J, Canals F, Horstmann J, Aviles FX, Fritz H, and Sommerhoff CP. A carboxypeptidase inhibitor from the medical leech *Hirudo medicinalis*. Isolation, sequence analysis, cDNA cloning, recombinant expression, and characterization. *J Biol Chem* 273: 32927–32933, 1998.

67. Ryan CA, Hass GM, and Kuhn RW. Purification and properties of a carboxypeptidase inhibitor from potatoes. *J Biol Chem* 249: 5495–5499, 1974.

68. Rydel TJ, Ravichandran KG, Tulinsky A, Bode W, Huber R, Roitsch C, and Fenton JW, 2nd. The structure of a complex of recombinant hirudin and human alpha-thrombin. *Science* 249: 277–280, 1990.

69. Salamanca S, Li L, Vendrell J, Aviles FX, and Chang JY. Major kinetic traps for the oxidative folding of leech carboxypeptidase inhibitor. *Biochemistry* 42: 6754–6761, 2003.

70. Sanglas L, Aviles FX, Huber R, Gomis-Ruth FX, and Arolas JL. Mammalian metallopeptidase inhibition at the defense barrier of *Ascaris* parasite. *Proc Natl Acad Sci U S A* 106: 1743–1747, 2009.

71. Sommerhoff CP and Schaschke N. Mast cell tryptase beta as a target in allergic inflammation: an evolving story. *Curr Pharm Des* 13: 313–332, 2007.

72. Sommerhoff CP, Sollner C, Mentele R, Piechottka GP, Auerswald EA, and Fritz H. A Kazal-type inhibitor of human mast cell tryptase: isolation from the medical leech *Hirudo medicinalis*, characterization, and sequence analysis. *Biol Chem Hoppe Seyler* 375: 685–694, 1994.

73. Stubbs MT, Morenweiser R, Sturzebecher J, Bauer M, Bode W, Huber R, Piechottka GP, Matschiner G, Sommerhoff CP, Fritz H, and Auerswald EA. The three-dimensional structure of recombinant leech-derived tryptase inhibitor in complex with trypsin. Implications for the structure of human mast cell tryptase and its inhibition. *J Biol Chem* 272: 19931–19937, 1997.

74. Thompson RC and Ohlsson K. Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor, a potent inhibitor of leukocyte elastase. *Proc Natl Acad Sci U S A* 83: 6692–6696, 1986.

75. van Mierlo CP, Darby NJ, Keeler J, Neuhaus D, and Creighton TE. Partially folded conformation of the (30–51) intermediate in the disulphide folding pathway of bovine pancreatic trypsin inhibitor. 1H and 15N resonance assignments and determination of backbone dynamics from 15N relaxation measurements. *J Mol Biol* 229: 1125–1146, 1993.

76. Venhudova G, Canals F, Querol E, and Aviles FX. Mutations in the N- and C-terminal tails of potato carboxypeptidase inhibitor influence its oxidative refolding process at the reshuffling stage. *J Biol Chem* 276: 11683–11690, 2001.

77. Waxman L, Smith DE, Arcuri KE, and Vlasuk GP. Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. *Science* 248: 593–596, 1990.

78. Weissman JS and Kim PS. Reexamination of the folding of BPTI: predominance of native intermediates. *Science* 253: 1386–1393, 1991.

79. Weissman JS and Kim PS. Kinetic role of nonnative species in the folding of bovine pancreatic trypsin inhibitor. *Proc Natl Acad Sci U S A* 89: 9900–9904, 1992.

80. Weissman JS and Kim PS. The pro region of BPTI facilitates folding. *Cell* 71: 841–851, 1992.

81. Welker E, Narayan M, Wedemeyer WJ, and Scheraga HA. Structural determinants of oxidative folding in proteins. *Proc Natl Acad Sci U S A* 98: 2312–2316, 2001.

Address correspondence to:
Dr. Salvador Ventura
Departament de Bioquímica i Biologia Molecular
Institut de Biotecnologia i Biomedicina
Universitat Autònoma de Barcelona
E-08193 Bellaterra (Barcelona)
Spain

E-mail: salvador.ventura@uab.es

Abbreviations Used

ACI = *Ascaris* carboxypeptidase inhibitor
BPTI = bovine pancreatic trypsin inhibitor
CS = connecting segment
CTD = C-terminal domain
LCI = leech carboxypeptidase inhibitor
LDTI = leech-derived tryptase inhibitor
MCoTI-II = *Momordica cochinchinensis* trypsin inhibitor II
N = native protein
NMR = nuclear magnetic resonance
NTD = N-terminal domain
PCI = potato carboxypeptidase inhibitor
R = fully reduced and unfolded protein
RNase A = ribonuclease A
SLPI = secretory leukocyte protease inhibitor
TAP = tick anticoagulant peptide
TCI = tick carboxypeptidase inhibitor
XS = molecules with X number of disulfide bonds