

VI- Treball 3. Monitoring the protein folding by D/H exchange followed by MALDI-TOF MS

VI.A Introduction

The folding of the small disulfide-rich proteins has been broadly studied as a good approach to the general understanding of the protein folding problem (Chang et al; 2000). Although protein folding basically depends on the amino acid sequence (Anfinsen, 1973), the rules that govern the folding process still remain to be established and detailed (Fersht and Daggett, 2002). The process is characterized by the appearance and disappearance of a number of partially structured folding intermediates and the isolation and characterization of these intermediates contribute significantly to the understanding of the folding process. In the case of the small disulfide-rich proteins, their folding is substantially driven and constrained by the disulfide-bond formation. Therefore, the characterization of the number and pairing of disulfide bonds of the folding intermediates and their kinetics of appearance and disappearance allows the characterization of important aspects of their folding process.

The acid-trapping method followed by reverse-phase high pressure liquid chromatography (RP-HPLC) is a simple and quick approach to isolate and study disulfide folding intermediates. Typically, the folding reactions are performed at pH 7-8.5, and then time-course samples are quenched with acid or the sulfhydryl groups are blocked to stop disulfide formation (Konishi et al, 1982). Afterwards they can be analyzed by RP-HPLC and characterized by mass spectrometry (MS), amino acid analysis and disulfide analysis (Thannhauser et al, 1998). Using this methodology the folding and unfolding process of an important number of small disulfide-rich proteins have been elucidated (Weissman and Kim; Chatrenet & Chang, 1993; Chang et al, 1994; Chang, 1995) and some differences among their folding mechanisms have been observed. For hirudin, potato carboxypeptidase inhibitor (PCI) and tick anticoagulant peptide (TAP) the folding process is characterized by a high heterogeneity of 1 and 2-disulfide intermediates and by the presence of 3-disulfide scrambled species, and their reductive unfolding follows an all-or-none mechanism (Chang, 2000). For epidermal growth factor (EGF) and leech carboxypeptidase inhibitor (LCI) a heterogeneity of 1 and 2-disulfide species and of scrambled forms are detected along their folding pathway, but there is also a stable intermediate that in the case of EGF contains all native disulfide bonds except one. This stable intermediate accumulates during both protein folding and reductive unfolding and acts as a kinetic trap (Chang, 2000 and Salamanca et al., unpublished results).

The acid-trapping method is a useful approach to study the disulfide folding mechanism, but it poorly contributes to elucidate the conformational stability of the trapped folding intermediates. Several attempts of resolving three-dimensional structures of disulfide folding intermediates by NMR were unsuccessful and only one has been reported (Gehrmann et al, 1998). Although NMR is also capable of analyzing the conformational stability of folding intermediates through the amide proton/deuteron (H/D) exchange, the large amounts of these species required and sometimes, their

poor solubility at the required concentration for analysis, severely limited its use for folding studies. However, other approaches with capability for such kind of H/D analysis, but less demanding in terms of sample quantities required, may be very useful. We propose the use of matrix-assisted laser desorption/ionization- time of flight (MALDI-TOF) mass spectrometry (MS) for this goal.

MALDI-TOF mass analysis is a broadly used method for mass measurements of biological molecules (Beavis & Chait, 1990). Recently, MALDI-TOF MS has been used to follow H/D exchange for the study of protein-protein interaction (Mandell et al, 1998) and for the characterization of the stability and conformation of proteins (Villanueva et al, 2000; Ghaemmaghami et al, 2000). While H/D exchange coupled to MALDI-TOF MS has been used to obtain structural and conformational stability information, only H/D exchange coupled to electrospray ionization (ESI) has been used to study protein folding (Miranker et al, 1996 and Jeng et al, 1990). The homogeneity and stability of the native state of proteins, make them relatively simple to describe. Contrarily, the characterization of folding intermediates and unfolded states is a more difficult issue, due to the short-life and conformational heterogeneity of these molecules.

In the present work, we have applied a methodology based on the hydrogen exchange of proteins monitored by MALDI-TOF MS, previously reported by our group, to characterize the protein folding process of disulfide-rich proteins (Villanueva et al, 2000). Two small disulfide rich proteins, hirudin core domain and LCI have been chosen, since their folding mechanism have been previously studied and they represent two different folding models of disulfide-rich proteins (Chang, 1993; Salamanca et al, unpublished results). The comparison of the number of slow exchanging protons of different samples along the folding pathway provides an easy and fast way to follow the protein folding process. The picture here stated for hirudin and LCI folding is in good agreement with those previously described by the acid-trapping method and RP-HPLC. Additionally, D/H exchange experiments on purified folding intermediates were performed in order to characterize their conformational stability. The data provided evidence that the conformational stability correlates well with protein folding progress, but not necessarily with the number of disulfide bonds of the folding intermediates. Since we are also able to quantify the extent of the unfolding of each of the folding intermediates, we can measure the contribution of the disulfide bonds versus non-covalent forces in the folding process of the small disulfide-rich proteins.

VI.B Experimental Procedures

VI.B.1 Folding experiments

0.5 mg lyophilized aliquots of protein (hirudin or LCI) were dissolved in 0.5 mL guanidinium chloride (GndCl) 8M, dithiotreitol (DTT) 30 mM, Tris-Cl 0.5 M, pH 8.5. After 3 hours at 25°C the sample was acidified to pH 3.0, adding trifluoroacetic acid (TFA) to prevent disulfide formation,

and the reduced and denatured protein was filtered through a PD-10 column (Pharmacia) equilibrated with 20 mM glycine pH 3.0, to remove the denaturant and the reducer. The protein was collected in 1.2 mL sample, split in three aliquots and lyophilized.

Folding experiments in a protonated buffer. An aliquot of lyophilized protein collected from the PD-10 column was dissolved in 400 μ l of 0.1 M TrisCl, pH 8.5. At these conditions, the disulfide formation is favored and the protein starts its refolding process. In the case of LCI, the refolding buffer contained 1mM reduced glutathion (GSH) and 0.5 mM oxidized glutathion (GSSG), and in the case of hirudin contained 0.1 mM GSSG. Samples of all reaction mixtures were collected in a time-course manner and trapped by mixing with equal aliquots of water containing 1%TFA and analyzed by reverse phase-high pressure liquid chromatography (RP-HPLC) on a C4 0.46x25 cm 5 μ m column (Vydac). The chromatographic conditions used were: solvent A, water containing 0.1% TFA and solvent B, acetonitrile containing 0.1% TFA. The gradient was linear, 20%-40% solvent B in 40 minutes for LCI and 15%-35% solvent B in 40 minutes for hirudin.

Folding experiments in a deuterated buffer. For LCI, an aliquot of lyophilized protein collected from the PD-10 column was dissolved in 100 μ l of deuterated buffer (glycine, pD 3.0) and incubated for 1.5 hours at 90°C in order to exchange completely all labile protons, and afterwards it was maintained for 30 minutes at room temperature. The deuterated sample was lyophilized and dissolved in 400 μ l of the deuterated refolding buffer (TrisHCl 0.1 M, pH 8.5, containing GSSG 0.5 mM and GSH 1 mM). Samples of the refolding mixture were taken in a time-course manner, trapped by sample acidification with DCl 1 M and lyophilized.

For hirudin, an aliquot of lyophilized protein collected from the PD-10 column was dissolved in 80 μ l of deuterated buffer (glycine 20 mM, pH 3.0 containing GndCl 2 M) and incubated for 1.5 hours at 70°C in order to exchange completely all labile protons and, afterwards, it was maintained at room temperature for 30 minutes and diluted to a final volume of 400 μ l of deuterated refolding buffer (TrisHCl 0.1 M, pD 8.5 containing 0.1 mM GSSG) and aliquots of the refolding mixture were taken in a time-course manner, trapped by acidification with DCl 1 M (final pD 2.0) and lyophilized.

VI.B.2 Isolation of refolding intermediates

For LCI, an aliquot of the sample eluted in the PD-10 column was dissolved in 400 μ l of refolding buffer (TrisHCl 0.1 M, pH 8.5) containing 0.5 mM GSSG. Half of the refolding mixture was trapped at 2 minutes by mixing with water containing 1% TFA and purified by RP-HPLC as previously described. The reduced form, and a mixture containing 1 and 2-disulfide forms were collected (Figure 4A1). The other half of the refolding mixture was trapped at 1 hour by mixing with an equal volume of water containing 1% TFA. The 3-native disulfide form, a mixture of scrambled species and the native form, were purified by RP-HPLC (Figure 4A2).

For hirudin, two aliquots of the sample eluted in the PD-10 column were dissolved in 400 μ l of refolding buffer. One of the aliquots was dissolved in TrisHCl 0.1 M, pH 8.5 containing 0.1 mM GSSG and the refolding process was stopped at 1 minute by addition of water containing 1% TFA. From this sample, the reduced form and a mixture of species of 1 and 2-disulfides were purified by RP-HPLC (Figure 2.A1). The second aliquot was dissolved in TrisHCl 0.1 M, pH 8.5 containing 0.1 mM GSSG and after 24 hours the refolding mixture was trapped by acidification with water containing 1% TFA and the native form and a mixture of scrambled species were collected from the RP-HPLC purification (Figure 2.A2).

All isolated intermediates were lyophilized in 5 μ g aliquots. LCI samples were resuspended in 5 μ l of deuterated buffer (glycine 20 mM, pH 3.0), incubated at 90°C for 1.5 hours to exchange all labile protons and afterwards were maintained at room temperature for 30 minutes to refold properly. Hirudin samples were dissolved in 5 μ l of deuterated buffer (glycine 20 mM, pH 3.0 containing GndCl 2 M), incubated for 1.5 hours at 70°C to exchange all labile protons, diluted 1:4 with deuterated water to a final 0.4 M GndCl concentration and maintained at room temperature for 30 minutes.

VI.B.3 Deuterium to proton exchange (D/H)

Samples were diluted 1:4 with ammonium citrate (50 mM, pH 4.0) to start the hydrogen exchange. Samples were taken in a time-course manner and mixed 1:1 with sinapic acid (pH 2.0) to be analyzed by MALDI-TOF MS. The sample acidification, together with the so-called crystallization process, slows down considerably the exchange rate.

VI.B.4 Analysis of disulfide species

Fractions of acid-trapped refolding samples (\approx 10 μ g) were lyophilized and free cysteines (cys) were blocked with iodoacetic acid (1 M) in 30 μ l of Tris.HCl buffer (0.5 M, pH 6.5) containing 40% (by volume) of dimethylformamide. The carboxymethylation reaction was allowed to proceed for 30 minutes at room temperature. At these conditions the equilibration of intermediate isomers during carboxymethylation is inhibited (Chang, 1993b). Samples were diluted 1:9 with water containing 0.1% TFA and were analyzed by MALDI-TOF MS. Each carboxymethylated free cys increases the molecular mass by 59 Da.

VI.B.5 MALDI-TOF MS analysis

Molecular masses were determined by MALDI-TOF MS on a BRUKER-BIFLEX spectrometer. Ionization was accomplished with a 337 nm pulsed nitrogen laser and spectra were acquired in the linear positive ion mode, using a 19 Kv acceleration voltage. Deflection of the low mass ions was

used to enhance the target protein signal. Samples were prepared mixing equal volumes of the protein solution and a saturated solution of sinapic acid, used as a matrix, in aqueous 30% acetonitrile with 0.1% TFA.

In the D/H exchange experiments, two sample aliquots were analyzed at each time and each sample was analyzed twice. The media of the mass values corresponding to the centroid of the peaks and the standard deviation was calculated for each time and compared to an external unlabeled control (next spot), whose mass was determined by duplicate. The number of retained deuterons was calculated by subtraction of the average mass of the control from the average mass of each sample and the accuracy in mass determination was calculated as:

$$\left. \begin{array}{l} \text{Sample: } x \pm a \\ \text{Control: } y \pm b \end{array} \right\} (x-y) \pm (a^2 + b^2)^{1/2}$$

VI.C Results

VI.C.1 Measurement of D/H exchange of hirudin during its oxidative refolding by MALDI-TOF MS

Hirudin is a thrombin specific inhibitor isolated from *Hirudo medicinalis*. The sequences of the natural variants contain from 64 to 66 residues and 6 cys that form 3 disulfide bonds in the native structure. Hirudin core domain (amino acids 1-49) is biologically active and is the only part of the protein that requires proper folding to maintain its biological activity (Chang et al, 1990). The folding pathway of hirudin (1-49) has been previously described and is characterized by a first stage of nonspecific disulfide pairing and a second stage of disulfide reshuffling (Chatrenet and Chang, 1993). In the first stage, the flow of disulfide species advances from fully reduced hirudin to equilibrated 1-disulfide, 2-disulfide and finally 3-disulfide (scrambled) forms. The second stage is the limiting step and it is characterized by the reshuffling of the scrambled species to reach the native disulfide pairing. Hirudin shares with PCI and TAP a similar oxidative folding mechanism (Chang et al, 1994 and Chang, 1996), with no evidence of stable 2-disulfide intermediates along the folding process.

Hirudin oxidative folding experiments were performed in the presence of 0.1 mM GSSG in Tris.HCl buffer (0.1 M, pH 8.5). Time-course acid-trapped samples were analyzed by RP-HPLC (Fig 1A) and the disulfide content of their folding intermediates was calculated by integration of the MALDI-TOF mass spectra of the carboxymethylated samples (carboxymethylation of free cys increases the protein molecular weight by 59). The refolding samples chosen to further characterize the stability of their intermediates were: the 1 min refolding sample, which contains reduced hirudin and 1 and 2-disulfide intermediates as major species; the 15 min sample, which contains 1 and 2-disulfide intermediates as major species; the 120 min sample, which contains 2 and 3-

disulfide intermediates, as major species and finally the 1440 min sample, which contains 100% 3-disulfide intermediates, with about 60% of native form (data obtained from the integration of the peak areas of the corresponding RP-HPLC profile)(Fig 1B).

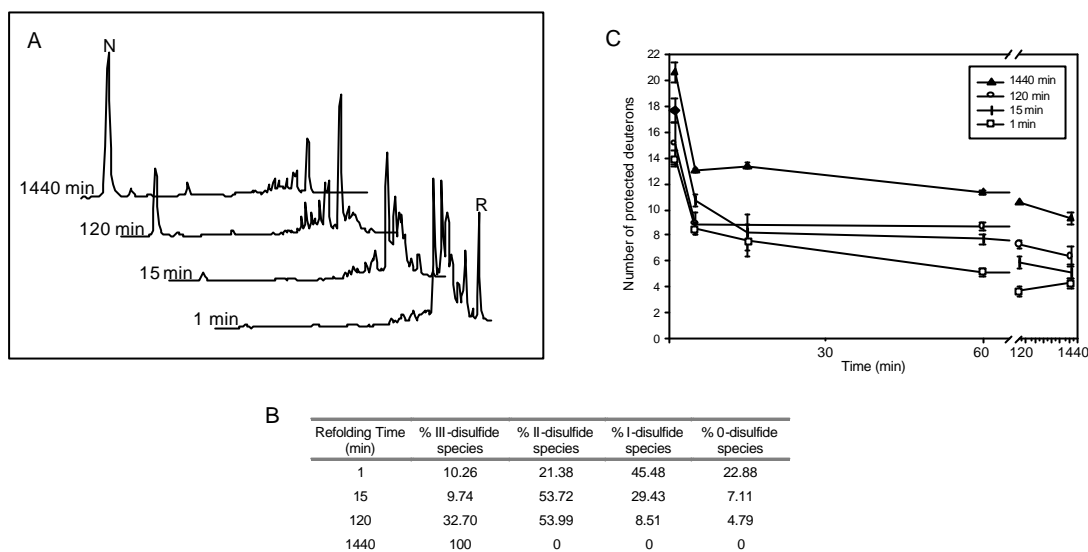


Figure 1. Characterization of time-course *in vitro* refolding samples of hirudin in the presence of 0.1 mM GSSG. (A) Samples were analyzed by RP-HPLC on a vydac C4 column. Native hirudin (N) and reduced form (R) are indicated. (B) The disulfide content of the samples was calculated by MALDI-TOF MS analysis of the carboxymethylated molecules. (C) D/H exchange experiments of the refolding samples detected by MALDI-TOF MS. The number of protected deuterons (increment in mass relative to the protonated protein) as a function of exchange time.

Lyophilized deuterated aliquots corresponding to these four refolding times were used to perform the D/H exchange experiments and the number of deuterons retained in the samples was calculated at different times by MALDI-TOF MS analysis (see experimental procedures). The kinetic plots of the D/H exchange of the refolding samples show that the number of retained protons increases along the folding pathway, indicating that the folding intermediates display a higher conformational stability when the folding process progresses (Fig 1C). We find the major differences between the 120 min and the 1440 min refolding samples, which suggests that exists an important difference in stability between the 3-disulfide isomers and the other disulfide intermediates. Additionally, we observe that the 1 min refolding sample presents 4 retained deuterons, although it contains basically reduced hirudin and 1 and 2-disulfide species. Since reduced hirudin is an almost unfolded form, such extra protection observed in this refolding time should be due to the mixture of 1 and 2-disulfide species.

VI.C.2 Measurement of D/H exchange of isolated intermediates of hirudin oxidative refolding by MALDI-TOF MS

Refolding samples are a mixture of hirudin species. In order to characterize the stability of the intermediates containing different number of disulfide bonds, the refolding samples were analyzed by RP-HPLC and the intermediates were isolated. Denatured and reduced hirudin was allowed to refold in the presence of 0.1 mM GSSG. From the 1 min refolding sample the reduced form and a mixture of 1 and 2-disulfide species were isolated (Fig 2.A1) and from the 1440 min sample the native form and the scrambled species were purified (Fig 2.A2). These four ensembles of molecules were lyophilized and deuterated and the D/H reactions were performed at room temperature, taking samples in a time-course manner and analyzing them by MALDI-TOF MS.

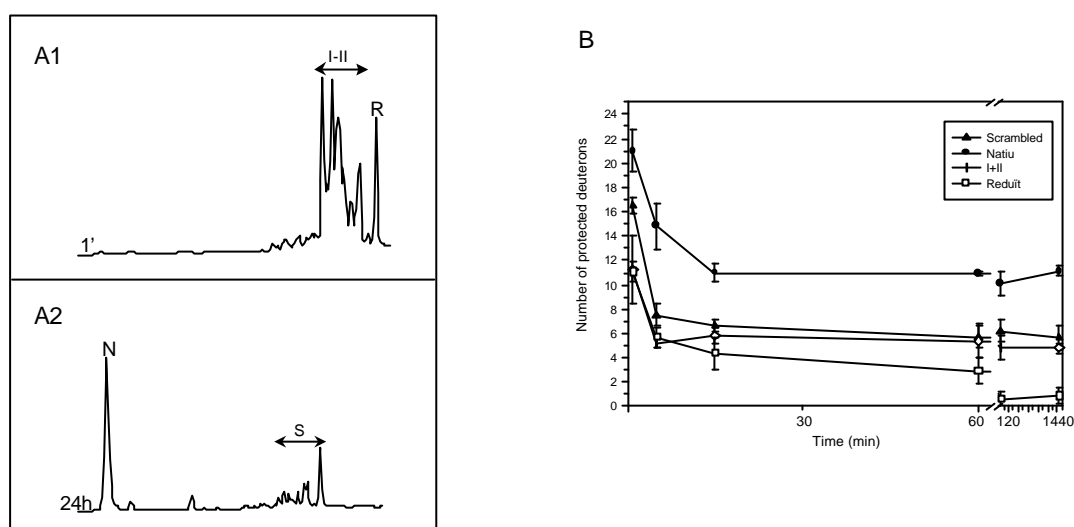


Figure 2. D/H exchange of isolated intermediates of hirudin folding. (A) RP-HPLC chromatographic profiles of refolding samples of hirudin in the presence of 0.1 mM GSSG. (A1) 1 min refolding sample of hirudin. The isolated species are indicated: Reduced form (R) and the mixture of 1 and 2-disulfide species (I-II). (A2) 24 hour refolding sample of hirudin. The isolated species are indicated: scrambled forms (S) and native form (N). (B) Exchange of fully deuterated intermediates of hirudin folding by MALDI-TOF MS as a function of time. The number of protected deuterons is the increment in mass (Da) relative to the protonated protein.

Native hirudin showed the highest number of protected deuterons, nearly twice the number of retained deuterons by the scrambled species (Fig 2B), which indicates that the conformational stability of native hirudin is not only due to the content of disulfide bonds, but also to other non-covalent interactions that are present in the protein when it adopts the proper disulfide pairing. The results also suggest that in the case of hirudin the disulfide-bond formation is absolutely necessary to achieve some degree of stability, since the reduced form hardly presents protected deuterons.

VII.C.3 Measurement of D/H exchange of LCI during its oxidative refolding by MALDI-TOF MS

LCI is a 65 amino acid carboxypeptidase inhibitor isolated from *Hirudo medicinalis* (Reverter et al, 1998), and its sequence contains 8 cys that form 4 disulfide bonds in its native structure (Reverter

et al, 2000). The *in vitro* disulfide folding pathway of LCI has been previously studied and it is characterized by a first stage of nonspecific disulfide pairing, which is characterized by the flow of fully reduced LCI through equilibrated 1-disulfide, 2-disulfide and 3-disulfide isomers to reach the scrambled species and a second stage of disulfide reshuffling to allow the native disulfide pairing formation (Salamanca et al, unpublished results). It has been hypothesized that the presence of a stable 3-disulfide isomer during the oxidative refolding of LCI could correspond to a kinetic trap. On these basis, LCI oxidative refolding would resemble that of EGF, the folding of which also progresses through a 2-disulfide stable isomer that contains two native disulfide bonds (Chang et al, 1995).

Time-course acid-trapped refolding samples of LCI were analyzed by RP-HPLC (Fig. 3A). Four representative refolding times were chosen and their disulfide contents were analyzed by MALDI-TOF MS (Fig 3B): the 5 min sample, which contains reduced LCI, 1, 2 and 3-disulfide isomers and scrambled forms as major species; the 15 min sample, which contains 1, 2 and 3-disulfide isomers and scrambled forms as major species; the 3 hour sample, which contains scrambled forms, a 3-disulfide stable isomer and native form; and the 7 hour sample, which contains native LCI and the 3-disulfide stable isomer as major species.

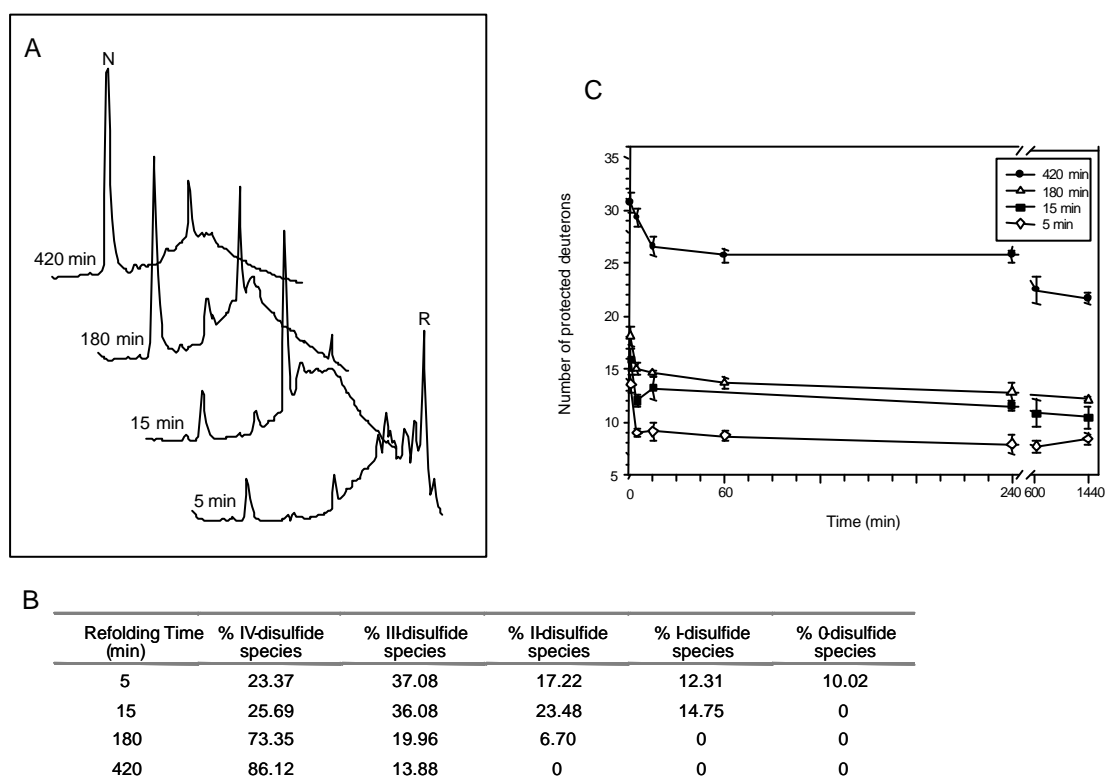


Figure 3. Characterization of time-course refolding samples of LCI. (A) Samples were analyzed by RP-HPLC on a Vydac C4 column. Native hirudin (N) and reduced form (R) are indicated. (B) The disulfide content of the samples was calculated by MALDI-TOF MS analysis of the carboxymethylated molecules. (C) D/H exchange experiments of the refolding samples detected by MALDI-TOF MS. The number of protected deuterons (increment in mass relative to the protonated protein) as a function of exchange time.

To perform the D/H exchange experiments, lyophilized deuterated aliquots corresponding to the 5 min, 15 min, 3 hour and 7 hour refolding samples were resuspended in 10 μ l of ammonium citrate (pH: 4.0) to initiate the D/H exchange reaction, and samples were taken in a time-course manner and analyzed by MALDI-TOF MS. The number of deuterons protected from D/H exchange was calculated in each time of the exchange reaction for all four folding samples (see experimental procedures). The kinetic plots of the D/H exchange show that the number of deuterons retained by the LCI species in the folding samples increased along the folding process (Fig 3C). The kinetics of the process was similar for all tested samples, and in all cases an exchange *plateau* was reached about 24 hours after the D/H exchange reaction was initiated. The number of retained deuterons increased from the 5 min refolding sample to the 15 min and finally to the 3 hour refolding sample. However, in no case the differences were as important as observed present between the 3 hour and the 7 hour refolding samples. In fact, both samples contain a similar disulfide content, but in the 3 hour sample about 50% of the 4-disulfide species are scrambled forms, and the 7 hour sample only contains a 10% of scrambled forms (data obtained from the integration of the peak areas of the corresponding RP-HPLC profiles). This suggest that there is an important difference in stability between the native and the scrambled forms.

VI.C.4 Measurement of D/H exchange of isolated intermediates of LCI oxidative refolding by MALDI-TOF MS

In order to characterize the D/H exchange kinetic plots of the most representative intermediates of LCI oxidative refolding, denatured and reduced LCI was refolded for 1 hour in the presence of 0.5 mM GSSG and afterwards it was acid-trapped and analyzed by RP-HPLC to isolate the 3-disulfide stable intermediate (Salamanca et al, unpublished results), scrambled species and the native form (Fig. 4A1). Another sample was refolded for 2 minutes in the presence of 0.5 mM GSSG and afterwards it was acid-trapped and analyzed by RP-HPLC to isolate a mixture of 1 and 2-disulfide isomers and the reduced form. Aliquots of each sample were lyophilized and deuterated and, afterwards, the D/H exchange experiments were initiated. Samples were taken in a time-course manner and analyzed by MALDI-TOF MS as previously described.

Significant differences were observed among the kinetic plots of the D/H exchange reaction of the 5 isolated intermediates. Native LCI showed the highest number of protected deuterons -about 27-, a fact which agrees with the 7-hour folding sample that contained native LCI as the major specie (about 75%) and presented 23 protected deuterons. Since these samples proceed from two different sources, and have been deuterated in different ways, this gives us indication that none of these parameters does interfere in the final number of protected deuterons.

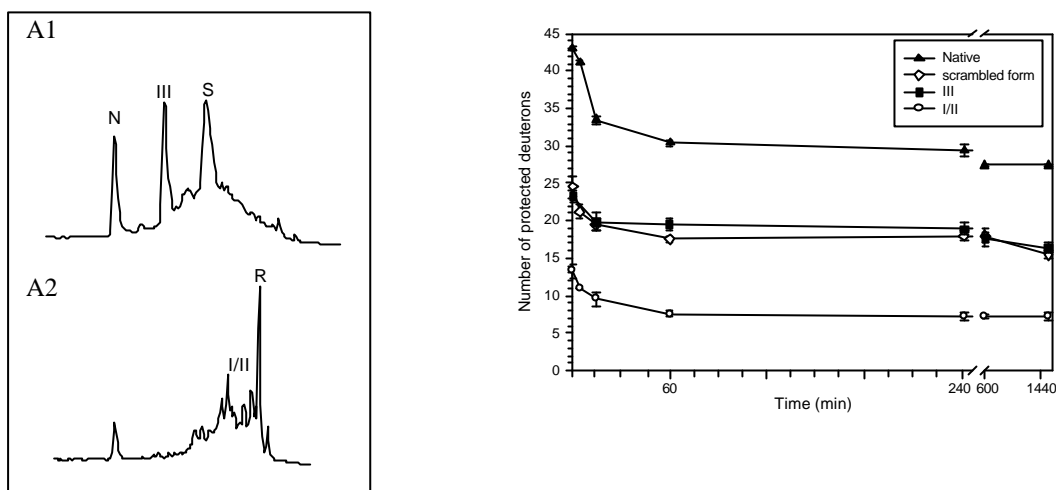


Figure 4. D/H exchange of isolated intermediates of LCI folding. (A1) RP-HPLC chromatographic profile corresponding to the 60-minute refolding sample in the presence of 0.5 mM GSSG. The isolated species are indicated: scrambled forms (S), 3-disulfide stable isomer (III), and native form (N). (A2) RP-HPLC chromatographic profile corresponding to the 2-min refolding sample in the presence of 0.5 mM GSSG. The isolated species are indicated: reduced form (R), 1 and 2 disulfide species (I/II). (B) Exchange of fully deuterated intermediates of LCI folding by MALDI-TOF MS as a function of time. The number of protected deuterons is calculated from the increment in mass relative to the protonated protein.

Interestingly, the scrambled forms and the 3-disulfide stable isomer showed a similar protection. Although the molecules possess a different number of disulfide bonds -three and four, respectively- they apparently present the same conformational stability. Finally, the mixture of 1 and 2-disulfide species shows 7 protected deuterons, about 1/4 of those found in the native form, indicating that these species are rather unstable.

VI.D Discussion

Results obtained in this work show that the use of the hydrogen exchange methodology previously reported by our group (Villanueva et al, 2000) can be applied to the study of protein conformational stability, providing an informative complement to the acid trapping method to study the folding pathway of disulfide-rich proteins. In this work, the acid trapping approach monitored by RP-HPLC, has been used to group and fractionate the different folding intermediates whereas the hydrogen exchange method monitored by MALDI-TOF MS, has been used to characterize the unfolding extent of the different species along the folding pathway. The comparison between both methods indicates that the increasing amount of deuterons retained during the refolding samples can be assigned to the disappearance of reduced, 1 and 2-disulfide forms and to the appearance of scrambled forms and, finally, to native form.

VI.D.1 Comparison of the folding pathways of LCI and hirudin.

When comparing the LCI and hirudin folding described by the acid-trapping RP-HPLC method it was concluded that these proteins were models of two different folding mechanisms (Chang et al, 2000; Salamanca et al., unpublished results). Folding of hirudin is characterized by a high heterogeneity of 1 and 2-disulfide intermediates and by the presence of scrambled 3-disulfide isomers as folding intermediates. Differently, during LCI folding, apart from highly heterogeneous 1 and 2-disulfide intermediates and scrambled 4-disulfide isomers, a stable 3-disulfide intermediate, that accumulates rapidly along the folding pathway, could be detected.

In the case of hirudin we observe a growing protection along the folding pathway: the 0, 1 and 2-disulfide intermediates showed a number of protected deuterons lower than the 3-disulfide scrambled forms, whilst the protection of these scrambled forms was inferior to that of native hirudin. In contrast, along LCI folding not all intermediates with a higher number of disulfide bonds showed a higher number of protected deuterons. Both, LCI scrambled forms -containing 4 disulfide bonds- and the stable 3-disulfide intermediate presented a similar protection in the D/H exchange reactions despite having a different disulfide bond content (Fig. 5). This apparently shocking result could imply that such LCI 3-disulfide intermediate displays a similar stability to that of the 4-disulfide scrambled forms.

VI.D.2 LCI kinetic trap

It has been suggested that the last NH hydrogens to exchange in a protein are located in the most rigid part of its structure and constitute a three-dimensional cluster called the slow exchange core, which is expected to be formed by segments of secondary structure that are tightly packed and collapse early during folding (Kim et al, 1993 and Woodward, 1993). Thus, we could assume from our hydrogen exchange results that the LCI 3-disulfide stable intermediate contains an amount of secondary structure higher than expected. This fact would contribute to the over-stabilization of the intermediate, promoting its accumulation in relation to the other 3-disulfide intermediates. Therefore, it would act as a kinetic trap in the same way that EGF-II acts in the EGF folding pathway (Chang et al, 1995). Whether this extra amount of secondary structure present in the III-disulfide stable intermediate is native-like, it is not clear yet. Actually, this hypothesis is being tested in our laboratory through structural characterization of such a 3-disulfide intermediate. In the case of EGF-II, this intermediate needs to undergo a general rearrangement and be transformed into different scrambled forms before acquiring the native disulfide bonds, a fact which seems to indicate that the EGF-II structure is not native-like.

VI.D.3 The intert ining dependence of the disulfide bonds and non-covalent forces

Disulfide bonds are known to enforce the stability of native proteins (Chang, 1997). Less certain is how disulfide bonds interact with non-covalent forces to stabilize the folded native conformation. For hirudin, PCI and TAP proteins, the non-covalent interactions do not actively drive or direct the folding until non-native disulfides are formed. In this first stage of folding the role of the nonspecific disulfide pairing is to cross-link the protein at its unfolded state to decrease its entropy. Otherwise, the second stage in the folding process, that is the conversion of scrambled disulfides to the native disulfides, is driven by the non-covalent interactions (Chang, 1997). For LCI and EGF the folding process is similar to that of hirudin, PCI and TAP. However, non-covalent interactions are probably involved in the formation of a stable intermediate during the first stage of the folding process, which facilitates the formation of scrambled species.

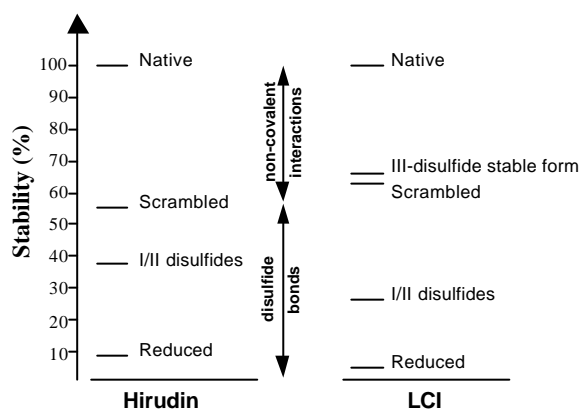


Figure 5. Stability of hirudin and LCI isolated refolding intermediates. The stability of the intermediates was calculated as the percentage of protected deuterons retained by each intermediate relative to the deuterons retained by the native form.

The acquisition of conformational stability along the folding pathway for hirudin and LCI takes place, similarly, in two stages. In the first stage a gradual increase in protein stability is observed. About 60% of the native state stability is achieved in the transition from the reduced form to the scrambled species (through folding intermediates with an increasing number of disulfide bonds) (Fig.5). In the second stage a step rise in conformational stability from the scrambled to the native form is detected. These data show that even though the scrambled forms contain the same number of disulfide bonds than the native form, they are much more unstable forms (about a 40% difference in conformational stability). This 40 % is an estimation of the contribution of non-covalent interactions to the folding process of the small disulfide-rich proteins. These results agree with the large differences observed between the inhibitory activity of native PCI and their scrambled forms. PCI is also a small disulfide-rich protein that inhibits carboxypeptidase A with an inhibition constant in the nanomolar range (Molina et al, 1994). When the inhibitory activities of PCI scrambled forms were tested, the results showed that the inhibition constants were three orders

of magnitude higher (in the micromolar range) (Pavía et al, unpublished results). In addition, several attempts of resolving their NMR three-dimensional structures have been unsuccessful. Finally, the NMR studies of α -conotoxin scrambled forms showed that these species present an increase in conformational heterogeneity -each scrambled form presents multiple conformers in solution- and a decrease in thermodynamic stability (Gehrmann et al, 1998). All these facts could be explained if we consider the scrambled forms as highly flexible unstable species that only present the disulfide bonds as major stability elements and without the contribution of non-covalent interactions.

VI.E References

- Anfinsen CB, Haber E, Sela M and White FHJ; (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain; *Proc Natl Acad Sci USA*; 47:1309-1314
- Beavis RC and Chait BT; (1990) Rapid, sensitive analysis of protein mixtures by mass spectrometry; *Proc Natl Acad Sci USA*; 68:6873-6877
- Chang J-Y, Li L and Bulychiev A; (2000) The underlying mechanism for the diversity of disulfide folding pathways; *J Biol Chem*; 275: 8287-8289
- Chang JY; (1997) A two-stage mechanism for the reductive unfolding of disulfide-containing proteins; *J Biol Chem*; 272:69-75
- Chang JY; (1996) The disulfide folding pathway of tick anticoagulant peptide (TAP), a Kunitz-type inhibitor structurally homologous to BPTI; *Biochemistry*; 35:11702-11709
- Chang JY, Schindler P, Ramseier U and Lai PH; (1995) The disulfide folding pathway of human epidermal growth factor; *J Biol Chem*; 270:9207-9216
- Chang JY; (1993b) Identification of productive folding intermediates which account for the flow of protein folding pathway; *J Biol Chem*; 268:4043-4049
- Chatrenet B and Chang J-Y; (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments; *J Biol Chem*; 268: 20988-20996
- Creighton TE; (1986) Detection of folding intermediates using urea-gradient electrophoresis; *Methods Enzymol*; 131:83-106
- Gehrmann J, Alewood PF and Craik DJ; (1998) Structure determination of the three disulfide bond isomers of α -conotoxin GI: a model for the role of disulfide bonds in structural stability; *J Mol Biol*; 278:401-415
- Ghaemmaghami S, Fitzgerald MC and Oas TG (2000) A quantitative, high-throughput screen for protein stability; *PNAS*; 97: 8296-8301
- Fersht AR and Daggett V; (2002) Protein folding and unfolding at atomic resolution; *Cell*; 108:573-582
- Jeng MF, Englander SW, Elove GA, Wand AJ and Roder H; (1990) Structural description of acid-denatured cytochrome c by hydrogen exchange and 2D NMR; *biochemistry*; 29:10433-10437
- Konishi Y, Ooi T and Scheraga HA; (1982) Regeneration of ribonuclease A from reduced protein. Rate-limiting steps; *Biochemistry*; 21: 4734-4740

- Mandell JG, Falick AM and Komives EA; (1998) Identification of protein-protein interfaces by decreased amide proton solvent accessibility; *Proc Natl Acad Sci USA*; 95: 14705-14710
- Miranker A, Robinson CV, Radford SE and Dobson CM; (1996) Investigation of protein folding by mass spectrometry; *FASEB J*; 10: 93-101
- Molina MA, Marino C, Oliva B, Aviles FX and Querol E; (1994) C-tail valine is a key residue for stabilization of complex between potato inhibitor and carboxypeptidase A; *J Biol Chem*; 269:21467-21472
- Reverter D, Vendrell J, Canals F, Horstmann J; Aviles FX, Fritz H and Sommerhoff CP; (1998) A carboxypeptidase inhibitor from the medical leech *Hirudo medicinalis*. Isolation, sequence analysis, cDNA cloning, recombinant expression and characterization; *J Biol Chem*; 273:32927-32933
- Reverter D, Fernandez-catalan C, Baumgartner R, Pfander R, Huber R, Bode W, Vendrell J, Holak TA and Aviles FX; (2000) Structure of a novel leech carboxypeptidase inhibitor determined free in solution and in complex with human carboxypeptidase A2; *Nat Struct Biol*; 7:322-328
- Roder H, Elove GA and Englander SW; (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR; *Nature*; 335:700-704.
- Thannhauser TW, Sherwood RW and Scheraga HA; (1998) Determination of the cysteine and cystine content of proteins by amino acid analysis: application to the characterization of disulfide-coupled folding intermediates; *J Protein Chem*; 17:37-43
- Udgaonkar JB and Baldwin RL; (1988) NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A; *Nature*; 335:694-699
- Villanueva J, Canals F, Villegas V, Querol E and Aviles FX; (2000) Hydrogen exchange monitored by MALDI-TOF mass spectrometry for rapid characterization of the stability and conformation of proteins; *FEBS Lett*; 472:27-33
- Weissman JS and Kim PS; (1992) The disulfide folding pathway of BPTI; *Science*; 256: 112-114
- Weissman JS and Kim PS; (1991) Reexamination of the folding of BPTI: Predominance of native intermediates; *Science*; 253:1386-1393

VII- Resum i discussió general

El PCI, la hirudina i el LCI formen part del grup de proteïnes petites i riques en ponts disulfur. Dins aquest grup de proteïnes trobem factors de creixement, toxines, inhibidors de proteases... que són molècules amb potencials aplicacions biotecnològiques. Des del punt de vista de la seva producció recombinant, són proteïnes de difícil obtenció, ja que l'elevat nombre de ponts disulfur fan que el seu rendiment de producció sigui baix, i que en molts casos s'acumulin formes mal plegades, o bé que es formin cossos d'inclusió. Des del punt de vista del seu plegament són proteïnes àmpliament estudiades, ja que el seu plegament està dirigit principalment per la formació dels ponts disulfur; i això fa que siguin una bona aproximació al procés general de plegament proteic.

En el primer treball d'aquesta tesi s'aprofundeix en l'estudi de la producció recombinant del PCI en flascó. S'avaluen tres sistemes d'expressió extracel·lular de PCI: un sistema eucariota (*P.pastoris*) i dos sistemes procariontes (*E.coli*). S'optimitzen diversos paràmetres per tal de millorar el seu nivell de producció recombinant; per exemple la font de carboni, l'agent inductor de l'expressió o el temps de producció, i es busquen sistemes que ens permetin simplificar el seu procés de purificació.

En el segon treball s'analitza el paper de les regions precursoras del PCI, especialment de la pro-regió N terminal, en el plegament de la forma madura. S'estudia el replegament *in vitro* de les diverses pro-formes del PCI per tal de caracteritzar la seva influència sobre la velocitat o la eficiència del procés. També s'estudia la influència de la pro-regió N terminal *in vivo* en *E.coli* per alguns mutants del PCI i finalment es duu a terme una caracterització estructural de la forma corresponent al PCI amb la pro-regió N-terminal.

En el tercer treball també s'analitza el procés del plegament proteic, però a través de l'estudi de dues proteïnes diferents, que també pertanyen al grup de proteïnes petites i riques en ponts disulfur: la hirudina i el LCI. Es determina el seu procés de plegament mitjançant la caracterització de l'estabilitat conformacional dels seus intermediaris de plegament usant la tècnica del bescanvi D/H i es compara amb els resultats obtinguts per altres aproximacions experimentals. S'estudien, des d'un punt de vista d'estabilitat, les diferències existents entre els dos models de plegament que representen les dues proteïnes, i es calcula la importància de les interaccions covalents i no-covalents en les proteïnes natives.

En el primer treball (descriu en l'apartat IV d'aquesta tesi) es desenvolupen diversos sistemes de producció de PCI mitjançant expressió heteròloga en flascó. Fins al moment d'iniciar-se el treball, el sistema descrit per a la producció de PCI en flascó era mitjançant el vector pIN3OmpAIII i la soca MC1061 d'*E.coli* i menava a una producció de PCI entorn de 1-2 mg/L (Molina i col., 1992). El primer sistema estudiat en aquest treball fou el del llevat metil·lotròfic *P. pastoris*, àmpliament descrit com a productor d'alts nivells de proteïnes heteròlogues (Cerghino i Cregg, 2000). A causa de l'elevat nombre de ponts disulfur de la nostra molècula es va escollir un vector d'expressió

extracel·lular (pPIC9) i es va comprovar l'estabilitat de la nostra proteïna en el medi de cultiu, fins a les 72 hores. En el triatge analític realitzat amb 20 colònies transformades a la soca GS115 es va observar que hi havia una variació important en la producció de la proteïna recombinant que abarcava des d'uns nivells pràcticament nuls fins a uns 5 mg/L. Aquesta vasta variació està probablement relacionada amb el fenotip d'utilització de metanol. En general, s'observava una tendència dels clons mut^s a una producció més elevada de PCI, la qual cosa indica que un creixement lent afavoreix la producció de la nostra proteïna.

L'expressió heteròloga del PCI en *P.pastoris* en condicions òptimes, a partir del clon escollit en el triatge analític i usant flascons de 2 litres, va conduir a una producció final de 5,5 mg/L de la nostra proteïna; unes tres vegades superior al sistema de partida (pIN3OmpAIII -PCI). El fet que a més del PCI també s'hagin descrit bons nivells de producció per altres proteïnes riques en ponts disulfur (Rosenfeld i col., 1996; Laroche i col., 1994; Brankamp i col., 1995) que s'han expressat extracel·lularment a *P.pastoris* sense problemes importants, sembla indicar que aquest és un sistema idoni per a l'obtenció d'aquest tipus de proteïnes. En altres sistemes, com ara *E.coli*, l'expressió extracel·lular d'aquest tipus de proteïnes acostuma a donar uns rendiments de producció de proteïna molt més baixos i obliga a abordar la producció a través d'una expressió intracel·lular, que condueix a la formació de cossos d'inclusió, obligant a un posterior replegament de la proteïna.

La segona estratègia que vam emprar per a l'expressió heteròloga del PCI fou la fusió a una cua d'histidines per l'extrem N-terminal i producció en la soca MC1061 d'*E.coli* mitjançant el vector pIN3OmpAIII. La principal finalitat d'aquest sistema era aconseguir purificar la proteïna en una única etapa cromatogràfica mitjançant una cromatografia d'afinitat de Ni²⁺, que es basa en l'afinitat que presenten les histidines per alguns ions metàl·lics, entre ells el níquel; i d'aquesta manera eliminar passos cromatogràfics que disminueixen el rendiment proteic. Entre la cua de 6 His i el PCI es va col·locar un residu d'Arg, que és una diana de tall de la tripsina, per tal de poder eliminar la cua després de purificar la proteïna, si fos necessari. Es va emprar el vector pIN3OmpAIII, el mateix que en el sistema de partida, que conté un promotor *lpp-lac* induïble mitjançant IPTG, i que integra el gen de resistència a l'ampicil·lina i el pèptid senyal (PS) OmpA que condueix a una expressió periplasmàtica. Tot i que es va aconseguir la producció de proteïna biològicament activa (mostrava activitat inhibidora vers la CPA), aquesta no presentava la cua de histidines N-terminal. El seguiment per MALDI-TOF MS de la producció de la nostra proteïna ens va permetre observar que es produïa una degradació ràpida i seqüencial a partir de l'extrem N-terminal de la cua d'histidines en el periplasma d'*E.coli*.

Aquest fet, juntament amb la degradació també observada en l'expressió extracel·lular del ProNtPCI (PCI amb la proseqüència N-terminal, descrit al treball 2, apartat V d'aquesta tesi) fa pensar que, a diferència de la gran estabilitat que caracteritza la proteïna madura, les extensions a la zona N-terminal de la proteïna generen variants molt làbils i susceptibles de degradació. Els tres ponts disulfur confereixen a la proteïna madura un nucli central molt estable, però l'addició de

residus a la zona N-terminal segurament provoca que l'estructura de la cua N-terminal esdeveningui més làbil, convertint-se en un candidat fàcil per a l'acció de les peptidases.

Finalment vam decidir expressar la nostra proteïna en un sistema de secreció extracel·lular però utilitzant un vector derivat dels pET (Studier i col., 1998), on la proteïna es troba sota el control d'un promotor fort, el de la T7 RNA polimerasa, i la seva expressió és induïda per addició d'IPTG al cultiu. El vector escollit fou el pBAT4 (Peranen i col, 1996) amb el PS OmpA, que en principi condueix a una expressió periplasmàtica. Es va observar una acumulació majoritària del PCI al medi extracel·lular, tal com també s'ha descrit per altres proteïnes (Hewinson i Russel, 1993 i Manosroi i col., 2001).

La soca d'*E.coli* emprada fou la BL21(DE3) que és deficient en algunes proteases, per tal de minimitzar els processos de degradació. Mitjançant estudis analítics d'expressió es varen optimitzar diversos paràmetres que afecten a la producció proteica, tals com la naturalesa o concentració de la font de carboni i el tipus d'agent inductor de l'expressió. Com en el cas abans mencionat (*P.pastoris*) es va observar que un creixement lent (en presència de glicerol) afavoria la producció de PCI; en canvi si aquest creixement tenia lloc de forma ràpida (en presència de glucosa) la producció disminuïa. Aquest fet, també observat en l'expressió d'altres proteïnes, s'ha relacionat amb la saturació de la maquinària d'exportació quan es treballa amb creixements molt ràpids, la qual cosa podria conduir a la formació de cossos d'inclusió per part de la proteïna estudiada. L'optimització de la producció de PCI en aquest sistema va permetre obtenir la proteïna a una concentració d'uns 15 mg/L, en flascó. Això representa un augment de 10 vegades respecte el sistema de partida.

És per tant evident, que en els dos sistemes d'expressió de PCI amb secreció extracel·lular aquí investigats, en *P.pastoris* i en *E.coli*, s'obté un major rendiment de producció de proteïna. En el cas de les proteïnes riques en ponts disulfurs, la secreció de les proteïnes recombinants és una estratègia a bastament utilitzada, ja que l'espai periplasmàtic, a diferència del citosol, és un compartiment oxidant i, per tant, permet la formació dels ponts disulfur. En el citosol d'*E.coli*, a causa del seu caràcter reductor, s'acumulen formes mal plegades, que en molts casos contenen superfícies hidrofòbiques exposades que promouen l'agregació de les molècules i la formació de cossos d'inclusió. En canvi, el caràcter oxidant del periplasma, juntament amb la presència de cisteïn òxido-reductases (Dsb) que catalitzen la formació dels ponts disulfurs, contribueixen al plegament correcte d'aquestes proteïnes. Cal remarcar que en ambdós casos es va utilitzar el mateix sistema de purificació, que és el mateix que el del sistema de partida: una primera etapa d'extracció en fase sòlida (Sep pak C18), una segona etapa cromatogràfica de bescanvi iònic (DEAE-SPW) i una última etapa cromatogràfica d'HPLC en fase reversa (RP-HPLC) (Vydac C4).

El rendiment de producció de PCI en *P.pastoris* fou inferior que a *E.coli*, però el primer sistema presenta algunes avantatges: el volum de mostra amb el qual es comença la purificació és molt més petit, el sistema d'inducció és molt barat i té una major estabilitat clonal, ja que el gen d'interès

s'integra al genoma de *P.pastoris*, i així s'eviten els problemes de segregació que es produeixen amb els plàsmids d'*E.coli*. El principal inconvenient d'aquest sistema és la duració del procés d'expressió, mentre que a *E.coli* només calen 2 dies per a l'expressió en flascó del PCI, a *P.pastoris* són necessaris 4 dies (2 dies per al creixement i dos dies per a l'expressió de la proteïna).

El rendiment de producció de 15 mg/L de PCI a *E.coli* en flascó representa el principal èxit d'aquest treball d'optimització de la seva expressió, ja que ens permet aplicar aquest sistema per a la producció de nous mutants o variants del PCI. Fins aquest moment calia escalar la producció d'aquestes proteïnes en un bioreactor de 2 litres per tal de poder obtenir la quantitat necessària de proteïna per a realitzar estudis estructurals (Raigs X, RMN...) o funcionals (determinació de paràmetres cinètics, cultius cel·lulars...). El sistema optimitzat ens permet arribar a aquestes quantitats en flascó, la qual cosa implica un estalvi molt important de temps i diners.

En el segon treball, descrit a l'apartat V d'aquesta tesi, s'analitza el possible paper de les pro-seqüències N i C-terminal del PCI sobre el plegament de la proteïna madura. Per tal d'obtenir un bon rendiment en l'obtenció de noves variants de PCI, a més d'un bon sistema de producció recombinant, cal que la proteïna que s'obtingui estigui ben plegada. El plegament i de splegament del PCI ha sigut anteriorment estudiat en el nostre grup (Chang i col., 1994; Chang i col., 2000) i s'ha observat que el plegament oxidatiu de la proteïna madura presenta una cinètica molt lenta, fins i tot a concentracions òptimes d'agents redox (2 mM Cys-Cys i 4 mM Cys) calen 3 hores per a replegar el PCI *in vitro* (Chang i col., 1994). En el cas d'altres proteïnes petites riques en ponts disulfur amb un plegament lent *in vitro* s'havia observat que la presència de les pro-seqüències ajudaven al plegament de la proteïna madura (Weissman i Kim, 1992; Hikaka i col., 2000). Això, juntament amb la lenta cinètica de plegament del PCI *in vitro*, ens va fer pensar que fora plausible una acció de les pro-regions del PCI en el plegament de la proteïna madura. La pro-seqüència C-terminal de 7 residus presenta similituds amb altres pro-pèptids de direcció a vacúol (Villanueva i col., 1998), per tant semblava que el pro-pèptid N terminal era el principal candidat a tenir un paper en el procés de plegament.

Amb aquest objectiu, els estudis de replegament oxidatiu *in vitro* es varen realitzar amb el PCI, el ProNtPCI (PCI amb la pro-regió N-terminal), Pro+PCI (PCI amb la pro-regió N-terminal en *trans*) i el ProPCI (PCI amb les pro-regions N i C-terminal). Tant l'eficiència de plegament (quantitat de proteïna nativa assolida en un determinat temps), com la cinètica de plegament (velocitat d'adquisició de la forma nativa) foren similars en tots els casos, indicant que el pro-pèptid no afecta al flux d'intermediaris de I i II ponts disulfur i de formes *scrambled* que segueix el PCI madur per arribar a la forma nativa. Tampoc es va observar cap influència en l'equilibri existent entre les formes *scrambled*, que es trobaven en les mateixes proporcions.

Aquests resultats ens feren replantejar parcialment la nostra hipòtesi de treball i vam pensar que potser l'efecte de la proseqüència no s'observava *in vitro* perquè necessitava la interacció d'algun altre component cel·lular i, per tant, que serien els estudis *in vivo* els que ens donarien una resposta a aquesta possibilitat. A causa de la impossibilitat de dur a terme els experiments *in vivo* en la planta vam decidir realitzar-los en *E.coli*. Per a aquests estudis es van utilitzar el PCI i 3 mutants que presentaven un baix rendiment de producció o poca eficiència de plegament -Y37G PCI, Δ 3 PCI, G35P&P36G PCI- (Venhudova i col., 2000) per tal de veure si l'efecte de la pro-seqüència s'observava millor en els casos on la producció i/o plegament de la proteïna estan desafavorits. Es van generar aquests mutants amb la pro-regió N-terminal i es va comparar el seu nivell d'expressió (producció total de proteïna) i l'eficiència del seu plegament (quantitat de forma nativa present en el cultiu) amb la dels mutants sense la pro-seqüència. Els resultats mostraren que el PCI madur i els seus mutants s'expressen al mateix nivell que les formes corresponents amb la pro-regió N-terminal. Tampoc es varen observar diferències entre la quantitat de forma nativa o la proporció entre forma nativa i *scrambleds*. Per tant, no sembla que la pro-seqüència N-terminal tingui una influència sobre el plegament *in vivo* a *E.coli* de la forma madura del PCI. En el cas que la pro-regió N-terminal influencés el plegament del PCI interaccionant amb alguna proteïna cel·lular, aquesta probablement no es troba present en *E.coli*.

Curiosament, la presència dels propèptids N i C-terminals no tenen una influència negativa sobre el plegament *in vitro* o *in vivo* a *E.coli* de la proteïna madura, tot i que de forma conjunta tenen una mida semblant al PCI madur. L'ús de tres tècniques d'anàlisi estructural: DC, proteòlisi limitada i bescanvi H/D seguits per espectrometria de masses MALDI-TOF, i RMN va indicar la presència d'elements d'estructura secundària i/o terciària dins del propèptid N-terminal del PCI. Així doncs, tot i que no s'hagi observat cap efecte de la pro-seqüència N-terminal del PCI *in vitro* o *in vivo* en *E.coli*, el fet que aquesta pro-regió no presenti una estructura tipus *random-coil* i que no afecti negativament al plegament de la proteïna madura fan pensar que no és descartable un possible efecte sobre el plegament *in vivo* en la planta, o bé intervenint en alguna altra funció, com la modulació de la secreció de la proteïna o bé interaccionant amb algun altre factor cel·lular (chaperones...).

En el tercer treball d'aquest tesi, recollit en l'apartat VI, es caracteritza l'estabilitat conformacional dels intermediaris de plegament de dues proteïnes petites riques en ponts disulfur. El plegament proteic té lloc de manera espontània i depèn de la seqüència aminoacídica de la proteïna (Anfinsen, 1973), però les regles que regeixen el procés són complexes i encara no s'han pogut determinar del tot. La complexitat del procés ha donat una gran rellevància a l'estudi del plegament de les proteïnes petites riques en ponts disulfur, ja que en aquestes proteïnes el plegament està regit principalment per la formació dels ponts disulfur i això permet caracteritzar una part dignificativa del seu procés de plegament d'una manera relativament senzilla, mitjançant la determinació del nombre i aparellament de ponts disulfur dels seus intermediaris.

Un dels mètodes de més àmplia utilització per a la caracterització del flux d'aparellament de disulfurs és la captura en medi àcid seguida per RP-HPLC (Weissman i Kim, 1991). Aquest mètode es basa en la diferent reactivitat dels grups tiol de les cys en funció del pH. En el rang de pH en què normalment s'estudien els replegaments proteics (7-8,5), els grups tiol reaccionen entre si formant enllaços disulfur. Si es baixa el pH de la reacció (pH 2-4), els grups tiol no reaccionen entre si i d'aquesta manera es poden capturar els intermediaris de plegament. Posteriorment es poden separar per RP-HPLC i aïllar-los, es pot caracteritzar el nombre de disulfurs mitjançant la derivatització química de les cys no oxidades (per exemple, per alquilació) o es pot caracteritzar la continuació del procés de plegament d'un intermediari purificat, ja que la captura en medi àcid és reversible (només cal tornar a pujar el pH).

Aquests estudis ens permeten analitzar els flux de disulfurs al llarg del plegament i també ens indiquen quins intermediaris són més productius (intermediaris que donen lloc a forma nativa amb una cinètica més ràpida). La poca quantitat de proteïna de què normalment es disposa i la heterogeneïtat de les mostres de plegament (ja que parlem d'intermediaris aïllats i realment es tracta d'un conjunt de formes amb un mateix aparellament de ponts disulfur però amb un gran ventall de configuracions) provoca que la caracterització de la seva estabilitat o estructura sigui difícil. Els estudis de bescanvi H/D seguits per MALDI-TOF MS ens permeten solventar aquest problema ja que es poden analitzar un gran nombre de mostres, amb poc temps i emprant quantitats mínimes de proteïna.

Els protons amida de les proteïnes bescanvien a unes velocitats que són mesurables experimentalment. La velocitat de bescanvi, a més de dependre de factors externs com la temperatura o el pH, depèn de l'entorn on es troba el protó: si està soterrat o es troba involucrat en una estructura secundària presenta una velocitat de bescanvi molt més lenta que si es troba accessible al dissolvent. Així doncs, la determinació del nombre de protons de bescanvi lent ens dóna informació de l'estabilitat conformacional de la molècula (Villanueva 2000; Ghaemmaghami et al, 2000). El fet que tant les reaccions de bescanvi H/D com les de formació de ponts disulfur siguin fortament dependents del pH, fa que calgui tenir molt en compte aquest paràmetre en el desenvolupament experimental. Per a la deuteració de la proteïna reduïda o d'un intermediari aïllat cal treballar a un pH baix (3-4) per evitar que hi hagi formació de ponts disulfur i si volem accelerar la reacció de bescanvi es pot pujar la temperatura o afegir un agent desnaturalitzant. Per a mesurar la cinètica del procés de bescanvi D/H hem treballat a pH 4.0, de manera que no hi ha formació de ponts disulfur i la reacció de bescanvi és mesurable.

La caracterització de l'estabilitat conformacional de mostres de replegament i d'intermediaris de plegament purificats de la hirudina i del LCI ens ha permès determinar que aquest paràmetre evoluciona de forma paral·lela a la formació dels ponts disulfur. Les dues etapes que caracteritzen el seu plegament (una primera etapa seqüencial on es passa de la forma reduïda a les formes *scrambled* i una segona etapa limitant on s'adquireix la forma nativa) també es poden observar en l'adquisició d'estabilitat per part dels seus intermediaris. Primer hi ha un augment gradual

d'estabilitat desde la forma reduïda, passant pels intermediaris de I i II -ponts disulfur fins a arribar a les formes *scrambled*, i després hi ha un salt d'estabilitat fins arribar a la forma nativa. El fet que les formes *scrambled* presentin més un 60% d'estabilitat respecte la forma nativa ens indica que en el procés de reaparellament de disulfurs es produeix una profunda reorganització estructural que permet que es donin una sèrie d'interaccions no-covalents en la forma nativa, que no es donaven abans i que són responsables de la seva estabilitat. Així doncs, tot i que la formació dels ponts disulfur sigui molt important per al plegament de la molècula (la forma reduïda presenta una estabilitat mínima), les interaccions no-covalents que es donen en la forma nativa també contribueixen de manera essencial en la seva estabilitat i constitueixen l'etapa limitant des del punt de vista termodinàmic.

En el cas concret del plegament del LCI es va observar que l'estabilitat conformacional dels intermediaris de plegament no està rigorosament lligada al nombre de ponts disulfur. Es va purificar un intermediari de III ponts disulfur que s'acumula durant el plegament del LCI i que s'havia postulat com a possible trampa cinètica (Salamanca i col, resultats no publicats) i es va comparar la seva estabilitat amb la de les formes *scrambled* que en contenen quatre. El fet que ambdós intermediaris presentin una estabilitat semblant, ens indica que la forma amb III-disulfurs presenta una sèrie d'interaccions no-covalents, que podrien ser similars a les que es donen en la forma nativa, que sobre-estabilitzen la molècula i compensen la manca d'un pont disulfur. La major estabilitat d'aquest intermediari de III ponts disulfur, probablement, és la responsable de la seva acumulació respecte la resta d'intermediaris de III ponts.

El plegament del LCI és similar al de l'EGF: ambdues molècules presenten al llarg del seu plegament l'acumulació d'un intermediari que conté un pont disulfur menys que la forma nativa; en el cas de l'EGF s'ha demostrat que els dos ponts que presenta aquest intermediari són nadius (Chang, 2000). Un cas extrem d'aquest tipus de plegament és el BPTI, que presenta de forma molt majoritària intermediaris de plegament que contenen exclusivament ponts disulfur nadius (intermediaris de I i II ponts disulfur) (Weissman i Kim, 1991). Això s'ha relacionat amb un plegament dirigit per la formació d'estructures secundàries i no pels ponts disulfur. L'adquisició d'interaccions no-covalents de tipus-natiu des d'un inici, forçaria a la formació dels ponts disulfurs presents en la forma nativa. En el cas del BPTI fins i tot s'ha observat que en una variant on dues cys han sigut substituïdes per serina (un dels ponts disulfur no es pot formar), la molècula aconsegueix plegar-se i adoptar una estructura de tipus natiu activa (Goldenberg, 1988).

En el cas de l'EGF i del LCI, no tots els intermediaris presents en el plegament contenen ponts nadius i a més, també s'observa la presència de formes *scrambled*. Això implica que tot i que s'acumuli de manera predominant un intermediari que és força estable, després aquest intermediari es reorganitza i trenca els seus ponts disulfurs per passar a les formes *scrambled* i finalment aquestes tornen a reorganitzar els seus ponts per donar lloc a la forma nativa (Chang i col, 1997). Així doncs, això sembla indicar que aquests intermediaris més estables presents en el plegament *in vitro* del LCI i l'EGF, probablement contenen una estructura que no és semblant a la de la forma

nativa; ja que si presentessin una estructura similar a la forma nativa i tinguessin tots els ponts natius formats, excepte un, el camí més probable per arribar a la forma nativa seria mitjançant la formació directa de l'últim pont, com passa en el cas del BPTI, i no mitjançant la reorganització total de la molècula per donar lloc a formes *scrambled*.

Així doncs, aquests estudis ens permeten determinar que l'estabilitat conformacional dels intermediaris de plegament es correlaciona amb el progrés del procés de plegament estudiat per RP-HPLC, però no està necessàriament lligada amb el nombre de ponts disulfurs dels intermediaris de plegament. A més, les dades d'estabilitat conformacional també ens permeten quantificar la contribució de les interaccions covalents i no-covalents en l'estabilitat final de la forma nativa.

VIII-Conclusions

1- S'ha clonat i expressat el PCI madur en el llevat metil.lotròfic *Pichia pastoris* mitjançant l'ús del vector pPIC9 que és d'expressió extracel.lular. S'ha observat que els clons de fenotip mut^s presenten produccions generalment més elevades que els de fenotip mut⁺. L'optimització de l'expressió del PCI en flascons de 2L ha permès l'obtenció de 5,5 mg/L de proteïna finals.

2- La construcció gènica d'una cua d'histidines fusionada a l'extrem N terminal del PCI, clonada en el vector pIN3OmpAIII, s'expressa normalment, però és degradada al periplasma d'*Escherichia coli* per acció de les proteases, que eliminen la cua d'histidines. Això impossibilita l'ús d'aquest sistema d'expressió per a facilitar la purificació de la proteïna mitjançant una columna d'afinitat de Ni²⁺.

3- S'ha clonat el PCI madur en el vector pBAT4-OmpA, observant-se que la seva expressió en *E.coli* és extracel.lular. Posteriorment s'ha optimitzat la seva expressió en les cèl.lules d'*E.coli* BL21(DE3), la qual cosa ha conduït a una producció final en flascó de 15 mg/L de proteïna, 10 cops superior al sistema de partida.

4- S'han clonat el ProPCI i el ProNtPCI en el vector pBAT4 i s'ha optimitzat la seva expressió en *E.coli*. La caracterització de l'activitat inhibidora del ProNtPCI ha mostrat que és similar a la de la molècula madura.

5- Els estudis de replegament oxidatiu realitzats amb el PCI, el ProPCI, el ProNtPCI, i el PCI amb la pro-seqüència N-terminal en *trans*, han demostrat que les pro-seqüències N i C-terminals del PCI no presenten cap influència positiva sobre la cinètica i/o eficiència de plegament de la molècula madura *in vitro*.

6- Els estudis de producció recombinant en *E.coli* duts a terme amb els mutants Y37G PCI, Δ3 PCI, G35P&P56G PCI, Y37 ProNtPCI, Δ3 ProNtPCI, G35P&P36G ProNtPCI i ProNtPCI han mostrat que la presència de la pro-seqüència N-terminal no afecta ni el nivell d'expressió ni l'eficiència de plegament del PCI o dels mutants d'aquest *in vivo*, en *E.coli*.

7- La molècula ProNtPCI ha estat sotmesa a estudis estructurals de dicroïsmes circulars, bescanvi H/D seguit per espectrometria de masses MALDI-TOF i ressonància magnètica nuclear. De forma conjunta, aquests estudis suggereixen que l'extensió N-terminal del PCI està estructurada.

Conclusions

8- Els estudis de bescanvi D/H seguits per espectrometria de masses MALDI-TOF de mostres de plegament de la hirudina i del LCI han permès caracteritzar l'adquisició d'estabilitat conformacional durant el plegament d'aquestes proteïnes.

9- La determinació de l'estabilitat d'intermediaris de plegament de la hirudina i del LCI ha corroborat que ambdues molècules segueixen models de plegament diferents. En la hirudina l'estabilitat es correlaciona amb el nombre de ponts disulfur dels intermediaris de plegament. En canvi, en el LCI, tant les interaccions no-covalents com el nombre de ponts disulfur determinen l'estabilitat dels seus intermediaris de plegament.

10-L'estabilitat de les formes *scrambled* de la hirudina i del LCI és d'aproximadament un 60% respecte la forma nativa. Per tant, la formació dels ponts disulfur no és suficient per a l'estabilització d'aquestes proteïnes, sinó que les interaccions no-covalents presents en la forma que presenta l'aparellament de disulfurs correcte és responsable, en gran part, de la seva estabilitat.

IX-Bibliografia

A

- Andrews BA and Asenjo JA; (1996) A tightly regulated high level expression vector that utilizes a thermosensitive lac repressor: production of the human T cell receptor U beta 5.3 in *Escherichia coli*; *Gene*; 182:101-109
- Anfinsen CB, Haber E, Sela M and White FHJ; (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain; *Proc Natl Acad Sci USA*; 47:1309-1314

B

- Bajzar L, Manuel R and Nesheim ME; (1995) Purification and characterization of TAFI, a thrombin-activable fibrinolysis inhibitor; *J Biol Chem*; 270:14477-14484
- Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J and Turk V; (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases; *Embo J*; 7:2593-2599
- Bode W and Huber R; (1992) Natural proteinase inhibitors and their interaction with proteinases. *Eur J Biochem*; 204:433-451
- Bode W and Huber R; (2000) Structural basis of the endoproteinase-protein inhibitor interaction; *BBA*; 1477:241-252
- Brankamp RG, Sreekrishna D, Smith PL, Blankenship DT and Cardin AD; (1995) Expression of a synthetic gene encoding the anticoagulant-antimetastatic protein ghilanten by the methylotrophic yeast, *Pichia pastoris*; *Protein Expression Purif*; 6:813-820
- Brinkmann U, Mattes RE and Buckel P; (1989) High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product; *Gene*; 85:109-114

C

- Chang JY, Li L and Lai P-H; (2001) A major kinetic trap for the oxidative folding of human epidermal growth factor; *J Biol Chem*; 276:4845-4852
- Chang JY, Li L and Bulychiev A; (2000a) The underlying mechanism for the diversity of disulfide folding pathways; *J Biol Chem*; 275:8287-8289
- Chang J-Y, Li L, Canals F and Aviles FX; (2000b) The unfolding pathway and conformational stability of potato carboxypeptidase inhibitor; *J Biol Chem*; 275:14205-14211

- Chang JY; (1999) Quantitative analysis of the composition of the native and scrambled ribonuclease A; *Anal Biochem*; 268:147-150
- Chang JY; (1997) A two-stage mechanism for the reductive unfolding of disulfide-containing proteins; *J Biol Chem*; 272:69-75
- Chang JY; (1996) The disulfide folding pathway of tick anticoagulant peptide (TAP), a kunitz-type inhibitor structurally homologous to BPTI; *Biochemistry*; 35:11702-11709
- Chang JY, Schindler P, Ramseier U and Lai PH; (1995) The disulfide folding pathway of human epidermal growth factor; *J Biol Chem*; 270:9207-90216
- Chang J-Y, Canals F, Schindler P, Querol E and Aviles FX; (1994) The disulfide folding pathway of potato carboxypeptidase inhibitor; *J Biol Chem*; 269: 22087-22090
- Chatrenet B and Chang J-Y; (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments; *J Biol Chem*; 268:20988-20996
- Cereghino JL and Cregg JM; (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*; *FEMA Microbiol Rev*; 24:45-66
- Chen L-H, Emory SA, Bricker AL, Bouvet P and Belasco JG; (1991) Structure and function of a bacterial mRNA stabilizer: Analysis of the 5' untranslated region of *OmpA* mRNA; *J Bacteriol*; 173:4578-4586
- Clore GM, Gronenborn AM, Nilges M and Ryan CA; (1987) Three-dimensional structure of potato carboxypeptidase inhibitor in solution. A study using nuclear magnetic resonance, distance geometry, and restrained molecular dynamics; *Biochemistry*; 26:8012-8023
- Corrales FJ and Fersht AR; (1996) Toward a mechanism for GroEL.GroES chaperone activity: an ATPase-gated and -pulsed folding and annealing cage; *Proc Natl Acad Sci USA*; 93:4509-4512
- Cregg JM, Cereghino JL, Shi J and Higgins DR; (2000) Recombinant protein expression in *Pichia pastoris*; *Mol Biotechnol*; 16(1):23-52
- Creighton TE, Bagley CJ, Cooper L, Darby NJ, Freedman RB, Kemmink J and Sheikh A; (1993) 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
- Creighton TE; (1992) The disulfide folding pathway of BPTI; *Science*; 256:111-114
- Creighton TE; (1986) Detection of folding intermediates using urea-gradient electrophoresis; *Methods Enzymol*; 131:83-106
- Creighton TE; (1978) Experimental studies of protein folding and unfolding; *Prog Biophys Mol Biol*; 33:231-297

D

Davies DR; (1990) The structure and function of the aspartic proteinases; *Annu Rev Biophys Biophys Chem*; 19:189-215

E

Englander SW, Mayne L, Bai Y and Sosnick TR; (1997) Hydrogen exchange: the modern legacy of Linderstrom-Lang; *Protein Sci*; 6:1101-1109

Englander SW, Sosnick TR, Englander JJ and Mayne L; (1996) Mechanisms and uses of hydrogen exchange; *Curr Op Struct Biol*; 6:18-23

F

Fenn JB, Mann M, Meng CK, Wong SF and Whitehouse CM; (1989) Electrospray ionization for mass spectrometry of large biomolecules; *Science*; 246:64-71

G

Ghaemmaghami S and Oas TG; (2001) Quantitative protein stability measurement in vivo; *Nat Struct Biol*; 8:879-882

Ghaemmaghami S, Fitzgerald MC and Oas TG; (2000) A quantitative, high-throughput screen for protein stability; *Proc Natl Acad Sci USA*; 97:8296-8301

Goldenberg DP; (1988) Kinetic analysis of the folding and unfolding of a mutant form of bovine pancreatic trypsin inhibitor lacking the cysteine-14 and -38 thiols; *Biochemistry*; 27:2481-2489

H

Hannig G and Makrides SC; (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*; *Tibtech*; 16:54-60

Harvey RP, Degryse E, Stefani L, Schamber F, Cazenave JP, Courtney M, Tolstoshev P and Lecocq JP; (1986) Cloning and expression of a cDNA coding for the anticoagulant hirudin from the bloodsucking leech, *Hirudo medicinalis*; *Proc Natl Acad Sci USA*; 83:1084-1088.

- Hass GM and Hermudson MA; (1981) Amino acid sequence of a carboxypeptidase inhibitor from tomato fruit; *Biochemistry*; 20:2256-2260
- Hass GM and Ryan CA; (1981) Carboxypeptidase inhibitor from potatoes; *Methods Enzymol*; 80:778-791
- Hass GM and Derr JE; (1979a) Purification and characterization of the carboxypeptidase isoinhibitors from potatoes; *Plant Physiol*; 64:1022-1028
- Hass GM and Derr JE; (1979b) Distribution of carboxypeptidase isoinhibitors in the potato plant; *Plant Physiol*; 64:1029-1031
- Hass GM, Ako H, Grahn DG and Neurath H; (1976) Carboxypeptidase inhibitor from potatoes. The effects of chemical modifications on inhibitory activity; *Biochemistry*; 15:93-100
- Hewinson RG and Russell WP; (1993) Processing and secretion by *Escherichia coli* of a recombinant form of the immunogenic protein MPB70 of *Mycobacterium bovis*; *J Gen Microbiol*; 139:1253-1259
- Hollander-Czytko H, Andersen JK and Ryan CA; (1985) Vacuolar localization of wound-induced carboxypeptidase inhibitor in potato leaves; *Plant Physiol*; 78:76-79
- Hillenkamp F, Karas M, Beavis RC and Chait BJ; (1991) Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers; *Anal Chem*; 63:1193-1203
- Homandberg GA, Litwiller RD and Peanasky RJ; (1989) Carboxypeptidase inhibitors from *Ascaris suum*: the primary structure; *Arch Biochem Biophys*; 270:153-161
- Huyghues-Despointes BM, Scholtz JM and Pace CN; (1999) Protein conformational stabilities can be determined from hydrogen exchange rates; *Nat Struct Biol*; 6:910-912

I

- Ikai A and Tanford C; (1971) Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins; *Nature*; 230:100-102

K

- Khan AR and James MN; (1998) Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes; *Protein Sci*; 7:815-836
- Kim KS, Fuchs JA and Woodward CK; (1993) Hydrogen exchange identifies native-state motional domains important in protein folding; *Biochemistry*; 32:9600-9608

Konishi Y, Ooi T and Scheraga HA; (1982) Regeneration of ribonuclease A from reduced protein. Rate-limiting steps; *Biochemistry*; 21: 4734-4740

L

Lanzer M and Bujard H; (1988) Promoters largely determine the efficiency of repressor action; *Proc Natl Acad Sci USA*; 85:8973-8977

Laroche Y, Storme V, De Meutter J, Messens J and Lauwereys M; (1994) High level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, *Pichia pastoris*; *Biotechnology*; 12:1119-1124

La Vallie ER, Di Blasio EA, Kovacic S, Grant KL, Schendel PF and McCoy JM; (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the cytoplasm; *Biotechnology*; 11:187-193

La Vallie ER and McCoy JM; (1995) Gene fusion expression systems in *Escherichia coli*; *Curr Opin Biotechnol*; 6:501-506

Lenarcic B and Turk V; (1999) Thyroglobulin type-I domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D; *J Biol Chem*; 274:563-566

Levinthal C; (1968) Are there pathways for protein folding?; *J Chim Phys*; 65:44-45

Lin SL and Nussinov R; (1995) A disulphide-reinforced structural scaffold shared by small proteins with diverse functions; *Nat Struct Biol*; 2:835-837

M

Makrides SC; (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*; *Microbiol Rev*; 60:512-538

Mandell JC, Falick AM and Komives EA; (1998) Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry; *Anal Chem*; 70:3987-3995

Mann M, Meng CK and Fenn JB; (1989) Interpreting mass spectra of multiply charged ions; *Anal Chem*; 61:1702-1708

Manosroi J, Tayapiwatana C, Götz F, Werner RG, and Manosroi A; (2001) Secretion of active recombinant human tissue plasminogen activator derivatives in *Escherichia coli*; *Appl Environ Microbiol*; 67: 2657-2664

Margis R, Reis EM and Villeret V; (1998) Structural and phylogenetic relationships among plant and animal cystatins; *Arch Biochem Biophys*; 359:24-30

Miranker A, Robinson CV, Radford SE, Aplin RT and Dobson CM; (1993) Detection of transient protein folding populations by mass spectrometry; *Science*; 262:896-899

Murby M, Uhlén M and Stahl S; (1996) Upstream strategies to minimize proteolytic degradation upon recombinant production in *Escherichia coli*; *Protein Expression Purif*; 7:129-136

Molina MA, Aviles FX and Querol E; (1992) Expression of a synthetic gene encoding potato carboxypeptidase inhibitor using a bacterial secretion vector; *Gene*; 116:129-138

N

Nagase H; (1996) Matrix metalloproteinases; *Zinc Metalloproteases in health and disease*; 153-204

Neira JL and Fersht AR; (1999) Exploring the folding funnel of a polypeptide chain by biophysical studies on protein fragments; *J Mol Biol*; 285:1309-1333

Neira JL, Itzhaki LS, Ladurner AG, Davis B, de Prat Gay G and Fersht AR; (1997) Following cooperative formation of secondary and tertiary structure in a single protein module; *J Mol Biol*; 268:185-197

Normant E, Martres MP, Schwartz JC and Gros C; (1995); Purification, cDNA cloning, functional expression, and characterization of a 26-kDa endogenous mammalian carboxypeptidase inhibitor; *Proc Natl Acad Sci USA*; 92:12225-12229

O

Olins PO and Lee SC; (1993) Recent advances in heterologous gene expression in *Escherichia coli*; *Curr Opin Biotechnol*; 4:520-525

P

Peranen J, Rikkonen M, Hyvonen M and Kaariainen L; (1996) T7 vectors with a modified *T7lac* Promoter for expression of proteins in *Escherichia coli*; *Anal Biochem*; 236:371-373

Poole ES, Brown CM and Tate WP; (1995) The identity of the base following the stop codon determines the efficiency of in vivo translational termination in *Escherichia coli*; *Embo J*; 14:151-158

Price-Carter M, Gray WR and Goldenberg DP; (1996a) Folding of ω -conotoxins.1.Efficient disulfide-coupled folding of mature sequence in vitro; *Biochemistry*; 35:15537-15546

Price-Carter M, Gray WR and Goldenberg DP; (1996b) Folding of ω -conotoxins.1. Influence of precursor sequences and protein disulfide isomerase; *Biochemistry*; 35:15547-15557.

R

Radford SE, Dobson CM and Evans PA; (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways; *Nature*; 358:302-307

Rawlings ND, O'Brien E and Barrett AJ; (2002) MEROPS: the protease database; *Nucleic Acids Res*; 30:343-346

Rees DD and Lipscomb WN; (1982) Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution; *J Mol Biol*; 160:475-498

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

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

Roder H, Elove GA and Englander SW; (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR; *Nature*; 335:700-704.

Robinson CV, Gross M, Eyles SJ, Ewbank JJ, Mayhew M, Hartl FV, Dobson CM and Radford SE; (1994) Conformation of GroEL-bound α -lactalbumin probed by mass spectrometry; *Nature*; 372:646-651

Ryan CA; (1989) Proteinase inhibitor gene families: strategies for transformation to improve plant defenses against herbivores; *Bioassays*; 10:20-24.

Rydel TJ, Tulinsky A, Bode W and Huber R; (1991) Refined structure of the hirudin-thrombin complex; *J Mol Biol*; 221:583-601

Rosenfeld SA, Nadeau D, Tirado J, Hollis GF, Knabb RM and Jia S; (1996) Production and purification of recombinant hirudin expressed in the methylotrophic yeast *Pichia pastoris*; *Protein Expression Purif*; 8:476-482

S

Shinde U and Inouye M; (2000) Intramolecular chaperone: polypeptide extensions that modulate protein folding; *Cell Developmental Biology*; 11:35-44

Somerville JE, Goshorn SC, Fell HP, and Darveau RP; (1994) Bacterial aspects associated with the expression of a single-chain antibody fragment in *Escherichia coli*; *Appl Microb Biotechnol*; 42: 595-603

Soritopoulou G, Anisowicz A and Sager R; (1997) Identification, cloning and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer; *J Biol Chem*; 272:903-910

Suzuki K and Sorimachi H; (1998) A novel aspect of calpain activation; *FEBS Lett*; 433:1-4

Swartz JR; (2001) Advances in *Escherichia coli* production of therapeutic proteins; *Curr Op Biotech*; 12:195-201

T

Tobias JW, Shrader TE, Rocap G and Varshavsky A; (1991) The N-end rule in bacteria; *Science*; 254:1374-1377

Turk V and Bode W; (1991) The cystatins: protein inhibitors of cysteine proteinases; *FEBS Lett*; 285:213-219

U

Udgaonkar JB and Baldwin RL; (1988) NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A; *Nature*; 335:694-699

V

Vestal ML; Juhasz P and Martin SA; (1995) Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry; *Rapid Commun Mass Spectrom*; 9:1044-1050

Villanueva J, (2002) *J Mass Spectr*; manuscrit acceptat per a publicació

Villanueva J, Canals F, Querol E and Aviles FX; (2001) Monitoring the expression and purification of recombinant proteins by MALDI-TOF mass spectrometry; *Enz micr tech*; 29:99-103

Villanueva J, Canals F, Villegas V, Querol E and Aviles FX; (2000) Hydrogen exchange monitored by MALDI-TOF mass spectrometry for rapid characterization of the stability and conformation of proteins; *FEBS Lett*; 472:27-33

Villanueva J, Canals F, Prat S, Ludevid D, Querol E and Aviles FX; (1998) Characterization of the wound-induced metallocarboxypeptidase inhibitor from potatoe. cDNA sequence, induction

of gene expression, subcellular immunolocalization and potential roles of the C terminal propeptide; *FEBS Lett*; 440:175-182

W

Weissman JS and Kim PS; (1992) The pro region of BPTI facilitates folding; *Cell* 71:841-851

Weissman JS and Kim PS; (1991) Reexamination of the folding of BPTI: Predominance of native intermediates; *Science*; 253:1386-1393

Wilkins TA and Raikhel NV; (1989) Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos; *Plant Cell*; 1:541-549

Woodward C; (1993) Is the slow exchange core the protein folding core?; *TIBS*; 18:359-360

Wright HT; (1996) The structural puzzle of how serpin serine proteinase inhibitors work; *Bioessays*; 18:453-464

Y

Yates Jr III; (1998) Mass spectrometry and the age of the proteome; *J Mass Spectrom*; 33:1-19

X-Apèndix

Anglicismes

acid-trapping	captura en medi àcid
clon(ing)	clon(atge)*
disulfide-quenching	captura de disulfurs
disulfide-reshuffling	reorganització de disulfurs
DNA spacer	ADN separador*
(un)folding	(des)plegament*
(cat/an)ionic exchange	bescanvi (cat/an)iònic*
labeling	marcatge*
loop	llaç*
screening	triatge*
stop codon	codó de terminació*
up/downstream	en direcció 5'/3'*
wild-type	tipus salvatge*

*Traduccions obtingudes a partir del centre de terminologia TERMCAT (ConSORCI integrat per la Generalitat de Catalunya, l'Institut d'Estudis Catalans i el Consorci per a la Normalització Lingüística).

<http://www.termcat.es/>