Characterisation and preliminary X-ray diffraction analysis of human pancreatic procarboxypeptidase A2

David Reverter[a], Isabel García-Sáez[b], Lluis Catasús[a], Josep Vendrell[a], Miquel Coll[b], Francesc X. Avilés[a,*]

[a]*Departament de Bioquímica, Facultat de Ciències and Institut de Biologia Fonamental, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*
[b]*Departament de Biologia Molecular i Cellular, Centre d'Investigació i Desenvolupament, CSIC, Jordi Girona 18–26, 08034 Barcelona, Spain*

Received 13 October 1997; revised version received 10 November 1997

Abstract Human procarboxypeptidase A2 has been expressed in a *Pichia pastoris* heterologous system and purified by hydrophobic interaction and anion exchange chromatographies. The hydrolytic action of carboxypeptidase A2 on peptide substrates with different lengths and residues at the C-terminus was analysed, and a preference towards long substrates with aromatic amino acids in their C-terminal end, particularly tryptophan, was found; with such substrates its activity is similar or higher than that of bovine carboxypeptidase A1. Procarboxypeptidase A2 has been crystallised using a vapour diffusion approach; the crystals obtained belong to the monoclinic system, spacegroup $P2_1$, and present one procarboxypeptidase A2 molecule per asymmetric unit. The crystals diffract beyond 1.8 Å resolution and are suitable for detailed X-ray analysis.

Key words: Carboxypeptidase A2; Proenzyme; Substrate specificity; Crystallography

1. Introduction

Human procarboxypeptidase A2 (proCPA2) is the inactive precursor of a pancreatic proteolytic enzyme that belongs to the family of metallo exoproteases [1]. Human pancreas secretes many of the proteolytic enzymes that participate in the digestive processes of alimentary proteins and peptides. The occurrence of these enzymes as inactive (pro) forms, and their proteolytic maturation upon arrival at their operating zone, is a control mechanism for the optimum expression of their biological activity [2,3]. Three forms of monomeric procarboxypeptidases, A1, A2 and B occur in human pancreas, together with a binary complex between the A1 form and proproteinase E [4]. The mature forms of these human procarboxypeptidases have a different role in their functional activity of removing C-terminal ends of polypeptide chains [4,5]: carboxypeptidase B preferentially hydrolyses basic residues, while carboxypeptidases A1 and A2 have a proteolytic activity on aliphatic and aromatic or only aromatic side-chain residues, respectively.

Procarboxypeptidase A2 is an isoform reported until now only in rat and in humans [4,6,7]. Some structural and functional differences between A1 and A2 isoforms have been described in the two systems. The A2 isoform shows an activation mechanism closer to the B form than to the A1 form [7]; this is surprising considering the closer homologies to the latter in sequence and specificity. The substrate preference for A1 and A2 isoforms has been studied in rat; rat CPA1 prefers smaller amino acids while rat CPA2 prefers bulkier amino acids [6]. The reports of the crystal structure of rat mature CPA2 [8] and of several crystal structures of procarboxypeptidases [9–11] have allowed a better understanding of the functional differences between these proenzymes based on their structural determinants. No three-dimensional structure of human proCPA2 is available until now.

Using a *Pichia pastoris* heterologous system we have overexpressed human proCPA2 in the supernatant medium of the yeast. In this report, we present results concerning the extensive purification of this proenzyme from the *Pichia pastoris* extracellular medium and the systematic enzymatic characterisation of its mature form with a series of synthetic substrates that differ in length and in C-terminal amino acid. In addition, the crystallisation and the crystal parameters of the recombinant human proCPA2 obtained by X-ray diffraction analysis are also reported. The derivation of the three-dimensional structure of this proenzyme could help to clarify the structural and functional differences between the activation processes and the substrate specificities for the different types of human pancreatic procarboxypeptidases.

2. Materials and methods

2.1. Protein purification

Human recombinant proCPA2 was secreted at high levels in a *Pichia pastoris* heterologous system (Reverter et al., to be published elsewhere). Purification from the supernatant was a crucial step due to non-specific and heterogeneous cleavages at the signal peptide that produced a different N-terminus in the recombinant protein. The first purification step was a hydrophobic interaction chromatography: 200 ml of *P. pastoris* extracellular medium were loaded onto an atmospheric pressure butyl column (butyl 650M Toyopearl, from Tosohaas, 1.5×15 cm) after equilibration with 30% ammonium sulphate. Chromatographic elution was performed at 4°C with a linear gradient of ammonium sulphate from 30% to 0% in 2 h. The fractions containing the zymogen were selected by enzymatic activity detection with the synthetic substrate *N*-(3-[2-furyl]acryloyl)-L-phenylalanyl-L-phenylalanine (FAPP) after activation with trypsin. The protein fraction was desalted by dialysis against 30 mM MES-acetate (pH 5.7) buffer. A second chromatography in the same buffer at pH 5.7 was performed in an anion exchange preparative column (TSK-DEAE 5PW, from Toyo-Soda, 2.5×15 cm) connected to an FPLC system (Pharmacia). A linear smooth salt gradient from 0 M to 0.45 M ammonium acetate was used to separate the distinct N-terminal processed forms of the recombinant proCPA2; the chromatography was performed at 22°C and the flow rate kept at 2.5 ml/min. Eluted proteins were kept precipitated with 45% of ammonium sulphate, at 4°C, to avoid proteolysis. The nature and purity of the eluted forms of the human proCPA2 was confirmed by SDS-acrylamide gel electrophoresis, automated

*Corresponding author. Fax: (34) (3) 581 2011.
E-mail: FX.Aviles@blues.uab.es

Abbreviations: proCP, procarboxypeptidase; Cbz, carbobenzoxy; FAPP, *N*-(3-[2-furyl]acryloyl)-L-phenylalanyl-L-phenylalanine

PII S0014-5793(97)01476-2

N-terminal sequence analysis, electrospray mass spectrometry and by activity measurements with FAPP, as previously described [4].

For crystallisation studies, the purified protein was immediately used to avoid autolysis. For kinetic studies, another purification step to separate the carboxypeptidase moiety from its pro-segment was required. The ammonium sulphate precipitate of proCPA2 was recovered by centrifugation at 13 000 rpm, redissolved in 50 mM Tris-HCl (pH 8.0) and subsequently desalted on a PD-10 Sephadex G-25M gel filtration column. The proenzyme was activated by trypsin at 40:1 (w/w) ratio, in 10 min at 25°C. An anion exchange chromatography (analytic TSK-DEAE column) in an FPLC system was used to separate the tryptic fragments with a linear salt gradient from 5% to 40% of 0.8 M ammonium acetate. The eluted CPA2 was dissolved at high salt concentration (0.5 M NaCl) to avoid aggregation.

2.2. Enzymatic assays

Different synthetic substrates were used to characterise the recombinant purified human CPA2. The rate of hydrolysis of each substrate was continuously measured spectrophotometrically in 50 mM Tris-HCl, 0.5 M NaCl (pH 7.5) at 25°C. The wavelengths used to monitor each reaction were as follows: Cbz-Gly-Gly-Phe, Cbz-Gly-Gly-Leu, Cbz-Gly-Gly-Val, Cbz-Gly-Gly-Ala, Cbz-Gly-Gly-Ser and Cbz-Gly-Phe, 224–228 nm; Cbz-Gly-Tyr and Cbz-Gly-Gly-Tyr, 236 nm; Cbz-Gly-Trp and Cbz-Gly-Gly-Trp, 300–304 nm. Initial rates, determined from the first 5–10% of each reaction, were obtained at substrate concentrations bracketing the K_m value whenever possible. The kinetic parameters, k_{cat} and K_m, were obtained using 6–10 experimental points by direct fit in a Michaelis-Menten curve using the Enzfitter program [12].

2.3. Crystallisation and X-ray data collection

The purified human proCPA2 was concentrated to 12–13 mg/ml using an Amicon Centriprep 10 concentrator. Protein concentration was confirmed spectrophotometrically and by Bradford Bio-Rad assay. Crystallisation was carried out using the hanging drop vapour diffusion method at 4°C and at 20°C. In a preliminary assay, conditions from the Hampton Crystal Screen I and Crystal Screen II [13] were carried all out. Crystals were obtained in 20% PEG MME 8000, 0.1 M Tris, pH 8.5, 0.01 M $NiCl_2$ mixing an equal volume of reservoir solution and protein solution. The pH was screened from 6.0 to 8.5 and the concentration of protein was optimised. The best crystals were obtained by mixing 5 μl of protein solution with 1 μl of reservoir solution at pH 8.0 and 4°C.

Crystals were harvested in a solution containing 40% of PEG 8000 MME, 0.1 M Tris, pH 8.0 and mounted in a thin-walled glass capillary tube. Diffraction analysis was carried out using a Rigaku RU-200 rotating anode generator, K_α Cu radiation, and a 300-mm Image Plate Area Detector from Mar Research. The data were evaluated with the program DENZO [14].

3. Results and discussion

The optimisation of the purification method of proCPA2 from the *Pichia pastoris* supernatant leads to the obtainment of pure protein through only two chromatographic steps in two days. The *P. pastoris* supernatant can be directly loaded onto a butyl column at atmospheric pressure after equilibration at 30% saturation with ammonium sulphate. The use of hydrophobic interaction chromatography avoids the need for desalting required in other chromatographic variants. The peak eluted at this step is dialysed against the buffer used in the second chromatographic step, which is performed in an anion exchange column (TSK-DEAE) due to the negatively charged character of the proenzyme.

In the selected expression system and conditions, *Pichia pastoris* produces several forms of proCPA2 with different N-terminal processing and glycosylated variants (Reverter et al., to be published elsewhere). To solve the problem of heterogeneity of proCPA2s it was necessary to investigate high resolution chromatographic conditions to separate the several expressed forms. After many trials, we found that the interaction of the protein variants with the TSK-DEAE column is

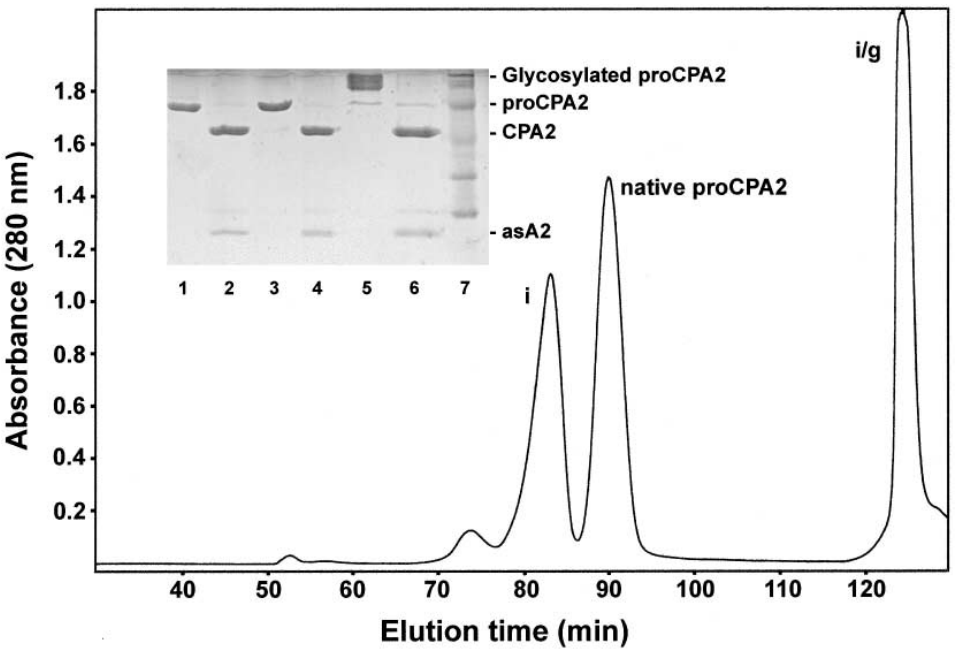

Fig. 1. Purification of the correctly and incorrectly processed recombinant forms of human procarboxypeptidase A2. Anion exchange chromatography on a TSK-DEAE column was performed in an FPLC system with a smooth gradient from 30 mM MES-acetic acid, pH 5.7 (solvent A) to 30 mM MES-acetic acid, 0.45 M ammonium acetate, pH 5.7 (solvent B). The gradient was: 0% solvent B from 0 to 7 min; 5% solvent B at 15 min; 15% solvent B at 100 min; 100% solvent B at 110 min. Inset: electrophoretic analysis of the peaks of the chromatogram; lanes 1 and 2: I, incorrectly processed proCPA2; lanes 3 and 4: native proCPA2; lanes 5 and 6: I/G, incorrectly processed and glycosylated proCPA2; lanes 1, 3, 5 and 2, 4, 6 contain samples before and after activation with trypsin, respectively. Lane 7: molecular mass markers. The position of the activation (pro-) segment of PCPA2 is indicated by asA2.

weak and slightly different at pH 6.0, the proteins still being activatable (the zinc atom is lost at lower pH). The combination of this pH with a smooth ammonium acetate gradient allowed the separation of the well-processed N-terminal proCPA2 from the other incorrectly processed or glycosylated forms. A typical profile obtained in the second chromatography is displayed in Fig. 1. The main peaks contained different forms of proCPA2 as shown by enzymatic measurements, electrophoretic analysis and N-terminal sequencing, and each peak generated active carboxypeptidase after activation with trypsin. The first and third large peaks were found to contain incorrectly processed proCPA2s, the latter being a glycosylated form. The second large peak (eluted at 90 min) was found to contain native – well processed – proCPA2, as confirmed by mass spectrometry when its molecular mass was compared with that derived from gene sequencing [5]. The purification of the native proCPA2 was a prerequisite for the subsequent enzymatic and crystallographic analyses.

Fully active and pure CPA2 was produced by limited proteolysis of any of the above proCPA2 variants with trypsin, followed by anion exchange chromatography on a TSK-DEAE column, as described in Section 2. The occurrence of the differential processing or glycosylation of proCPA2 during expression at the N-terminus of the pro-region did not prevent the homogeneous generation of CPA2 by trypsin cleavage at the boundary between the pro-segment and the active enzyme moiety.

A functional characterisation of the enzyme was performed by means of a series of substrates of the type Cbz-(Gly)n-L-X that differed in their lengths and in their C-terminal amino acids. In Table 1 the kinetic constants derived for human CPA2 are compared with those for bovine CPA1, a model enzyme of type A carboxypeptidases. In the case of shorter substrates ($n = 1$), human CPA2 was only able to hydrolyse those containing a C-terminal aromatic residue, with the exception of a residual activity detected for the substrate with C-terminal leucine; no activity was detected with substrates containing other residues at this position, such as alanine, serine, arginine or lysine. For substrates with two glycine residues, the higher activity was observed with C-terminal tryptophan followed by those with tyrosine and phenylalanine, respectively. The k_{cat}/K_m values for these longer substrates for human CPA2 are comparable with those derived for bovine CPA1 for C-terminal tyrosine, smaller for phenylalanine and much larger for tryptophan; in the latter case, the activity is not measurable in bovine CPA1. These results agree with the hypothesis of a large binding pocket proposed for rat CPA2 that would be able to accommodate the large aromatic substrates better [8]. For short substrates with only one glycyl residue, there is also a correlation between the activity and the size of the aromatic C-terminal residue in the substrate, although the k_{cat}/K_m values are lower compared with the substrates with one more glycine. Thus, it seems clear that longer substrates can be placed better in the binding pocket of the enzyme, allowing for a better filling of the different subsites, as proposed for other carboxypeptidases [1].

The kinetic constants in Table 1 confirm the results that define the A2 isoform as an enzyme that preferentially catalyses the removal of aromatic and bulkier residues from the C-terminus of peptides and proteins. In contrast to the A1 isoform, CPA2 is the only reported carboxypeptidase with specificity for C-terminal tryptophan residues; this fact increases the importance of this exoprotease that can be used in controlled digestions removing C-terminal tryptophan ends of peptides and proteins in vitro or in vivo. It is also worth mentioning that the proenzyme shows intrinsic (remnant) activity against the above mentioned substrates. The percentages of activity of the proenzyme compared to the active enzyme against a number of substrates at 1 mM concentration are as follows: 3% for Cbz-Gly-Trp, 1.5% for Cbz-Gly-Gly-Trp, 10% for Cbz-Gly-Tyr, 5% for Cbz-Gly-Gly-Tyr, and 0.45% for FAPP. This is an indication that the active site is preformed in the zymogen and that its accessibility is more restricted for longer substrates.

Crystallisation of recombinant proCPA2 was only possible after an accurate chromatographic purification of the well-processed N-terminal proCPA2 in an anion exchange TSK-DEAE column, performed at low pH and with smooth gradient conditions, as previously described. Repurification of the peak under the same chromatographic conditions was necessary to avoid minor contamination from other variants that could interfere with the crystallisation of the enzyme. To avoid autolysis processes in the purified proCPA2, crystallisation experiments were started just after the repurification of the protein. The proenzyme was concentrated up to 12.5 mg/ml and dialysed against 2 mM MOPS, pH 7.0 using an Ami-

Table 1
Kinetic constants for peptide substrate hydrolysis by human CPA2 and bovine CPA1, assayed in aqueous buffer (50 mM Tris-HCl, 0.5 M Na Cl, pH 7.5) at 25°C

Substrate	Human CPA2			Bovine CPA1		
	K_{cat} (s^{-1})	K_m (μM)	K_{cat}/K_m ($\times 10^{-5}$)	K_{cat} (s^{-1})	K_m (μM)	K_{cat}/K_m ($\times 10^{-5}$)
Cbz-Gly-Gly-Phe	58.3 ± 2.4	372 ± 30	1.57	131.5 ± 3.1[a]	172 ± 12[a]	7.62
Cbz-Gly-Gly-Tyr	70.0 ± 5.3	125 ± 15	5.60	56.3 ± 2.0[a]	102 ± 2[a]	5.51
Cbz-Gly-Gly-Trp	90.3 ± 7.0	146 ± 9	6.18	NM[b]	NM[b]	
Cbz-Gly-Gly-Leu	11.8 ± 1.1	5300 ± 1400	0.03	63.4 ± 2.5[b]	1180 ± 93[b]	0.54
Cbz-Gly-Gly-Val	NM	NM		19.5 ± 2.0	3720 ± 390	0.52
Cbz-Gly-Gly-Ala	NM	NM		NM	NM	
Cbz-Gly-Gly-Ser	NM	NM		NM	NM	
Cbz-Gly-Phe	16.1 ± 1.3	2270 ± 200	0.07	41.7 ± 2.8[a]	1093 ± 154[a]	0.38
Cbz-Gly-Tyr	9.7 ± 0.6	175 ± 10	0.56	16.0 ± 0.6[a]	394 ± 29[a]	0.41
Cbz-Gly-Trp	33.8 ± 1.1	261 ± 12	1.29	50.0 ± 4.3[a]	3310 ± 430[a]	0.15

NM, non-measurable; the extremely low rates of hydrolysis of such substrates did not permit derivation of Michaelis-Menten parameters.
[a]Kinetic values taken from [6].
[b]Kinetic values taken from [15].

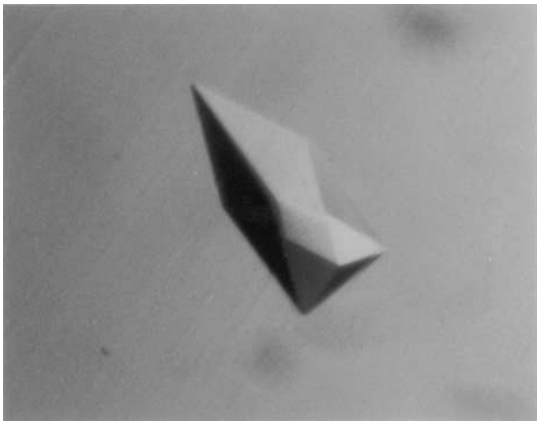

Fig. 2. Crystal specimen of recombinant human procarboxypeptidase A2. The crystals were obtained in 20% PEG and 10 mM $NiCl_2$, belonged to the monoclynic system, and the largest ones were of approximate dimensions of $1\times1\times0.6$ mm.

con centriprep 10 concentrator. Large crystals were obtained using PEG MME as precipitant. After growing during one week, the crystals achieved a final size of $1\times1\times0.6$ mm.

The obtainment of well-ordered crystals of proCPA2, suitable for X-ray diffraction analysis, was found to be critically dependent on an efficient and fast purification, performed just before crystallisation. The crystals (Fig. 2) diffract beyond 1.8 Å resolution using a rotating anode X-ray source and are stable under the X-ray beam. They belong to the monoclinic space group $P2_1$ with unit cell parameters $a=42.2$ A, $b=87.1$, $c=59.0$ and $\beta=99.3°$. The V_m value is 2.23 $Å^3$/dalton with one proCPA2 molecule per asymmetric unit. A 1.8 Å resolution data set has been collected resulting in 37 420 unique reflections at the 2σ (F_0) level. This data set is 94% complete (88% complete in the 1.86–1.8 outermost resolution shell). The overall $I/\sigma(I)$ is 32.6, being 9.1 in the 1.86–1.8 resolution shell. The determination of the native structure of the proCPA2 will be performed by means of the molecular replacement method, using the structure of porcine pancreatic procarboxypeptidase A1 [10] as initial model.

The resolution of the human proCPA2 structure could clarify the functional and structural differences between procarboxypeptidase isoforms and should provide insight into the inhibition mechanism and trypsin activation process of the proenzyme, as well as on the substrate preferences through the study of its specificity pocket. It could also be of great use in the redesign of the specificity pocket of the active enzyme, a subject of biotechnological interest nowadays [15].

Acknowledgements: This work has been supported by Grants BIO95 0848 from CICYT (Ministerio de Educación y Ciencia, Spain) and PB95–0224 from the Ministerio de Educación y Ciencia, and by the Centre de Referència en Biotecnologia (Generalitat de Catalunya). D. Reverter is recipient of a fellowship from CIRIT (Generalitat de Catalunya).

References

[1] Avilés, F.X., Vendrell, J., Guasch, A., Coll, M. and Huber, R. (1993) Eur. J. Biochem. 211, 391–399.
[2] Puigserver, A., Chapus, C. and Kerfelec, B. (1986) in: Molecular and Cellular Basis of Digestion (Desnuelle, P., Sjostrom, H. and Noren, O., Eds.) pp. 325–247, Elsevier, Amsterdam.
[3] Auld, D.S. and Vallee, B.L. (1987) New Compr. Biochem. 16, 201–256.
[4] Pascual, R., Burgos, F.J., Salvà, M., Soriano, F., Méndez, E. and Avilés, F.X. (1989) Eur. J. Biochem. 179, 609–616.
[5] Catasús, Ll., Vendrell, J., Avilés, F.X., Carreira, S., Puigserver, A. and Billeter, M. (1995) J. Biol. Chem. 270, 6651–6657.
[6] Gardell, S.J., Craik, C.S., Clauser, E., Goldsmith, E., Stewart, C.-B., Graf, M. and Rutter, W.J. (1988) J. Biol. Chem. 263, 17929–17936.
[7] Oppezzo, O., Ventura, S., Bergman, T., Vendrell, J., Jörnvall, H. and Avilés, F.X. (1994) Eur. J. Biochem. 222, 55–63.
[8] Famming, Z., Kobe, B., Stewart, C.B., Rutter, W.J. and Goldsmith, E. (1991) J. Biol. Chem. 266, 24606–24612.
[9] Coll, M., Guasch, A., Avilés, F.X. and Huber, R. (1991) EMBO J. 9, 1–9.
[10] Guasch, A., Coll, M., Avilés, F.X. and Huber, R. (1992) J. Mol. Biol. 224, 141–157.
[11] Gomis-Rüth, F.X., Gómez, M., Bode, W., Huber, R. and Avilés, F.X. (1995) EMBO J. 14, 4387–4394.
[12] Leatherbarrow, R.J. (1987) Enzfitter: a non-linear regression data analysis program for the IBM-PC. Elsevier Biosoft, Cambridge, UK.
[13] Jankarik, J. and Kim, S.H. (1991) J. Appl. Crystallogr. 24, 409–411.
[14] Otwinowski, Z. (1993) in: Proceedings of the CCP4 Study Weekend: Data Collection and Processing, 29–30 January 1993. Compiled by L. Sawyer, N. Isaacs and S. Bailey, SERC Daresbury Laboratory, England, pp. 56–62.
[15] Smith, G.K., Banks, S., Blumenkopf, T.A., Cory, M., Humphreys, J., Laethem, R.M., Miller, J., Moxham, C.P., Mullin, R., Ray, P.H., Walton, L.M. and Wolfe III, L.A. (1997) J. Biol. Chem. 272, 15804–15816.