Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 1041–1045
Carbohydrate RESEARCH

Indiscriminate glycosylation of procarboxypeptidase Y expressed in *Pichia pastoris*

Haruko Maeda, Eri Chatani,* Takahiro Koyama, Masatoshi Sugiura, Hirohisa Izumi and Rikimaru Hayashi

Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Received 4 November 2003; received in revised form 8 January 2004; accepted 14 February 2004

Abstract—To obtain large amounts of deglycosylated procarboxypeptidase Y (proCPY), in which all of the N-glycosylation sites were replaced by alanine residue by the point mutation method, an expression system was constructed using *Pichia pastoris*. The secreted enzyme was characterized by SDS-PAGE, native PAGE, MALDI-TOF mass spectrometry, and dynamic light scattering, and the results indicated heterogeneity. The recombinant proCPY contained 29 mol of glucose per mole of protein in average, according to the carbohydrate analysis by the phenol–sulfuric acid method. A large part of the recombinant enzyme absorbed on a Con A column: even the break-through fraction of the column contained 3 mol of glucose per mole of protein. These carbohydrates were removed by the mild alkaline treatment. Since the entire N-glycosylation site had been destructed in the present expression system, the carbohydrates contained in the recombinant proCPY are concluded to be O-linked ones, which bound indiscriminately to serine and/or threonine residues.

Keywords: Procarboxypeptidase Y; Carboxypeptidase Y; *Pichia pastoris*; N-Glycosylation site; O-Glycosylation site

1. Introduction

Procarboxypeptidase Y (proCPY), the inactive precursor of carboxypeptidase Y (CPY) from *Saccharomyces cerevisiae*, contains a pro-peptide of 91 amino acid residues in its N-terminal region and a mature region comprised of 421 amino acid residues. It has been proposed that the pro-region is involved in three functions, that is as a transport signal, an intramolecular chaperon, and an intramolecular inhibitor[1,2] but to prove this, an X-ray crystallographic analysis will be required.

CPY contains 16% (w/w) carbohydrates, which have been reported to be bound to four N-glycosylation sites (Asn13, Asn87, Asn168, and Asn368).[3] Crystallization and X-ray crystallography of CPY succeeded after the partial removal of these carbohydrates by treatment with endoglycosidase H.[4] Deglycosylated CPY (Δgly CPY), in which four asparagine residues corresponding to N-glycosylation sites were replaced with alanine residues by site-directed mutagenesis, was obtained, for use in further refining the crystallographic structure of the molecule. Δgly CPY was first expressed in *S. cerevisiae* strain BJ2168. However, in this system, it is highly susceptible to proteases and is partially digested in the vacuole, resulting in a low yield.

Since the purification and crystallization of Δgly proCPY is similar to that of Δgly CPY, we constructed an expression system of proCPY, in which the N-glycosylation sites were replaced with alanine, using *Pichia pastoris*. The rationale for this was that, since the protein would be secreted into the culture medium it could be purified easily and in high yield. In this paper, the production and purification of proCPY expressed by *P. pastoris* are described. The proCPY, N-glycosylation sites (Asn 13, 87, 168, and 368 in CPY) were replaced by

Abbreviations: BT*p*NA, *N*-benzoil-L-tyrosine-*p*-nitroanilide; Con A, concanavalin A; CD, circular dichroism; EDTA, ethylenediaminetetraacetic acid; PEG, polyethylene glycol

* Corresponding author. Tel.: +81-75-753-6111; fax: +81-75-753-6128; e-mail: chatani@protein.osaka-u.ac.jp

0008-6215/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2004.02.003

alanine residues, and the product simply referred to hereafter as the recombinant proCPY.

2. Results and discussion

2.1. Expression and purification of the recombinant proCPY

The recombinant proCPY, in which the asparagine residues at the N-glycosylation sites were replaced with alanine residues was expressed in *P. pastoris* strain GS-115 and secreted into the culture medium. The purification procedures were simple and the protein was obtained in high yield. That is, successive column chromatographies on Butyl-Toyopearl and Poros HQ/M resulted in the pure protein by SDS-PAGE with a molecular mass of approximately 60 kDa, which was smaller by approximately 10 kDa than that of wild-type proCPY (Fig. 1, lane 1). Using this procedure, 2.04 mg of pure protein could be obtained from 1 L of culture in 44% yield, a 50-fold increase over the *S. cerevisiae* system. From these results, it can be concluded that the *P. pastoris* expression system permitted us to efficiently obtain relatively large amounts of recombinant proCPY.

2.2. Properties of the recombinant proCPY

The CD spectrum of the recombinant proCPY was indistinguishable from that of the wild-type proCPY. The specific activity of BT*p*NA after the activation treatment with proteinase K was identical with that of wild-type proCPY (1.5 mmol/min/mg enzyme). The sequence of the five N-terminal residues of the recombinant proCPY was identical with that of wild-type proCPY.

MALDI-TOF mass spectrometry of the recombinant proCPY, however, revealed one broad peak with a molecular mass 57998.58 Da, which was approximately 360.12 Da larger than the theoretical value of a carbohydrate-free proCPY. This suggests that the recombinant proCPY may contain several species with a higher molecular weight. Native-PAGE of the recombinant proCPY supported the molecular heterogeneity of the recombinant proCPY as shown in Figure 1. Dynamic light scattering analysis also indicated molecular heterogeneity (Fig. 3A). These unexpected results prompted us to analyze the carbohydrate of the recombinant proCPY.

2.3. Carbohydrate analysis of the recombinant proCPY

When the purified recombinant proCPY was applied on a Con A column, it was separated into three fractions I, II, and III (Fig. 2). Fraction I was eluted in void volume, followed by fraction II. Fraction III was eluted with an α-methyl-D-mannoside solution, indicating that fraction III contains carbohydrates. All three fractions showed a specific activity identical to that of CPY (1.5 mmol/min/mg enzyme for BT*p*NA) after the activation treatment by proteinase K. Fractions I, II, and III corresponded to 1.4%, 48.7%, and 9.4%, respectively, of the originally loaded sample. The low recovery of the protein (60%) by Con A chromatography might be due to irreversible absorption of part of the protein sample.

When the carbohydrate content of the three fractions was determined as glucose by the phenol–sulfuric acid method, fraction I was found to contain 29 mol of glucose-equivalent per mole of protein. Since the finally

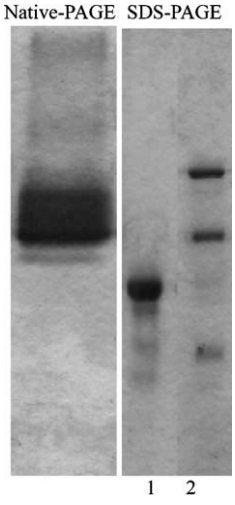

Figure 1. Native-PAGE and SDS-PAGE of the purified recombinant proCPY expressed in *P. pastoris*. Lanes 1 and 2 of SDS-PAGE show the recombinant proCPY and molecular-mass markers (rabbit muscle phospholirase B for 97.4 kDa, bovine serum albumin for 66.2 kDa, hen egg white ovalbumin for 45 kDa, and bovine carbonic anhydrase for 31 kDa, respectively).

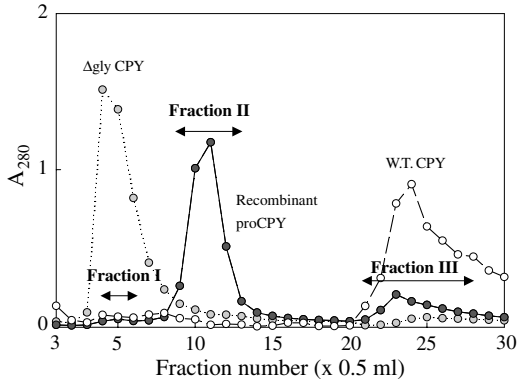

Figure 2. Con A column chromatography, wild-type CPY (---), Δgly CPY (···), and the recombinant proCPY (—). Double headed arrows indicate the pooled fractions.

purified recombinant proCPY contains 29 mol of glucose per mole of protein, this suggests that fraction I contains aggregates or leaked proteins, which might be due to overloading of the column. Fractions II and III contained 3 and 201 mol of glucose per mole of protein (Table 1). These results indicate that the present recombinant proCPY is a mixture of heterogeneously glycosylated proteins.

Since all of the N-linked glycosylation sites of proCPY (Asn 13, 87, 168, 368 in CPY) had been replaced with alanine by site-directed mutagenesis, the carbohydrates contained by recombinant proCPY represent O-linked carbohydrates, and are not N-linked carbohydrates. This conclusion was also obtained from the result of the mild alkaline treatment. The carbohydrate content of the recombinant proCPY was decreased from 2.99 to 0.309 mol of glucose per mole of protein by the mild alkaline treatment.

Table 1. Carbohydrate content of CPY and proCPY, as determined by the phenol–sulfuric acid method[a]

Sample	Carbohydrate detected[b]
	Mole glucose/mole protein
Wild-type CPY	62.28[c]
Δgly CPY expressed in *S. cerevisiae*	0.33
Recombinant proCPY expressed in P. pastoris	
Before Con A chromatography	28.9
Chromatographic fraction I	28.6
Chromatographic fraction II	2.99
Chromatographic fraction III	201.6

[a] Protein-bound carbohydrates were expressed as glucose.
[b] Average values of the duplicate analysis. In reference experiments, it was confirmed that a simple protein, RNase A, contained 0.32 mol of glucose per mole of protein and a glycoprotein, RNase B contained 5.75 mol of glucose per mole of protein (theoretical value, 5.64, see Ref. 8).
[c] The literature value, 58.[9]

2.4. Dynamic light scattering of the fraction II

Dynamic light scattering of fraction II, which absorbed weakly to the Con A column revealed a single distribution (Fig. 3B), which was different from that of the recombinant proCPY before chromatography (Fig. 3A). Although high molecular species with a hydrodynamic radius (Rh) of approximately 100 were removed by the Con A column, fraction II still contained a small amount of carbohydrates (Table 1). These results suggest that the carbohydrate chains on the recombinant proCPY resulted in a heterogeneous molecular species. ProCPY contains 34 serine and 21 threonine residues, all of which are candidates for O-glycosylation sites, therefore determining the location of the carbohydrates will be a difficult task.

As has been shown previously,[5,6] O-glycosylation with heterogeneous lengths of carbohydrate chains, which randomly binds to serine and threonine may be a general phenomenon in the *P. pastoris* expression system and, therefore, O-linked carbohydrates may play some currently unidentified roles in protein synthesis and/or protein transport in *P. pastoris*.

3. Experimental

3.1. Materials

P. pastoris strain GS-115 (*his4*), pHIL-S1 plasmid, and zymolyase (300 units/mL) were purchased from Invitrogen (California, USA). *Escherichia coli* strain TG1 (*K12*, Δ*(lac-pro)*, *supE*, *thi*, *hsdD5/F' traD36*, *proA+B+*, *laclq*, *lacZΔM15*) was from Stratagene (California, USA). pTSY3 plasmid coding *PRC1* gene was generously provided by Dr. K. Breddam, Carlsberg Laboratory (Copenhagen, Denmark).[7] p-Bluescript II SK+ plasmid was a generous gift from Dr. and Professor N. Kato, Graduate School of Agriculture, Kyoto

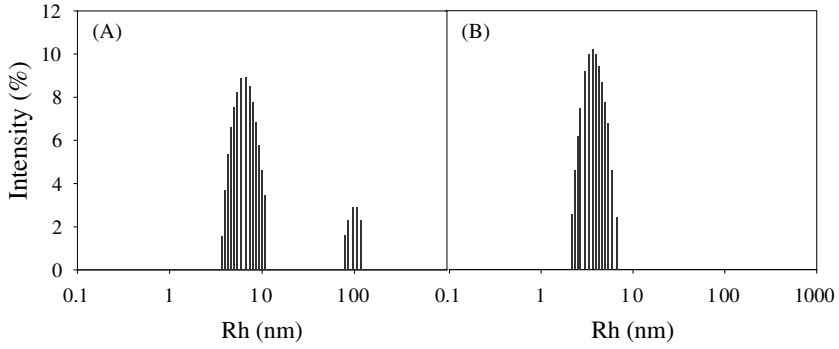

Figure 3. Dynamic light scattering of the recombinant proCPY before (A) and after (B) the Con A chromatography (fraction II). The purified recombinant proCPY and fraction II (see Fig. 2 for details) from the Con A chromatography were analyzed in 0.1 M MES at pH 6.5 and 6.0 for A and B, respectively. The *X* and *Y* axes indicate the radius and distribution of the molecules observed, respectively. The average molecular masses were calculated to be 28474.8 and 74.0 kDa for A and B, respectively.

University (Kyoto, Japan). Proteinase K from *Tritirachium album*, BT*p*NA, and sulfuric acid (97% purity) were from Nacalai Tesque (Kyoto, Japan). Con A-agarose and phenol (99.5% purity) were from Wako Pure Chemical Ind. (Osaka, Japan).

3.2. Construction of expression plasmid and transformation into *P. pastoris* strain GS115

A DNA fragment coding proCPY was amplified by PCR using the Gene Amp PCR System 2400 of Perkin Elmer (California, USA) by using the mutant (N13A, N87A, N168A, and N368A) pTSY3 plasmid as a template, which was previously constructed with the QuickChange Site-Directed Mutagenesis Kit (Stratagene Cloning Systems, California, USA) in our laboratory. Two oligonucleotides (5′-CGG AAT TCA TCT CAT TGC AAA GAC CGT T-3′ and 5′-CGG AAT TCT TAT AAG GAG AAA CCA CCG T-3′) were used as primers to attach the new EcoRI sites at both ends of the DNA fragment. The amplified DNA fragment was digested by treatment with EcoRI, followed by subcloning into p-Bluescript II SK+ and introduced into *E. coli* strain TG1. The plasmid, purified from *E. coli* strain TG1, was treated with EcoRI and cloned into the corresponding site of pHIL-S1 that had been treated with calf intestine alkaline phosphatase, which was again introduced into *E. coli* strain TG1. The constructed expression plasmid contains PHO1, a secretion signal sequence connected directly to the proCPY sequence; 5′ AOX1, promoter fragment; 3′ AOX1 (TT), transcriptional termination fragment; HIS4, *Pichia* wild type gene coding for histidinol dehydrogenase; 3′ AOX1, a sequence from the AOX1 gene that is further 3′ to the TT sequence; ColE1, the *E. coli* origin of replication; and Ampicillin, the ampicillin resistance gene.

The constructed expression plasmid (10 μg) was mixed with *P. pastoris* strain GS-115 *(his4)*, that had been previously spheroplasted with 100 μL zymolyase, and incubated at 25 °C for 10 min. After the incubation, 1 mL PEG/CaT solution, which consisted of 20% (w/v) PEG 3350 in water, 10 mM Tris–HCl, pH 7.5, and 10 mM $CaCl_2$, was added to the mixture with gentle mixing and further incubated at 25 °C for 10 min. After centrifugation of the mixture at 750*g* for 10 min, the supernatant was carefully removed by aspiration. The pellet was suspended in 150 μL of SOS medium, which consisted of 1 M sorbitol, 0.3% yeast extract, 0.6% peptone, 0.6% glucose, and 10 mM $CaCl_2$, incubated at 25 °C for 20 min. After the incubation, the suspension was mixed with 850 μL of 1 M sorbitol and plated on RDB plates. Using the above procedure, the expression plasmid was introduced into the corresponding site of the *P. pastoris* strain GS-115 genome DNA. Transformed cells were grown on minimal medium.

3.3. Expression and purification of the recombinant proCPY

After they were fully grown in 300 mL of YPD medium, which consisted of 1% yeast extract, 2% peptone, and 2% glucose, pH 7.0, the transformed cells of *P. pastoris* strain GS-115 were collected by centrifugation and suspended in 4.5 L of BMMY medium, which consisted of 1% yeast extract, 1.34% yeast nitrogen base, 2% peptone, 0.1 M sodium phosphate buffer, pH 7.0, 0.07% EDTA, 0.01% adecanol, and 1% MeOH, followed by shaking at 28 °C for 24 h to induce the expression and secretion of proCPY into the culture medium.

The culture medium was centrifuged at 24,000*g* for 20 min to remove the yeast cells and the supernatant was concentrated to 1/10 volume with a Millipore Pellicon-2 ultrafiltration system (Massachusetts, USA) using a 10,000 M_r cutoff polyether sulfone membrane. The concentrated medium was applied on a Tosoh Butyl-Toyopearl column (3 × 10 cm) (Tokyo, Japan) equilibrated with 10 mM sodium phosphate buffer, pH 7.0. ProCPY was eluted by the same buffer containing 50 mM NaCl. After reducing the NaCl concentration of the pooled eluate by ultrafiltration, the protein solution was loaded onto a Poros HQ/M (4.6 × 100 mm) column of the PerSeptive Biosystems BioCAD SPRINT system (Massachusetts, USA), equilibrated with 10 mM potassium phosphate buffer, pH 7.0, and eluted with a linear gradient of KCl (0–2 M).

3.4. Activation of proCPY and the activity assay of CPY

proCPY (100 μL, approximately 10 μg) were mixed with 10 μL of proteinase K (5 μg) in 50 mM MES buffer, pH 6.5, containing 20 mM $CaCl_2$ at 25 °C for 1 h.[1] Activated enzyme (100 μL) were mixed with 0.8 mL of 0.1 M sodium phosphate buffer, pH 7.0, and pre-incubated at 37 °C for 5 min. BT*p*NA (100 μL of 3 mM) was added to the mixture and the increase in A_{410} was measured at 37 °C for 5 min.

3.5. CD spectroscopy

CD spectra were measured from 190 to 250 nm using 0.1 cm light path photocell at 25 °C on a Jasco J-720W spectrophotometer (Tokyo, Japan). The thermal stability of proCPY and CPY was measured by monitoring the CD at 222 nm in 20 mM MES buffer, pH 7.0, containing 0.1 M KCl. Temperature was elevated from 25 to 75 °C at a rate of 0.5 °C/min.

3.6. Dynamic light scattering

The dynamic light scattering of protein solutions at 20 °C was measured using a DynaPro99 of Protein Solutions Inc. (Lakewood, USA). Protein was dissolved

in 0.1 M buffer (sodium acetate for pH 4.0 and 4.5, sodium citrate for pH 5.0 and 5.5, MES for pH 6.0 and 6.5, imidazole for pH 7.0, HEPES for pH 7.5 and 8.0, Tris for pH 8.5 and 9.0). The data were analyzed by the Dynals software program.

3.7. Polyacrylamide gel electrophoresis

SDS-PAGE with a separating gel (10% acrylamide in 0.38 M Tris–HCl, pH 8.8) and a stacking gel (4% acrylamide in 0.38 M Tris–HCl, pH 8.8) was performed at 200 V and 25 °C in 25 mM Tris–HCl containing 192 mM glycine, and 0.1% SDS. Native-PAGE with a separating gel (10% acrylamide in 0.5 M Tris–HCl buffer, pH 7.5) and a stacking gel (4% acrylamide in 0.4 M Tris–phosphate buffer, pH 5.5) was performed at 100 V and 4 °C in 0.01 M Tris–barbital buffer, pH 7.0 for 3 h. Protein was stained with Coomassie Brilliant Blue.

3.8. Analysis of Con A affinity

A Con A-agarose column (2 × 100 mm) was washed with 0.1 M boric acid and equilibrated with 20 mM Tris–HCl, pH 7.5, containing 0.5 M NaCl, 1 mM $CaCl_2$, and 1 mM $MnCl_2$. Protein, dissolved in the same buffer, was loaded onto the column and eluted with the washing buffer, followed by the elution buffer containing 0.5 M α-methyl-D-mannoside from fraction no 21, while the absorbance was monitored at 280 and 214 nm. The flow rate was 0.5 mL/min.

3.9. Mass spectrometry

MALDI-TOF mass spectrometry was performed by a PerSeptive Biosystems Voyager RP mass spectrometer (California, USA). A protein solution, containing 30% acetonitrile, 10 mg/mL of 3,5-dimethoxy-4-hydroxycinnamic acid, and 0.3% trifluoroacetic acid, was dried on the sample plate and ionized by laser desorption. The data were calculated by PerSeptive Biosystems GRAMS/386.

3.10. Determination of carbohydrate content

Sample solution (100 μL) was mixed with 100 μL of 5% aqueous phenol solution and immediately agitated thoroughly by means of a vortex mixer. Five hundreds microliters of concd H_2SO_4 was added and the mixture was incubated at room temperature for 10 min under gentle shaking, followed by a further calm incubation for 20 min. After the incubation, the absorbance at 490 nm was measured. This permitted the detection of 1–10 μg glucose.

3.11. Mild alkaline treatment

Two hundred microliters of sample solution were mixed with 800 μL of 0.17 M KOH in a 5:2:1 (v/v) mixture of dimethyl sulfoxide–water–ethanol and incubated for 60 min at 45 °C. The reaction was stopped by an addition of 975 μL of 0.17 M HCl. This reaction mixture was applied onto a Con A-agarose column (2 × 100 mm) and protein fractions were collected for carbohydrate analysis by the phenol–sulfuric acid method as previously described (3.10. Determination of carbohydrate content).

Acknowledgements

The authors express their sincere gratitude to Dr. Kenji Yamamoto, Graduate School of Biostudies, Kyoto University, for the carbohydrate analysis, and Drs. Yasuo Hata and Tomomi Fujii, Institute for Chemical Research, Kyoto University, for the dynamic light scattering measurement. The authors owe Dr. Giman Jung, Graduate School of Agriculture, Kyoto University, for his initiation work on the construction of the expression plasmid.

References

1. Winther, J. R.; Sorensen, P. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9330–9334.
2. Ramos, C.; Winther, J. R.; Kielland-Brandt, M. C. *J. Biol. Chem.* **1994**, *269*, 7006–7012.
3. Aibara, S.; Hayashi, R.; Hata, T. *Agric. Biol. Chem.* **1971**, *35*, 658–666.
4. Sorensen, S. B.; Raaschou-Nielsen, M.; Mortensen, U. H.; Remington, S. J.; Breddam, K. *J. Am. Chem. Soc.* **1995**, *117*, 5944–5950.
5. Duman, J. G.; Miele, R. G.; Liang, H.; Grella, D. K.; Sim, K. L.; Castellino, F. J.; Bretthauer, R. K. *Biotechnol. Appl. Biochem.* **1998**, *28*, 39–45.
6. Chatani, E.; Tanimizu, N.; Ueno, H.; Hayashi, R. *Biosci. Biotechnol. Biochem.* **2000**, *64*, 2437–2444.
7. Stevens, T. H.; Rothman, J. H.; Payne, G. S.; Schekman, R. *J. Cell. Biol.* **1986**, *102*, 1551–1557.
8. Bernard, B. A.; Newton, S. A.; Olden, K. *J. Biol. Chem.* **1983**, *258*, 12198–12202.
9. Ballou, L.; Hernandez, L. M.; Alvarado, E.; Ballou, C. E. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3368–3372.