Overexpression of Human Procarboxypeptidase A2 in Pichia pastoris and Detailed Characterization of Its Activation Pathway*

David Reverter, Salvador Ventura, Virtudes Villegas, Josep Vendrell, and Francesc X. Avilés

From the Departament de Bioquímica i Biologia Molecular, Unitat de Ciències and the Institut de Biologia Fonamental, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

The cDNA of human procarboxypeptidase A2 has been overexpressed in the methylotriprotrophic yeast Pichia pastoris and secreted into the culture medium by means of the α-mating factor signal sequence, yielding a major protein of identical size and N-terminal sequence as the wild-type form. Two other forms containing the proenzyme have also been overexpressed: one of them resulted from an incomplete processing of the signal peptide, whereas the other was a glycosylated derivative. Recombinant procarboxypeptidase A2 was purified to homogeneity, and it was shown that its mature active form displays functional properties similar to those of the enzyme directly isolated from human pancreas. The overall yield was ~250 mg of proenzyme or 180 mg of mature enzyme/liter of cell culture. The proteolysis-promoted activation process of the recombinant proenzyme has been studied in detail. During maturation by trypsin, the increase in activity of the enzyme is a rapid and monotonic event, which reflects the rate of the proteolytic release of the inhibitory pro-segment and the weaker nature of its interactions with the enzyme moiety compared with procarboxypeptidases of the A1 type. Three main forms of the prosegment (96, 94, and 92 amino acids), with no inhibitory capability in the severe state, and a single mature carboxypeptidase A2 are produced during this process. No further proteolysis of these pro-segments by the generated carboxypeptidase A2 occurs, in contrast with observations made in other procarboxypeptidases (A1 and B). This differential behavior is a result of the extreme specificity of carboxypeptidase A2.

Pancreatic carboxypeptidases (CPs) are digestive metalloenzymes involved in the hydrolysis of alimentary proteins and peptides from their C-terminal end. Their participation as proenzymes in the digestive cascade (promoted by limited proteolysis) is a well documented process (1–3). Also, their specificity classification between the A forms (CPA, with preference for apolar C-terminal residues) and the B forms (CBP, with preference for basic C-terminal residues) and the tertiary structures of both forms are well known (4). In recent years, there has been an increasing interest in the study of the synthesis, storage, activation, and three-dimensional structure of procarboxypeptidases (pro-CPs), the precursors of such enzymes (4–6).

The classification of metalloprocarboxypeptidases has been widened in the last few years with reports about new non-digestive pancreatic-like carboxypeptidases in different extra-pancreatic tissues and biological fluids, with the same evolutionary ancestors as pancreatic carboxypeptidases (4, 7–10). Additionally, the traditional classification of pancreatic carboxypeptidases and their zymogens into the A and B forms has been expanded with the identification of the A1 and A2 isoforms in rat and humans (11, 12). CPA1 and CPA2 differ in specificity for peptide substrates: the former (assignable to the traditional A form) shows a wider preference for aliphatic and aromatic residues, whereas the latter is more restrictive for aromatic residues; this reflects significant differences in the specificity pocket of the enzymes (13). Recently, we have reported the cloning and sequence analysis of the human pro-CPA2 cDNA as well as its three-dimensional modeling (14). CPA2 isoforms have also been reported in rat extra-pancreatic tissues such as brain, testis, and lung (15); these CPA2 isoforms are shorter and have a distinct role from the pancreatic isoform. The high sequence identity found between human pro-CPA2 and rat pro-CPA2 (89% homology) as compared with human pro-CPA1 (64% homology) corroborates the proposal that locates the appearance of the two isoforms by gene duplication before speciation of mammals (11).

Comparison of the prodomain structures in the family of pancreatic proenzymes shows close similarities in conformation between the A1 and A2 forms in regions assumed to be critical for their inhibition and proteolytic activation (6, 14) and significant differences from the corresponding regions in the B form (5). Accordingly, the A2 proenzyme could be expected to show a bimodal and slow proteolytic activation behavior, as previously reported for the A1 form (16), and to differ from the monotonic and quick activation behavior found for the B proenzyme (17). However, earlier proteolytic activation experiments carried out on natural pro-CPA2 isolated from human or rat pancreas (12, 18) do not fit with these expectations and assumptions. Therefore, this is an issue that requires clarification.

In this work, pro-CPA2 has been overexpressed in Pichia pastoris to produce the protein in quantities amenable to the study of the structural and functional determinants of its behavior and activation. The methylotriprotrophic yeast P. pastoris was chosen because of its high yield and capacity of secreting heterologous proteins when linked to the appropriate secretion signal (19). The development of the system reported here to obtain large quantities of fully activable human pro-CPA2 should facilitate not only the characterization of this form, but...
probably also that of other structurally related forms to which the same procedure could be applied. It could also facilitate its potential biotechnological use, such as the large-scale production of carboxypeptidases able to act as activators of antitumor prodrugs (20, 21). The efficient expression of human pro-CPA2 in *P. pastoris* has allowed us to investigate the different events in the proteolytic activation and processing of this proenzyme in detail and to compare them with the processes described in other pancreatic carboxypeptidases. An overall maturation scheme of such zymogenes emerges from this study.

**EXPERIMENTAL PROCEDURES**

**Materials**—Restriction endonucleases, T4 DNA ligase, Vent polymerase, deoxynucleotide stocks, and N-glycosidase F were purchased from Boehringer Mannheim. Salts and media for *Escherichia coli* and *P. pastoris* growth were purchased from Difco. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanine chloromethyl ketone) was from Worthington. Trifluoroacetic acid, N-p-tosyl-L-lysine chloromethyl ketone (TLCK), and N-(3-(2-furyl)acryloyl)-L-phenylalanine (FAPP) were from Sigma.

Electrophoretic studies were carried out in a Bio-Rad Mini-Protean system. HPLC studies were carried out in a Waters chromatograph.

**Transformation and Selection of the Productive Clones**—To linearize the corresponding *P. pastoris* pPIC9-pro-CPA2 clone was digested with Merpol as the vector for selection of the productive clones. After digestion enzymatic digestion of the polymerase chain reaction product, the cDNA was cloned and subcloned into M13 phBluescript to confirm the entire sequence and the homologous recombination with the pBluescript-pro-CPA2 clone was digested and ligated to the pPIC9 vector to introduce an **EXPERIMENTAL PROCEDURES**

**Materials**—Restriction endonucleases, T4 DNA ligase, Vent polymerase, deoxynucleotide stocks, and N-glycosidase F were purchased from Boehringer Mannheim. Salts and media for *Escherichia coli* and *P. pastoris* growth were purchased from Difco. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanine chloromethyl ketone) was from Worthington. Trifluoroacetic acid, N-p-tosyl-L-lysine chloromethyl ketone (TLCK), and N-(3-(2-furyl)acryloyl)-L-phenylalanine (FAPP) were from Sigma.

Electrophoretic studies were carried out in a Bio-Rad Mini-Protean system. HPLC studies were carried out in a Waters chromatograph.

**Transformation and Selection of the Productive Clones**—To linearize the corresponding *P. pastoris* pPIC9-pro-CPA2 clone was digested with Merpol as the vector for selection of the productive clones. After digestion enzymatic digestion of the polymerase chain reaction product, the cDNA was cloned and subcloned into M13 phBluescript to confirm the entire sequence and the homologous recombination with the pBluescript-pro-CPA2 clone was digested and ligated to the pPIC9 vector to introduce an

**Experimental Procedures**

**Materials**—Restriction endonucleases, T4 DNA ligase, Vent polymerase, deoxynucleotide stocks, and N-glycosidase F were purchased from Boehringer Mannheim. Salts and media for *Escherichia coli* and *P. pastoris* growth were purchased from Difco. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanine chloromethyl ketone) was from Worthington. Trifluoroacetic acid, N-p-tosyl-L-lysine chloromethyl ketone (TLCK), and N-(3-(2-furyl)acryloyl)-L-phenylalanine (FAPP) were from Sigma.

Electrophoretic studies were carried out in a Bio-Rad Mini-Protean system. HPLC studies were carried out in a Waters chromatograph.

**Transformation and Selection of the Productive Clones**—To linearize the corresponding *P. pastoris* pPIC9-pro-CPA2 clone was digested with Merpol as the vector for selection of the productive clones. After digestion enzymatic digestion of the polymerase chain reaction product, the cDNA was cloned and subcloned into M13 phBluescript to confirm the entire sequence and the homologous recombination with the pBluescript-pro-CPA2 clone was digested and ligated to the pPIC9 vector to introduce an

**Experimental Procedures**

**Materials**—Restriction endonucleases, T4 DNA ligase, Vent polymerase, deoxynucleotide stocks, and N-glycosidase F were purchased from Boehringer Mannheim. Salts and media for *Escherichia coli* and *P. pastoris* growth were purchased from Difco. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanine chloromethyl ketone) was from Worthington. Trifluoroacetic acid, N-p-tosyl-L-lysine chloromethyl ketone (TLCK), and N-(3-(2-furyl)acryloyl)-L-phenylalanine (FAPP) were from Sigma.

Electrophoretic studies were carried out in a Bio-Rad Mini-Protean system. HPLC studies were carried out in a Waters chromatograph.

**Transformation and Selection of the Productive Clones**—To linearize the corresponding *P. pastoris* pPIC9-pro-CPA2 clone was digested with Merpol as the vector for selection of the productive clones. After digestion enzymatic digestion of the polymerase chain reaction product, the cDNA was cloned and subcloned into M13 phBluescript to confirm the entire sequence and the homologous recombination with the pBluescript-pro-CPA2 clone was digested and ligated to the pPIC9 vector to introduce an

**Experimental Procedures**

**Materials**—Restriction endonucleases, T4 DNA ligase, Vent polymerase, deoxynucleotide stocks, and N-glycosidase F were purchased from Boehringer Mannheim. Salts and media for *Escherichia coli* and *P. pastoris* growth were purchased from Difco. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanine chloromethyl ketone) was from Worthington. Trifluoroacetic acid, N-p-tosyl-L-lysine chloromethyl ketone (TLCK), and N-(3-(2-furyl)acryloyl)-L-phenylalanine (FAPP) were from Sigma.

Electrophoretic studies were carried out in a Bio-Rad Mini-Protean system. HPLC studies were carried out in a Waters chromatograph.

**Transformation and Selection of the Productive Clones**—To linearize the corresponding *P. pastoris* pPIC9-pro-CPA2 clone was digested with Merpol as the vector for selection of the productive clones. After digestion enzymatic digestion of the polymerase chain reaction product, the cDNA was cloned and subcloned into M13 phBluescript to confirm the entire sequence and the homologous recombination with the pBluescript-pro-CPA2 clone was digested and ligated to the pPIC9 vector to introduce an

**Experimental Procedures**

**Materials**—Restriction endonucleases, T4 DNA ligase, Vent polymerase, deoxynucleotide stocks, and N-glycosidase F were purchased from Boehringer Mannheim. Salts and media for *Escherichia coli* and *P. pastoris* growth were purchased from Difco. The *P. pastoris* expression kit was purchased from Invitrogen. Trypsin (treated with tosylphenylalanine chloromethyl ketone) was from Worthington. Trifluoroacetic acid, N-p-tosyl-L-lysine chloromethyl ketone (TLCK), and N-(3-(2-furyl)acryloyl)-L-phenylalanine (FAPP) were from Sigma.
without histidine supplementation. Nearly 20% of the His+ clones from the transformation showed reduced growth on methanol as the sole carbon source (slow growing, Mut-), indicating the integration of the expression cassette into the AOX1 gene.

**Overexpression and Purification of Recombinant Human Pro-CPA2**—The clones transformed with the α-MF-prepro-CPA2 fusion product secreted a dominant 45-kDa protein in the *P. pastoris* supernatant upon induction by methanol. Those transformed with pHIL-D2-prepro-CPA2 did not secrete any protein into the medium. The 45-kDa protein was found to correspond to human pro-CPA2 by Western blot analysis. Upon induction with 0.5% methanol, the Mut+ phenotypes expressed more protein than the Mut- phenotypes. However, a 3-fold increase in protein secretion was observed in the Mut+ phenotypes when induction was assayed with 5% methanol (data not shown). One of these high productivity clones was selected for large-scale production of human pro-CPA2.

Analysis of the production of human pro-CPA2 in the above system by SDS-PAGE and Western blotting is shown in Fig. 1. After 2 days of induction by methanol, a strong band of the proenzyme and a faint band corresponding to the 34-kDa active form were detected in the intracellular soluble part of the culture. This observation is in accordance with previous reports about the existence of proteases with trypsin-like activity in *P. pastoris* (25). After 30 h of methanol induction, in addition to the dominant 45-kDa pro-CPA2 protein, a smear of high relative molecular mass, immunodetected by human procarboxypeptidase A2 (Fig. 2), was visible (Fig. 1tive molecular mass, immunodetected by human procarboxypeptidase antiserum, was visible (Fig. 1tive molecular mass, immunodetected by human procarboxypeptidase antiserum, was visible (Fig. 1)

A chromatographic approach was developed to separate the differently processed forms of human recombinant pro-CPA2 and to isolate the native form. It was based upon a two-step purification scheme, with atmospheric hydrophobic interaction chromatography on a butyl column applied first, followed by fast protein liquid anion-exchange chromatography to separate the different forms (see “Experimental Procedures”). Starting from a cell culture of a Mut+ clone that achieved a cell density of 20 g of cells (dry weight)/liter, an overall yield of 250 mg of total pure pro-CPA2 was normally obtained. The correctly processed pro-CPA2 form was generally the principal one, with yields varying from 80 to 200 mg/liter, although in certain cultures, the glycosylated and incorrectly processed forms could account for as much as 60% of the purified material. In contrast, treatment of total pro-CPA2 with trypsin before the anion-exchange chromatography yielded ~180 mg of homogeneous, correctly processed CPA2/liter of cell culture. Both the purified recombinant pro-CPA2 and CPA2 forms were fully functional, with enzymatic properties similar to those previously reported for the natural forms isolated from human pancreas.

**FIG. 1. Production of recombinant human pro-CPA2 in *P. pastoris*.** Shown are the results from analysis of the production of the recombinant proenzyme throughout the course of fermentation by SDS-polyacrylamide gel electrophoresis. Parallel 12% polyacrylamide gels were directly stained with Coomassie Blue (A) or immunoblotted (B) and stained. The samples were as follows: Cg, cell extract; M, cell-free medium after 4-day growth in 1% glycerol; Cm, cell extract after 48 h of 0.5% methanol induction; M1, M2, and M3, cell-free medium after 6, 24, and 48 h of 0.5% methanol induction, respectively. The cell extract samples contained the protein solubilized after treating the cells with electrophoresis loading buffer (62 mM Tris, 10% glycerol, 2% SDS, 0.7% β-mercaptoethanol, and bromphenol blue) for 1 min at 100°C. In the case of the cell-free medium samples, a volume equivalent to 5 µl of the fermentation medium was loaded in the electrophoresis wells. A control sample (st) containing 1.5 µg of pro-CPA2 isolated from human pancreas was included.

**FIG. 2. Deglycosylation and activation with trypsin of the α-MF-prepro-CPA2 fusion protein.** SDS-8% polyacrylamide gel electrophoresis was used to analyze the different products containing pro-CPA2 purified from the *P. pastoris* culture. Samples were analyzed before (lanes 1, 3, and 5) and after (lanes 2, 4, and 6) trypsin treatment. Lanes 1 and 2, correctly processed recombinant human pro-CPA2; lanes 3 and 4, samples containing a glycosylated derivative of the α-MF-prepro-CPA2 fusion product; lanes 5 and 6, the α-MF-prepro-CPA2 fusion product after deglycosylation. 2 µg of CPA2 purified from correctly processed recombinant pro-CPA2 were loaded in lane 7 as a control.
pro-CPA2 indicates that trypsin is, by far, the most efficient and elastase (the major active endoproteolytic counterparts in proteases—Analysis of the action of trypsin, chymotrypsin A, activator (data not shown). At the pro-CPA2/endoprotease ratio
pro-CPA2 is denoted by follow-up of the activation process on Tricine/SDS-polyacrylamide gel. The electrophoretic band corresponding to the activation sequence of sites (dashed arrows) found by N-terminal sequencing of recombinant products are indicated with
activation of pancreatic procarboxypeptidases (3). Recombinant human pro-CPA2, at 1 mg/ml in 50 mM Tris-HCl and 0.01 mM ZnCl2, pH 8.0, was treated with trypsin at 0 °C and at a ratio of 400:1 (w/w). At controlled times, aliquots were withdrawn and analyzed for carboxypeptidase A2 appearance by activity measurements and SDS-PAGE. A, carboxypeptidase activity values obtained with 0.2 mM FAPP as substrate; units are expressed as μmol of substrate hydrolyzed per min/mg of protein. B, quantitation of the relative amounts of pro-CPA2 (■) and CPA2 (▲) by densitometry after SDS-PAGE and Coomassie Blue staining. The inset shows the electrophoretic follow-up of the activation process on Tricine/SDS-polyacrylamide gel. The electrophoretic band corresponding to the activation sequence of pro-CPA2 is denoted by asA2.

creas (12). Recombinant CPA2 and the enzyme isolated from pancreas showed the same maximum specific activity of ~85 μmol of FAPP substrate hydrolyzed per min/mg of protein.

Activation of Recombinant Human Pro-CPA2 with Different Proteases—Analysis of the action of trypsin, chymotrypsin A, and elastase (the major active endoproteolytic counterparts in pancreatic secretion) upon correctly processed recombinant pro-CPA2 indicates that trypsin is, by far, the most efficient activator (data not shown). At the pro-CPA2/endoprotease ratio used, 90% activation of the former was achieved by trypsin in ~10 min, whereas the same level of activation was only achieved after 125 min of treatment with elastase; in the latter time span, chymotrypsin was able to generate only 10% of active CPA2. This is in agreement with previous studies that reported trypsin as the main enzyme responsible for the activation of pancreatic procarboxypeptidases (3).

Trypsin Activation of Recombinant Human Pro-CPA2—The action of trypsin at 37 °C on recombinant pro-CPA2 at a 40:1 (w/w) pro-CPA2/trypsin ratio is excessively quick for detailed mechanistic analysis, in agreement with previous studies on natural human proenzymes (12). The study of the activation process was therefore carried out at 0 °C and at a 400:1 ratio. Under these conditions, the appearance of carboxypeptidase activity was measured with the synthetic substrate FAPP; it followed a quick and monotonic activation curve, which can be fitted to a pseudo first-order kinetics (Fig. 4A). Using Tricine/SDS-PAGE, the rapid proteolytic severing of the pro-segment, the concomitant appearance of CPA2, and a straight correlation of this proteolysis with the monotonic activity curve were observed (Fig. 4B). Nearly all pro-CPA2 was converted into its active form in 20 min, and a maximum value in the activity curve was achieved. The generated pro-segment appeared as a single band on Tricine/SDS-PAGE (Fig. 4B), although mass spectrometry analysis indicated heterogeneity for this species (see below).

The N-terminal sequence of the electrophoretic CPA2 band, obtained from different samples at different activation times and transferred to a polyvinylidene difluoride membrane, was always Ser-Gly-Asn-, a fact that indicates that the Arg96–Ser97 peptide bond is the first point of cleavage of the proenzyme by trypsin. From this result and from the increase in the intensity of the CPA2 band in parallel with the appearance of CPA2 activity in the medium and the corresponding decrease in the intensity of the pro-CPA2 band (Fig. 4B), it can be assumed that the cleavage of the Arg96–Ser97 peptide bond is sufficient to generate all the carboxypeptidase activity. As a consequence, it can be concluded that the severed pro-segment does not inhibit CPA2, in contrast to what has been previously reported for the pro-CPA1 forms (12, 16, 27).

To study the maturation process in more detail, the time course of degradation of the pro-segment was followed by reverse-phase HPLC. It was expected that this degradation would take place only at the C-terminal arginine-rich end of the pro-segment since N-terminal sequence analyses of the activation mixtures indicated the appearance of only two N termini throughout the process, one corresponding to the original proenzyme and one from the generated mature enzyme. The chromatographic analyses of the activation mixtures at different times are represented in Fig. 5, where the appearance and disappearance of three different protein fragments arising from
the proteolysis of the pro-segment (labeled α, β, and γ) can be observed. Mass spectrometry (MALDI-TOF) was used for the characterization of the molecular masses of these fragments (Table I). Their identification was facilitated by the knowledge of the sequence of pro-CPA2 and by analysis of peptides and amino acids that were released into the medium by the action of trypsin and CPA2. In particular, it is worth mentioning that two dipeptides (Glu-Arg and Arg-Arg) were detected by mass spectrometry at early and late activation times and that no significant free amino acids were detected by parallel amino acid analysis of the protein-free activation medium. All of this information was used to establish and confirm the sequence of the cleavages during the proteolytic processing.

According to the above information, during the tryptic maturation of pro-CPA2, the first cleavage is at the C terminus of Arg96, producing the α-fragment (residues 1–96), as shown in Fig. 6. Subsequently and while some proenzyme molecules are still intact, a rapid cleavage occurs at Arg84 of the α-fragment, giving rise to the β-fragment (residues 1–94) and to the release of a Glu-Arg dipeptide. The third cleavage (also attributable to trypsin, as are the former cleavages) occurs slowly at Arg92 of the β-fragment, giving rise to the γ-fragment (residues 1–92) and to the release of a second dipeptide, Arg-Arg. The lack of detection of free amino acids in the medium indicates the null action of generated CPA2 on the α-, β-, or γ-fragment. The γ-fragment is resistant to further proteolysis under these conditions. Only a single CPA2 species was detected throughout the proteolytic process by HPLC, N-terminal, and mass spectrometry analyses.

![Fig. 5. Analysis by reverse-phase HPLC of the fragments generated during tryptic activation of human pro-CPA2.](image)

![Fig. 6. Schematic representation of the cleavage points in the activation process of human pro-CPA2 promoted by trypsin.](image)

### Table I

| Activation time (min) | Pro-CPA2 molecular mass (expected: 44,928) Da | CPA2 molecular mass (expected: 37,787) Da | Pro-segment fragment molecular mass (expected: α-fragment, 11,158; β-fragment, 10,873; γ-fragment, 10,561) Da |
|----------------------|-----------------------------------------------|------------------------------------------|------------------------------------------------------------------------------------------------|
| 0                    | 44,906                                        | 37,765                                   | 11,157 (α-fragment)                                                                 |
| 0.5                  | 44,950                                        | 37,710                                   | 11,148 (α-fragment)                                                                 |
| 1.5                  | 44,919                                        | 37,710                                   | 10,860 (β-fragment)                                                                 |
| 8                    | 37,700                                        | 11,148                                   | 10,867 (β-fragment)                                                                 |
| 20                   | 37,760                                        | 10,855                                   | 10,561 (γ-fragment)                                                                 |
| 50                   | 37,750                                        | 10,855                                   | 10,564 (γ-fragment)                                                                 |
| 260                  | 37,770                                        | 10,855                                   | 10,573 (γ-fragment)                                                                 |

### DISCUSSION

One of the aims of this study was the development of a highly efficient recombinant expression system for a native and activable form of human procarboxypeptidase A2 in the P. pastoris heterologous system to study its proteolytic activation and maturation pathway in vitro. It is worth mentioning that pancreatic carboxypeptidases and their precursors are difficult to express in native and soluble forms in E. coli (28) and that reports about the expression of their precursor forms in eukaryotic cells (i.e. Saccharomyces cerevisiae) have indicated a moderate yield (<10 mg/liter) until now (21, 29). In fact, the primary production of the proenzymes by recombinant approaches followed by proteolytic activation is probably the best strategy to obtain active carboxypeptidases given that it takes advantage of the high folding capability of the pro-segment in heterologous systems (29, 30). The recent biotechnological interest in metallo-carboxypeptidases as prodrug activators for cancer therapy (21, 31, 32) makes the availability of large-scale procedures to obtain pro- and carboxypeptidases, particularly the human forms, very useful. The efficient heterologous pro-

---

*Fig. 5.* Analysis by reverse-phase HPLC of the fragments generated during tryptic activation of human pro-CPA2. At different times, aliquots were taken from the activation mixture under the same conditions as described in the legend to Fig. 4 and analyzed by reverse-phase chromatography in an HPLC system. The action of trypsin was stopped by adding 0.1% trifluoroacetic acid. 100 μl of each sample were loaded on a Vydac C8 reverse-phase column, and elution was followed at 214 nm. The α-fragment corresponds to residues 1–96, the β-fragment to residues 1–94, and the γ-fragment to residues 1–92. PCPA2, pro-CPA2.

*Fig. 6.* Schematic representation of the cleavage points in the activation process of human pro-CPA2 promoted by trypsin. The scheme shows the sequence of the α-helical connecting region between the globular activation domain and the active enzyme. Arrows indicate the tryptic cleavage points observed under the same conditions as described in the legend to Fig. 4, which released three activation fragments (α, β, and γ) of 96, 94, and 92 residues, respectively, as deduced by reverse-phase HPLC, N-terminal sequencing, and mass spectrometry studies. No cleavages were detected in other parts of the molecule. The activation sequence of pro-CPA2 is denoted by asA2.
duction of extra-pancreatic metalloprocarboxypeptidases, with many of them involved in important biological functions and some with a tertiary structure or domains structurally related to the pancreatic ones (4), could also facilitate their detailed characterization. Hence, the development of efficient heterologous expression systems for enzymes and proenzymes such as the one reported here is highly desirable.

The experiments designed to establish an easy and robust expression system capable of providing large amounts of soluble human pro-CPA2 were restricted to shake-flask cultures, which rendered a production of ~250 mg of recombinant proenzyme/liter or 180 mg of active enzyme/liter. This production is more than enough for studies such as those presented here and for subsequent structural determinations as well as for other analytical and semipreparative purposes. However, *P. pastoris* can increase this production 10-fold by scaling up from shake-flask to high density fermentation (19).

The original target for the removal of the α-MF propeptide is the sequence KREAEAEIA, which is cleaved by the yeast endopeptidase KEX2 after the dibasic peptide and undergoes subsequent elimination of Glu-Ala dipeptides by dipeptidyl aminopeptidase A (33). Thus, according to the design of the pPIC9-pro-CPA2 expression plasmid, the pro-segment of the α-mating factor was expected to be cleaved after the dibasic residues Lys-Arg by a single KEX2 endopeptidase action. However, the N-terminal sequence of expressed recombinant pro-CPA2 displayed microheterogeneity at the amino terminus, indicating that cleavage by the KEX2 endopeptidase is only partial and that other endoproteinas might be responsible for the unspecific cleavages (see Fig. 3). The microheterogeneity observed did not become an important problem since the principal product found in all of the cultures was native pro-CPA2. Part of the secreted fusion product appeared glycosylated; however, this product was converted to normal pro-CPA2 after deglycosylation. In any case, all of these products can generate fully active CPA2 by limited proteolysis.

A correctly processed α-MF propeptide has been reported for recombinant proteins expressed in *P. pastoris* without the C-terminal Glu-Ala extension (34, 35); other reports indicated heterogeneity at the N terminus of the secreted proteins (i.e. aprotinin and coffee bean galactosidase), also resulting from different cleavage points in the processing of the α-MF pro-segment (36, 37). Thus, the microheterogeneity observed in some cases could be dependent on the nature of the recombinant protein expressed (33).

The trypsin-promoted maturation mechanism of human pro-CPA2, as derived from our studies, is shown in Fig. 6. Due to the presence of several arginines at the boundary region between the pro-segment and the enzyme moiety, a rapid tryptic cleavage at the most exposed arginine of this region and subsequent cleavages at the remaining trypsin targets should be expected. Alternatively, simultaneous cleavages could occur, but our results indicate that it is the first mechanism that takes place. In contrast to other procarboxypeptidases studied, the Arg96–Ser97 peptide bond is the first target for trypsin action observed in pro-CPA2, releasing a pro-segment of 96 residues into the medium (α-fragment in Figs. 5 and 6). From sequence alignments and structure comparisons (14), this peptide bond is considered to belong to the enzyme moiety in porcine pro-CPA1 and pro-CPB, where the cleavage occurs two residues farther along N-terminally (16, 17). The released primary activation fragment and mature enzyme are therefore two residues larger and shorter, respectively, in pro-CPA2.

The second step in the proteolytic processing is the rapid elimination of the C-terminal Glu-Arg dipeptide end from the α-fragment. The product of this cleavage (β-fragment in Figs. 5 and 6) is due to a tryptic action on arginine 94 of the α-fragment, as shown by mass spectrometry analysis and by the lack of release of free amino acids. The action of trypsin in this conversion is not as fast as in the first cleavage, probably due to the less efficient tryptic endoprotease activity near the C-terminal end of proteins, and takes place while some proenzyme molecules are still intact. The third and last sequence of 92 residues (γ-fragment in Figs. 5 and 6), resistant to further proteolysis under these conditions, is generated by trypsin action in the long run, releasing an Arg-Arg dipeptide into the activation medium. It is also worth commenting upon that no release of free amino acids into the medium is observed in the course of activation, in contrast to previous observations in the corresponding activation processes of pro-CPA1 and pro-CPB (16, 17, 38). This indicates the high specificity shown by human CPA2, which is unable to trim the C-terminal arginines from its activation peptides, even at the high concentration of these species in the activation medium.

According to this study, it seems clear that the pro-segments of pro-CPA2 released into the medium do not inhibit the activity of the active enzyme, even in its longer (primary) form (96 amino acids), since CPA2 reaches total activity following a rapid hyperbolic activation curve at activation times when a substantial concentration of the α-fragment sequence is still present in solution. In contrast to this behavior, porcine pro-CPA1, which shows a slower and biphasic activation process, needs subsequent trimmings at the C-terminal end and a second tryptic cleavage inside the globular domain of the prosequence to achieve full carboxypeptidase activity. In this sense, pro-CPA2 behaves in a way that is more similar to pro-CPB (17).

X-ray crystallography and modeling studies have shown the structural similarities between pro-CPA1 and pro-CPA2, both sharing a long C-terminal α-helix at the region connecting the globular activation domain with the active enzyme, and the differences of these two proenzymes from pro-CPB, whose corresponding connecting region has a much shorter α-helix and is less structured overall (5, 6, 14). These differences were proposed to be the primary determinant responsible for the diversity observed in the rates of activation of pro-CPA1 and pro-CPB, arguing that the larger regular structure of the connecting region in pro-CPA1 would render the interactions with the enzyme more stable and make the structural relaxation needed for full release slower (4, 39). Modeling studies with pro-CPA1 (14) have shown that such a connecting region is also structured in a long α-helix, at least the same size as in porcine pro-CPA1. Taking into account the former evidence and the observation of a rapid and monotonic activation curve for pro-CPA2, it can be concluded that the folding of the connecting region of procarboxypeptidases in a long α-helix does not give rise to inhibitory activation fragments and to slow activation processes by itself. Other structural determinants, such as surface complementarity and electrostatic and Van der Waals interactions, should be considered to evaluate the stability of the bimolecular complex between the pro-segment and the enzyme moiety once the first proteolytic cleavage has taken place.

Preliminary x-ray diffraction studies in human pro-CPA2 show that the connecting region is highly structured and that extensive interactions take place between the enzyme moiety and both the globular domain region and the connecting region of the pro-segment. These observations confirm that the latter two regions are responsible for the activation behavior of pro-CPA2 and for its functional differences from other procarboxypeptidases.

2 I. García-Sáez, personal communication.
Acknowledgments—We thank Dr. S. Bartolome (Laboratory for Image Analysis) for help in the elaboration of the figures and Dr. F. Canals (Servei de Sequenciació i Biocomputació) for performing the N-terminal sequence analysis of the proteins.

REFERENCES

1. Neurath, H. (1986) J. Cell. Biochem. 32, 35–49
2. Puigserver, A., Chaps, C., and Kerfelec, B. (1986) in Molecular and Cellular Basis of Digestion (Dessauille, P., Sjöström, H., and Noren, O., eds) pp. 235–247, Elsevier Science Publishers B.V., Amsterdam
3. Auld, D. S., and Vallee, B. L. (1987) New Compr. Biochem. 3, 222–246
4. Avilés, F. X., Vendrell, J., Guasch, A., Coll, M., and Huber, R. (1993) Eur. J. Biochem. 211, 391–399
5. Coll, M., Guasch, A., Avilés, F. X., and Huber, R. (1991) EMBO J. 10, 1–9
6. Guasch, A., Coll, M., Avilés, F. X., and Huber, R. (1992) J. Mol. Biol. 224, 141–157
7. Reynolds, D. S., Stevens, R. L., Gurley, D. S., Lane, W. S., Austen, K. F., and Serafin, W. E. (1989) J. Biol. Chem. 264, 20094–20099
8. Tan, A. K., and Eaton, D. L. (1995) J. Biol. Chem. 270, 141–157
9. Skidgel, R. A. (1996) in Zinc Metalloproteases in Health and Disease (Hooper, N. M., ed) pp. 241–283, Taylor & Francis Ltd., London
10. Song, L., and Fricker, L. D. (1997) J. Biol. Chem. 272, 10543–10550
11. Gardell, S. J., Craik, C. S., Clausen, E., Goldsmith, E., Stewart, C. B., Graf, M., and Rutter, W. J. (1989) J. Biol. Chem. 263, 17828–17836
12. Pascual, R., Burgos, F. J., Salva, M., Soriano, F., Méndez, E., and Avilés F. X. (1993) Eur. J. Biochem. 179, 609–616
13. Fanning, Z., Kobe, B., Stewart, C. B., Rutter, W. J., and Goldsmith, E. (1991) J. Biol. Chem. 266, 24606–24612
14. Catasús, L., Vendrell, J., Avilés, F. X., Carreira, S., Puigserver, A., and Billette, M. (1995) J. Biol. Chem. 270, 6651–6657
15. Normant, E., Gros, C., and Schwartz, J. C. (1995) J. Biol. Chem. 270, 20543–20549
16. Vendrell, J., Cuchillo, C. M., and Avilés, F. X. (1990) J. Biol. Chem. 265, 6949–6953
17. Villegas, V., Vendrell, J., and Avilés, F. X. (1995) Protein Sci. 4, 1792–1800
18. Oppeezo, O., Ventura, S., Bergman, T., Vendrell, J., Jurnvall, H., and Avilés, F. X. (1994) Eur. J. Biochem. 222, 55–63
19. Cregg, J. M., Vedvick, T. S., and Rasche, W. C. (1993) Bio/Technology 11, 905–910
20. Huynh, K. F. M. (1994) Trends Biotechnol. 12, 234–239
21. Laethem, R. M., Blumenkopf, T. A., Cory, M., Elwell, L., Moxham, C. P., Ray, P. H., Walton, L. M., and Smith, G. K. (1996) Arch. Biochem. Biophys. 332, 3–18
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
23. Schagger, H., and Von Jagow, G. (1987) Anal. Biochem. 166, 368–379
24. Vendrell, J., and Avilés, F. X. (1986) J. Chromatogr. 358, 401–413
25. Ohi, H., Ohtani, W., Okazaki, N., Furuhata, N., and Ohmura, T. (1996) Yeast 12, 31–40
26. Kjeldsen, T., Brandt, J., Andersen, A. S., Egol-Mitani, M., Hach, M., Petterson, A. P., and Vad, K. (1996) Gene (Amst.) 170, 97–112
27. Comis-Ruth, F. X., Gómez, M., Bode, W., Huber, R., and Avilés, F. X. (1995) EMBO J. 14, 4387–4394
28. Villegas, V. (1995) Detailed Characterization of the Activation Process of Procarboxypeptidase B through the Use of Inhibitors: Design and Construction of Mutants of the Activation Segment. Ph.D. thesis, Universitat Autònoma de Barcelona
29. Phillips, M. A., and Rutter W. J. (1996) Biochemistry 35, 6771–6776
30. Villegas, V., Azuaga, A., Catasús, L., Revertor, D., Mateo, P. L., Avilés, F. X., and Serrano, L. (1995) Biochemistry 34, 15105–15110
31. Vitols, K. L., Haag-Zeino, B., Baer, T., Montejano, Y. D., and Huynh, F. M. (1995) Cancer Res. 55, 478–481
32. Rowell, S., Paupit, R. A., Tucker, A. D., Melton, D. M., and Brick, P. (1997) Structure 15, 337–347
33. Brake, A. J., Merryweather, J. P., Heberlein, V. A., Masahir, F. R., Mullenbach, G. T., Urdea, M. S., Valenzuela, P., and Barr, P. J. (1994) Proc. Natl. Acad. Sci. U. S. A. 81, 4642–4646
34. Wagner, S. L., Siegel, R. S., Vedvick, T. S., Raschke, W. C., and Van Nostrand, W. E. (1992) Biochem. Biophys. Res. Commun. 196, 1139–1145
35. Vozza, L. A., Wittler, L., Higgins, D. R., Purcell, T. J., Bergsmedh, C., Edberg, C. M., LaVallie, E. R., and Hoeffler, J. P. (1996) Bio/Technology 14, 77–91
36. Vedvick, T., Buchholz, R. G., Engel, M., Urrica, M., Kinney, J., Prevot, S., Giesing, R., and Thill, G. P. (1991) J. Ind. Microbiol. 7, 197–202
37. Zhu, A., Monahan, C., Zhang, Z., Hurst, R., Leng, L., and Goldstein, J. (1995) Arch. Biochem. Biophys. 324, 65–70
38. Burgos, F. J., Salva, M., Villegas, V., Soriano, F., Méndez, E., and Avilés F. X. (1991) Biochemistry 30, 4092–4099
39. Vendrell, J., Catasús, L., Oppeezo, O., Ventura, S., Villegas, V., and Avilés, F. X. (1993) in Innovations in Proteases and Their Inhibitors (Avilés, F. X., ed) pp. 279–297, Walter de Gruyter & Co., Berlin
Overexpression of Human Procarboxypeptidase A2 in *Pichia pastoris* and Detailed Characterization of Its Activation Pathway

David Reverter, Salvador Ventura, Virtudes Villegas, Josep Vendrell and Francesc X. Avilés

*J. Biol. Chem. 1998, 273:3535-3541.*
doi: 10.1074/jbc.273.6.3535

Access the most updated version of this article at http://www.jbc.org/content/273/6/3535

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 34 references, 9 of which can be accessed free at http://www.jbc.org/content/273/6/3535.full.html#ref-list-1