Mutations in the N- and C-terminal Tails of Potato Carboxypeptidase Inhibitor Influence Its Oxidative Refolding Process at the Reshuffling Stage*

Gabriela Venhudová‡, Francesc Canals, Enrique Querol, and Francesc X. Aviles§

From the Institut de Biologia Fonamental and Departament de Bioquimica i Biologia Molecular, Universitat Autonoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

A comparative study of the oxidative refolding for nine selected potato carboxypeptidase inhibitor (PCI) mutants was carried out using the disulfide quenching approach. The mutations were performed at the N- and C-terminal tails of PCI outside its disulfide stabilized central core. The differences between the refolding of wild type and mutant proteins were observed in the second phase of the reshuffling process, the reshuffling of disulfide bridges, although the first phase, nonspecific packing, was not greatly affected by the mutations. Point mutations at the C-tail or deletion of up to three C-terminal residues of PCI resulted in a lower efficiency of the reshuffling process. In the case of the mutants lacking five N-terminal or four or five C-terminal residues, no “native-like” form was observed after the refolding process. On the other hand, the double mutant G35P/P36G did not attain a native-like form either, although one slightly more stable species was observed after being submitted to refolding. The disulfide pairing of this species is different from that of the wtPCI native form. The differences between the refolding process of wild type and mutant forms are interpreted in the light of the new view of protein folding. The results of the present study support the hypothesis that the refolding of this small disulfide-rich protein, and others, is driven by noncovalent interactions at the reshuffling stage. It is also shown that the interactions established between the N- and C-tail residues and the core of PCI are important for the proper refolding of the protein.

Small proteins that are rich in disulfide bridges are good models for studying the folding process using the disulfide quenching method, where the folding intermediates that form during their oxidative refolding are trapped at low pH, separated on HPLC, and analyzed (1, 2). Among the first and more thoroughly studied proteins using this technique was bovine pancreatic trypsin inhibitor (BPTI). Initially, mixed intermediates with native and non-native disulfide bridges were detected along the folding process (3). In later studies of BPTI folding (4), all of the well populated intermediates were found to contain only native disulfide bonds. Such results would support the idea of a single pathway for BPTI folding. Recently, it has been shown that BPTI also unfolds through a unique mechanism (5).

In contrast, the studies of hirudin (6), potato carboxypeptidase inhibitor (PCI) (7), epidermal growth factor (EGF) (8), and tick anticoagulant peptide (9) showed a high heterogeneity of folding intermediates containing both native and non-native disulfide bonds. In these proteins, the folding of the fully reduced protein to the native form occurs in a first phase as a flow of equilibrated 1-disulfide intermediates through equilibrated 2-disulfide intermediates to reach the equilibrated 3-disulfide scrambled species. In the second phase, the rate-limiting step, the scrambled species are reshuffled to finally form the native species. Moreover, the folding process seems to follow multiple parallel pathways. This mechanism would be consistent with the new view of protein folding processes describing folding funnels and energy landscapes (10, 11), i.e. parallel folding events, rather than the classical model of sequential folding via unique routes.

Further studies to characterize in depth the folding/unfolding process of the above mentioned or other disulfide rich proteins, based on such disulfide-quenching method, could help to clarify why and how multiple pathways take place not only in those particular proteins but also in other protein types, allowing comparisons with those lacking disulfides. With this purpose in mind, we have concentrated our attention on PCI.

PCI is a 39-residue globular protein that competitively inhibits several metallo carboxypeptidases with a $K_I$ in the nanomolar range (12). Its three-dimensional structure is known in aqueous solution (13) and in crystal state in complex with carboxypeptidase A (CPA) (14). The 27-residue core, reinforced by three disulfide bridges, forms a T-knot scaffold, also found in other proteins with different functions such as serine protease inhibitors from the squash family, the $\omega$-conotoxins from snake venoms, or the EGF superfamily of growth factors (15). PCI has 7 amino acid residues at the N-tail of the hydrophobic central core and a 5-residue C-tail, which is the primary binding site of PCI to carboxypeptidase A. The energetic contribution of each residue of the C-tail in the PCI-CPA complex has been evaluated by site-directed mutagenesis studies (16, 17). The biological function of PCI is probably the inhibition of insect digestive metallo carboxypeptidases as part of the defense system of the potato plant against insect attack (12). Recently, we have shown that PCI possesses interesting anti-tumoral properties that are related to its structural similarity to EGF and other related growth factors (18), all of them having the so-called
Influence of Mutations at the N- and C-tails on PCI Folding

T-knot topology (19).

Previous studies on the effect of denaturants on the folding pathway of PCI (7) suggested that specific noncovalent interactions mainly direct the reshufffing stage of the process. From the results of oxidative folding and reductive unfolding of hirudin, the intertwining dependence of the disulfide bonds and noncovalent forces was also concluded (20). Thus, in the proposed model for the folding of these small disulfide-rich proteins (6–9), until the scrambled species are formed the noncovalent interactions would not direct the folding, whereas the reshuffling step would be driven by noncovalent forces.

We have also recently characterized the unfolding pathway of PCI (21), showing that the bead-form scrambled isomer (presumably the most extended) is a significant intermediate in the process and that, under physiological conditions, scrambled isomers of PCI exist in equilibrium with the native form.

In the course of preparing recombinant PCI mutants constructed for functional studies, we found that some of the mutants were obtained in very low yield when expressed in Escherichia coli, even when the sequence changes did not affect the globular core of the protein. These observations prompted us to perform a comparative study of the oxidative refolding of selected mutants. The results of the study presented here show that the residues in the N- and C-tails of PCI are important for its folding, by affecting the reshuffling stage of the process to a different extent. The knowledge on the involvement of the noncovalent interactions between PCI tails and central core in its folding can be of importance in the future redesign of this small protein for pharmaceutical or other biotechnological purposes.

EXPERIMENTAL PROCEDURES

Production and Purification of the Mutant PCI Forms—The Y37G, P36G, G35/P36G, ΔVal35-Gly39, ΔGly35-Gly39, ΔPro35-Gly39, and ΔGly35-Gly39 PCI mutant genes were constructed by site-directed mutagenesis of wild type PCI synthetic gene (22) cloned in PINIII-OmpA3 vector (23). The proteins were expressed in E. coli and purified from the culture medium by ion exchange fast protein liquid chromatography and reverse-phase HPLC. Details regarding the construction of the mutants vectors and purification of the expressed proteins have been published elsewhere (17).

ΔGly39PCI was obtained by the incubation of wtPCI with CPA at 1:1 molar ratio for 4 h followed by purification on a Vydac C-4 HPLC column. CPA is known to cleave the Gly39 residue of PCI rapidly after its N-terminal side of the cyanylated cysteines for one h at 37 °C in a solution of 1 M ammonium hydroxide containing 6 M GdnCl, and finally the remaining masses were reduced to 5 M tris(2-carboxyethyl)phosphine for 30 min at 37 °C. The reaction mixture was subjected to reverse-phase HPLC and the three cyanylated isomers resulting from reduction of a single disulfide bond, as identified by MALDI-TOF mass spectrometry, were isolated. They were subjected to cleavage of the peptide bonds on the N-terminal side of the cyanylated cysteines for one h at 37 °C in a solution of 1 M ammonium hydroxide containing 6 M GdnCl, and finally the remaining masses were reduced to 5 M tris(2-carboxyethyl)phosphine for 30 min at 37 °C. The resulting mixtures of peptides were analyzed by MALDI-TOF mass spectrometry, and the disulfide bonds were deduced from the obtained mass maps.

Protease Digestions—Selected PCI scrambled species, scrambled A form of wtPCI (defined in Ref. 7), the G35/P36G PCI most stable form, or scrambled species of mutants ΔPro35–Gly39 and ΔGly35–Gly39 were isolated by HPLC on a Vydac C-18 column from the corresponding mixture of refolding intermediates, obtained as detailed under “Refolding Experiments.” 1 μg of the purified scrambled species, or the native wtPCI form as a control, were incubated for 1 h with 1 μg of elastase in 0.1 M Tris, pH 8.8, buffer or with 0.05 μg of thermolysin in 50 mM N-ethylmorpholine, pH 6.5, buffer. The reaction digestes were analyzed by MALDI-TOF mass spectrometry.

CPA Inhibitory Assays—The inhibitory activities of the different mutant PCI forms were assayed according to Hass and Ryan (12). Benzoyl-glycyl-t-phenylalanine was used as a substrate at 0.1 M concentration, and the enzyme (bovine carboxypeptidase A) was at 42.5 nM. The assays were performed at pH 7.5.

RESULTS

Refolding Behavior—The oxidative refolding in vitro of different N- and C-tail mutant forms of PCI was studied by reverse-phase HPLC analysis of the acid-trapped disulfide intermediates present in samples collected at different times during the refolding processes, these processes being followed either in the absence or presence of an external thiol added (cysteine) (see “Experimental Procedures”). The assays were always performed in parallel to a control refolding experiment of wtPCI under the same conditions. The refolding conditions used in this study (i.e. protein concentration, buffer conditions, the type of thiol reagent, and its concentration) were selected so that the qualitative or quantitative differences in the refolding process between the wild type and mutant PCI forms could be distinguished (number and type of folding intermediates, kinetics of the process) within a few hours’ time span, being slightly different from those of a previous paper (7).

The typical results observed for the refolding of wtPCI under the selected conditions are shown in Fig. 1. The protein was initially reduced and denatured in the presence of excess dithiothreitol and 5 M GdnCl and then quickly brought to the refolding buffer by gel filtration on a small PD-10 column. The HPLC profile of a sample collected immediately after the gel filtration (t = 0, Fig. 1) mainly contained a mixture of 3-disulfide paired scrambled species, as identified according to their chromatographic behavior, well characterized in previous studies on the disulfide refolding of PCI (7). Under the above conditions, it was observed (Fig. 1) that the first stage of nonspecific packing of wtPCI to reach 3-disulfide paired species was
Influence of Mutations at the N- and C-tails on PCI Folding

very quick and that the second stage of disulfide reshuffling of scrambled intermediates was rate-limiting. The evolution of these intermediates thereafter was very slow in the buffer without added cysteine (Fig. 1, left), probably being catalyzed only by the small amount of protein free thiols still present and leading to a final trapped mixture with less than 10% of the native species. In contrast, when 1 mM cysteine was added, a much faster reshuffling of the scrambled species led to a mixture in which more than 80% of the protein had the native disulfide pairing (Fig. 1, right).

When the study was performed on the N- and C-tail mutant forms of PCI, all of the mutants analyzed showed a first stage of refolding as fast as wtPCI under the same conditions. In the chromatograms obtained immediately after passing the samples through the PD-10 column, the main peaks corresponded to 3-disulfide scrambled species (results not shown), the subsequent evolution being very small when free cysteine was lacking in the buffer (see Figs. 2, 3, and 4 for the patterns observed after 3 h).

However, significant differences between mutants were observed in the second step of the refolding process, i.e. in the reshuffling step, when 1 mM cysteine was present. According to the evolution of their chromatographic pattern, the mutants were classified in three groups. Mutants of a first group (Group I) resembled wtPCI in their behavior (Fig. 2). These included point mutations in the C-terminal tail (P36G and Y37G) and consecutive deletions of up to three residues of the C-tail. It seems clear that the isomers marked “N” are what could be called native-like forms of each mutant, i.e. they possess the same disulfide bridging as wtPCI native form and presumably have a compact stable structure, as supported by the following findings: they all maintained CPA inhibitory activity, their retention time in the HPLC column was almost identical to that of the wtPCI form, and their relative amounts in the refolded mixtures increased in the course of refolding, so they seemed to be the most stable forms. As in the case of wtPCI, there was a remarkable difference between the mixtures with and without cysteine added. Little or no native-like form could be observed in the samples refolded in the buffer alone, whereas a significant amount of this form was seen in the mixtures with cysteine. The percentages of native-like forms in ∆Gly39, ∆Val38–Gly39, and Y37G mutants are close to wtPCI (more than 55%), whereas the trapped mixtures of P36G and ∆Tyr37–Gly39 mutants after 3 h contained only about 30% of such form. The percentages of native-like forms obtained for each protein under these conditions indicate a different kinetic efficiency for their reshuffling process. If a higher concentration of reducing agent is present in the refolding buffer, more than 90% of native-like species is formed for each of the mutants from this Group I (data not shown).

A second group was represented by mutants with a very different refolding behavior and which are not able to reach a native-like conformation (Group II). Their chromatograms after 3 h of refolding are shown in Fig. 3. Their mutations represent the largest changes made in the sequence of PCI tails (although they do not affect the amino acid residues of the core stabilized by the disulfide bonds), being the result of the deletion of four or five residues in the C-terminal tail or the deletion of five residues in the N-terminal tail of PCI. In this Group II no differences in the chromatographic profile could be observed between the 3-h trapped mixtures with and without cysteine. In none of these mixtures was a significant peak that could be considered a native-like form seen.

It is worth mentioning that the ∆Glu1–Asp5 mutant was initially prepared by direct mild acid hydrolysis from native wtPCI, leading to a protein that maintained the disulfide pairing and CPA inhibitory activity. Therefore, the HPLC elution time of such an N-terminal truncated form can be considered that which corresponds to the actual native-like species. However, when this truncated protein was reduced and allowed to refold, none of the resulting “refolded” chromatographic forms showed either a peak at the former elution time (Fig. 3) or any inhibitory activity. Thus, this protein variant lacking the N-terminal tail, once reduced, is not able to refold to its original native-like conformation.

It is not obvious whether some of the disulfide isomers of the mutants lacking four or five amino acids in the C-tail of PCI have the same disulfide pairing as the native PCI. As these mutants do not contain the residues that constitute the primary binding site to CPA, the activity assay does not give information about the conformation of the proteins. Besides, the HPLC retention times of the “correctly folded” forms of these mutants are not known. However, as we observe no evolution of the “scrambled” forms to one HPLC peak during refolding in the presence of cysteine, it seems likely that these mutations cause the inability of PCI to form a native-like stable conformation. The species present in the final refolding mixtures of these two mutants were isolated by HPLC and assayed for susceptibility to protease digestion. All of the species were found to be extensively degraded by elastase or thermolysin under conditions in which the native form of wtPCI is resistant to digestion. This behavior is also consistent with these species...
not having a native-like compact conformation.

For all three mutants of Group II attempts to obtain the correctly folded proteins in the presence of higher concentrations of reducing agent were carried out. The formation of a native-like form was not detected even in the presence of 4 mM cysteine, conditions under which wtPCI forms more than 90% of native species in less than 1 h (data not shown).

The refolding process and characteristics of the G35P/P36G double mutant are different, and therefore it was included in a third group (Group III). The chromatograms of its 3-h trapped refolding mixtures are shown in Fig. 4. A clear difference is observed between the chromatograms of the trapped mixtures in the absence or presence of added cysteine in a way similar to that observed for wtPCI or the mutants of Group I. In this case, the relative amount of one species (peak labeled S in Fig. 4) also increased when cysteine was added to the refolding buffer. Although the overall amino acid content of this double mutant is the same as that of wtPCI, the chromatographic behavior on the reverse-phase HPLC of this species (S) is very different from that of the wtPCI native form and resembles that of the scrambled species A of wtPCI (7), suggesting that such a species S can have a disulfide bond pairing different from the native form of wtPCI. To confirm this possibility, the disulfide bond pairing of such a most stable form of G35P/P36G (peak S) was determined by a partial reduction-mass spectrometry method (see “Experimental Procedures”). Disulfide bridges between cysteines I–V, II–III, and IV–VI (see Fig. 4) were found, confirming a different pairing from that of the native wtPCI.
Interestingly, the same disulfide pairing has also been determined for the scrambled species A of wtPCI.

Unlike results observed for the mutants of Group I, when refolding was attempted under more favorable conditions (higher cysteine concentrations) or longer times, no further evolution of the mixture of intermediates was observed for the G35P/P36G mutant. This seems to indicate that the species S, which increases in the refolding mixture with cysteine, is not actually native-like, in the sense that its thermodynamic stability is not much higher than that of the other coexisting 3-disulfide species. Susceptibility to elastase hydrolysis was used as an additional tool to check whether this species S has a compact, stable structure, the hydrolysis being followed in parallel with wtPCI native and scrambled A forms. After a 1.5-h incubation, only Gly39 was cut from the wtPCI native form, whereas the scrambled A and S species were degraded to fragments of less than 1000 Da, as observed by MALDI-TOF mass spectrometry.

In conclusion, the G35P/P36G mutant seems not to be able to refold into a native, compact structure, just as the mutants of Group II. However, unlike what was observed for the mutants of Group I or II, in this case the formation of one of the scrambled species, S, has a kinetic barrier greater than the rest of the scrambled species, and, accordingly, its proportion in the mixture of trapped intermediates depends on the addition of thiols to the refolding buffer (see "Discussion").

**CD Spectroscopy**—Although the wtPCI native form does not have regular secondary structures, except for a short 5-residue helix, we found that it presents features that have made the far-UV CD spectroscopy helpful as a tool to indicate the folding state of the mutants. The far-UV CD spectrum of wtPCI showed a characteristic positive ellipticity band at 228 nm and a minimum at 204 nm. The positive band at 228 nm was not found in some PCI mutants (Fig. 5). For instance, the so-called native-like refolded form of the Y37G mutant, lacking the Tyr37 residue, which presents a strong CPA inhibitory activity (K_i, 3.5 nM) and hence probably has a three-dimensional structure very similar to the wtPCI native form, does not display the characteristic positive band on its CD spectra. Therefore, Tyr37 side chain seems to be related to the presence of this band. It is

---

2 S. Pavia, unpublished observations.
known that aromatic residues may cause optically active bands in the far-UV region, their environment affecting the position of CD bands of proteins (27). When the CD spectrum of wtPCI was recorded at different pH values (pH 2, 9, and 11) it was found that the maximum at 228 nm disappears at pH 11 (Fig. 5a), probably when the tyrosine hydroxyl groups are ionized, in agreement with the potential involvement of Tyr37 in correctly folded PCI. Indeed, the scrambled forms of wtPCI, which possess all the residues of the native form but are folded in different ways, do not display either the maximum at 228 nm (not shown). Hence, the environment of Tyr37 in correctly folded PCI is probably responsible for the characteristic maximum at 228 nm, and so the PCI forms lacking the Tyr37 residue or having a different three-dimensional structure would not display the characteristic CD spectrum.

Indeed, all of the main protein species found in the HPLC peaks of the final refolding mixtures for the mutants from the first folding group (that is, peaks N), except for Y37G and ΔTyr37–Gly39 lacking the Tyr37, show the maximum at 228 nm on their CD spectra (Fig. 5b). In contrast, none of these bands in the HPLC peaks of the final refolded mixtures of the mutants from the second group displayed the 228 nm maximum (data not shown). In the case of ΔPro36–Gly39 and ΔGly35–Gly39 mutants, which have lost most or all of the C-tail and do not contain the Tyr37 residue, the maximum at 228 nm would not be expected even if the correctly folded protein (at the globular core) was present. Also, in the most stable form of the G35P/ P36G mutant (species S), the Tyr37 residue is present but, as noted above, the conformation of this form is different from that of the native wtPCI, and, accordingly, the positive band at 228 nm is not present (Fig. 5c).

**DISCUSSION**

All of the mutations performed in this study alter the PCI sequence out of its central globular core. The substitution of one residue at the C-tail at its boundary with the core (in mutants P36G and Y37G) and deletion of up to three residues from the C terminus significantly affected the PCI behavior during refolding but did not prevent the attainment of a native-like state. As in the case of wtPCI, in the presence of cysteine, the equilibrated mixture of scrambled forms in such mutants decreased to give rise to the native-like species (although the molar proportion between the scrambled forms is maintained along the refolding process). Yet, the yield of a native-like form in the final mixture of species was lowered in such mutants in relation to wtPCI. All of these mutations therefore affect the refolding process by diminishing the efficiency of the reshuffling step.

The mutations that have the strongest influence on the folding of PCI are the deletion of five residues at the N terminus or four or five residues at the C terminus, giving rise to a "trimmed" core with 2/5, 7/1, and 7/0 residues at the N/C tails, respectively. These mutants also quickly reached the equilibrium of scrambled species in the first stage of folding, but no reshuffling occurred in the second stage even in the presence of 1 mM cysteine. These results show that the reshuffling step is clearly influenced by the residues of the C- and N-terminal tails of PCI. It is worth mentioning that the first three residues at the PCI N terminus are disordered in the NMR structure (13), and interactions being observed between them and the PCI core. Only the fourth residue, alanine, forms a backbone hydrogen bond with the Cys4 residue, the one that establishes the boundary with the core, as observed in the crystal structure of the PCI-CPA complex (14). In addition, three main-chain/sidechain hydrogen bonds involving the N-terminal residues of PCI are found in the crystal structure of its complex with CPA: Ala1–Trp22, Ile7–Asp5, and Phe23–Asp5. None of these interactions is supposed to be important for the PCI inhibitory activity, because when five N-terminal residues are cut off from an already folded protein, the trimmed PCI is still fully active (25). However, the present results show that this N-tail could play a role in the PCI folding process.

As mentioned above, when up to three residues from the PCI C terminus are removed, the yield of the native-like form in the refolded mixture of species is lowered, in the conditions used in this study. If the fourth residue of the C terminus, Pro36, is also deleted, no formation of the native-like species is detected. Both in the NMR and x-ray structure of PCI, a hydrogen bond between the carbonyl oxygen of Pro36 (in the tail) and the amide nitrogen of Trp28 (in the core) is observed. Our results seem to indicate that this hydrogen bond, which could be formed by any residue in position 36, could participate significantly in directing the native disulfide pairing of PCI. When no residue is present in position 36, that is when the C-tail only has one residue (Gly35), no native-like form is seen in the course of refolding. Substitution of Pro36 by glycine in the wild type form (in P36G mutant) causes only the formation of a lower relative amount of the native-like form in the final mixture of species.

The results observed in the refolding of the G35P/P36G double mutant emphasize the participation of the Gly35 and Pro36 residues in driving the native PCI folding process. The substitution of Gly35 by proline probably influences the orientation of the PCI C-tail and thus prevents the formation of the hydrogen bond between Trp28 (in the core) and the new residue at position 36 in the double mutant, a glycine. It is important to mention that Gly35 in wtPCI establishes two hydrogen bonds with Ala26, Ala26(N–Gly35(O) and Gly35(N–Ala26(O). These two backbone hydrogen bonds are observed both in the NMR structure of wtPCI in solution and in the crystal structure of its complex with CPA. The substitution of Gly35 by proline causes suppression of at least one of these hydrogen bonds, as proline lacks the amide hydrogen. In the course of refolding of the G35P/P36G double mutant native-like species is formed. On the other hand, formation of a slightly more stable scrambled species that differs from the native form of wtPCI is observed. The importance of Gly35 and its two hydrogen bonds with Ala26 is reflected also in the fact that Gly35 and Ala26 are two of the eight PCI residues with less than 10 Å² of accessible surface area (14). The other six residues buried in the wtPCI core are Ala3 and five cysteines forming the disulfide bridges.

As previously reported, the cystine knot (or T-knot) structural motif gives rise to compact structures found in several proteins with different functions, such as EGF-like molecules, ω-toxins, and proteinase inhibitors from plants (15, 19). Some of these T-knots structures are included as domains within multidomain proteins, but many of them, such as the serine protease inhibitors from the squash plants, PCI, or the ω-conotoxins from snake venom are found in nature as isolated monodomain small disulfide-rich proteins. The question of whether these proteins need the tails for correct refolding, as does PCI, is an interesting issue. The trypsin squash inhibitors have a very similar structural topology to PCI and the same disulfide bridge pattern (28). The sequence alignment of a large number of the squash inhibitors shows that the cysteine-reinforced core contains 25 amino acids (27 in PCI); there are zero to five amino acids at the N terminus and, in most cases, only one glycine residue at the C terminus, after the last cysteine of the T-knot core. This glycine is conserved among all of the 41 known members of the squash inhibitor family (28). In the crystal structure of the Cucurbita maxima trypsin inhibitor-I (CMTI-I), in complex with trypsin, such a glycine forms two internal backbone hydrogen bonds with Val21, analogously to...
PCI. Refolding of the chemically synthesized proteins is a common procedure used to obtain squash inhibitors and their variants (29). Some of these variants lacked any amino acid out of the core enclosed by the disulfide bridges. However, the yield of correctly folded forms during in vitro refolding ranged between 5 and 10%, both for wild type and mutant proteins, independently of the presence of the tail residues. These low yields may be explained by the possible requirement of pro-sequences for the correct folding of the trypsin squash inhibitors.

On the other hand, the ω-conotoxins have all of their 23–26 amino acids enclosed by three disulfide bridges. Refolding studies in vitro revealed that under optimal conditions the yield of the correctly folded proteins is between 16 and 50%. The presence of an additional glycine residue at the C terminus of the correctly folded proteins is between 16 and 50%. The barriers between the intermediates of each group are small, so the scrambled species, which would represent kinetic traps, which can be overcome by the addition of reducing agents to allow their reshuffling to the thermodynamically favored native state.

The results presented here indicate that the energy landscape has some differences for the various mutants. The mutants of Group I have an energy barrier to the native state that is higher than that of wtPCI, a fact that results in the observed slower kinetics for the reshuffling step of these mutants or the need of higher concentrations of reducing agent to reach the native state efficiently. In the case of mutants of Group II, their inability to reach a native state can be explained in two ways. On one hand, the mutations introduced can cause a situation in which no conformation possesses a far higher thermodynamic stability than the rest, i.e., no truly native state exists. On the other hand, in the case that a hypothetical native state does exist, the "uphill step" necessary to leave the trap would be energetically too unfavorable, and therefore even at high concentrations of reducing agent the native state would not be obtained. In the case of the ΔGlu–Asp N terminus trimmed protein, the existence of a native form is known, because the protein, obtained by hydrolysis of wtPCI, presents all of the structural characteristics (CD, inhibitory activity, resistance to proteolysis) of the compactly folded form of wtPCI. Thus, at least for this protein, the inability to refold to the native state is because of kinetic rather than thermodynamic reasons.

The inability of these mutants to reach a native-like state could result from the formation of aggregates of the misfolded forms. However, the scrambled forms of these proteins show a behavior in reverse-phase HPLC very similar to wtPCI scrambled species, which would indicate a similar degree of exposure of the hydrophobic regions and consequently a similar tendency to aggregation as wtPCI. Moreover, when the refolding of these mutants was performed in more diluted solutions, no native form could be observed either (data not shown). Therefore it does not seem probable that aggregation plays a major role in the misfolding of these proteins, although it cannot be ruled out based on the present evidence.

For all proteins studied here, with the unique exception of the G35P/P36G mutant, the energy barriers between the different scrambled species seem to be much lower than those leading to the native form, and therefore the equilibrium proportions of the different forms are maintained throughout the folding process regardless of the addition of external thiols. In the case of G35P/P36G PCI, the scrambled S form presents a higher energy barrier, and its equilibrium proportion can only be attained in the presence of added thiol.

Given that in the last few years PCI has become a molecule with biomedical potentialities, having been suggested as a drug for blood fibrinolysis (36) and for antitumoral strategies (18) and given that its N- and C-tails seem to be involved in such potentialities, it is interesting to fully characterize the influence of these tails in the folding of the protein and, therefore, in its recombinant production and redesign. The same applies to other topologically related molecules, such as EGF, transforming growth factor-α, toxins, and defensins (19), several of them having clear biotechnological applications.

Acknowledgments—We are indebted to Dr. Miquel Pons (Universitat de Barcelona) for kindly helping us in the CD studies and to Dr. R.-Y.-Y. Chang (Institute of Molecular Medicine, Houston) for helpful discussions. The support and advice of Drs. Cristina Marino and Silvana Pavia are also gratefully acknowledged.

REFERENCES

1. Creighton, T. E. (1986) Methods Enzymol. 131, 83–106
2. Creighton, T. E. (ed.) (1992) Protein Folding, pp. 301–351, EMBL, Heidelberg
Influence of Mutations at the N- and C-tails on PCI Folding

3. Creighton, T. E., and Goldenberg, D. P. (1984) J. Mol. Biol. 179, 497–526
4. Weissman, J. S., and Kim, P. S. (1991) Science 253, 1386–1393
5. Chang, J. Y., Li, L., and Bulcherc, A. (2005) J. Biol. Chem. 275, 8287–8299
6. Chatrenet, B., and Chang, J. Y. (1993) J. Biol. Chem. 268, 20988–20996
7. Chang, J. Y., Canals, F., Schindler, P., Querol, E., and Aviles, F. X. (1994) J. Biol. Chem. 269, 22867–22894
8. Chang, J. Y., Schindler, P., Rameser, U., and Lai, P. H. (1995) J. Biol. Chem. 270, 9207–9216
9. Chang, J. Y. (1996) Biochemistry 35, 11702–11709
10. Baldwin, R. L. (1994) Nature 369, 183–184
11. Dill, K. A., and Chan, H. S. (1997) Nat. Struct. Biol. 4, 10–19
12. Hass, G. M., and Ryan, C. A. (1981) Methods Enzymol. 80, 778–791
13. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) Biochemistry 26, 8012–8023
14. Rees, D. C., and Lipscomb, W. N. (1982) J. Mol. Biol. 160, 475–498
15. Lin, S. L., and Nussinov, R. (1995) Nat. Struct. Biol. 2, 835–837
16. Molina, M. A., Marino-Buslje, C., Oliva, B., Aviles, F. X., and Querol, E. (1994) J. Biol. Chem. 269, 21467–21472
17. Marino-Buslje, C., Vehnudova, G., Molina, M. A., Oliva, B., Jorba, X., Canals, F., Aviles, F. X., and Querol, E. (2000) Eur. J. Biochem. 267, 1503–1509
18. Blanco-Aparicio, C., Molina, M. A., Fernández-Salas, E., Fraizer, M. L., Mas, J. M., Querol, E., Aviles, F. X., and de Llorens, R. (1998) J. Biol. Chem. 273, 12370–12377
19. Mas, J. M., Aloy, P., Martí-Renom, M. A., Oliva, B., Blanco-Aparicio, C., Molina, M. A., de Llorens, R., Querol, E., and Aviles, F. X. (1998) J. Mol. Biol. 284, 541–548
20. Chang, J. Y. (1997) J. Biol. Chem. 272, 69–75
21. Chang, J. Y., Li, L., Canals, F., and Aviles, F. X. (2000) J. Biol. Chem. 275, 14205–14211
22. Molina, M. A., Aviles, F. X., and Querol, E. (1992) Gene 116, 129–138
23. Ghrayeb, J., Kimura, H., Takahara, M., Hsiung, H., Masui, Y., and Inouye, M. (1984) EMBO J. 3, 2437–2442
24. Hass, G. M., and Ryan, C. A. (1980) Biochem. Biophys. Res. Commun. 97, 1481–1486
25. Hass, G. M., Ako, H., Grahn, D. T., and Neurath, H. (1976) Biochemistry 15, 93–100
26. Wu, J., and Watson, J. T. (1997) Protein Sci. 6, 391–398
27. Adler, A. J., Greenfield, N. J., and Fasman, G. D. (1973) Methods Enzymol. 27, 675–735
28. Otlewski, J., Krowarsch, D. (1996) Acta Biochim. Pol. 43, 431–434
29. Rolka, K., Kupryszewski, G., Różycki, J., Ragnarsson, U., Zbyryt, T., and Otlewski, J. (1992) Biol. Chem. Hoppe-Seyler 373, 1055–1060
30. Price-Carter, M., Gray, W. R., and Goldenberg, D. P. (1996) Biochemistry 35, 15547–15557
31. Shinde, U., and Inouye, M. (1993) Trends Biochem. Sci. 18, 442–446
32. Villanueva, J., Canals, F., Prat, S., Ludevid, D., Querol, E., and Aviles, F. X. (1998) FEBS Lett. 440, 175–182
33. Martinneau, B., McBride, K. E., and Houck, C. M. (1991) Mol. Gen. Genet. 228, 281–296
34. Ling, M. H., Qi, H. Y., and Chi, C. W. (1993) J. Biol. Chem. 268, 810–814
35. de Prat, Gay, G., Ruiz-Sanz, J., Neira, J. L., Corrales, F. J., Otzen, D. E., Ladurner, A. G., and Fersht, A. R. (1995) J. Mol. Biol. 254, 968–979
36. Klement, P., Liao, P., and Bajzar, L. (1999) Blood 94, 2735–2743