Available online at www.sciencedirect.com

SCIENCE DIRECT

Matrix Biology 24 (2005) 59–68

MATRIX BIOLOGY

www.elsevier.com/locate/matbio

High-level production of human collagen prolyl 4-hydroxylase in *Escherichia coli*

Antje Neubauer[a], Peter Neubauer[b], Johanna Myllyharju[a,*]

[a]*Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, P. O. Box 5000, University of Oulu, FIN-90014 Oulu, Finland*

[b]*Bioprocess Engineering Laboratory, Department of Process and Environmental Engineering, University of Oulu, Oulu, Finland*

Received 12 October 2004; received in revised form 23 November 2004; accepted 23 November 2004

Abstract

The collagen prolyl 4-hydroxylases (C-P4Hs), enzymes residing within the lumen of the endoplasmic reticulum, play a central role in the synthesis of all collagens. The vertebrate enzymes are $\alpha_2\beta_2$ tetramers in which the two catalytic sites are located in the α subunits, and protein disulfide isomerase serves as the β subunit. All attempts to assemble an active C-P4H tetramer from its subunits in in vitro cell-free systems have been unsuccessful, but assembly of a recombinant enzyme has been reported in several cell types by coexpression of the two types of subunit. An active type I C-P4H tetramer was obtained here by periplasmic expression in *Escherichia coli* strains BL21 and RB791. Further optimization for production by stepwise regulated coexpression of its subunits in the cytoplasm of a thioredoxin reductase and glutathione reductase mutant *E. coli* strain resulted in large amounts of human type I C-P4H tetramer. The specific activity of the C-P4H tetramer purified from the cytoplasmic expression was within the range of values reported for human type I C-P4H isolated as a nonrecombinant enzyme or produced in the endoplasmic reticulum of insect cells, but the expression level, about 25 mg/l in a fermenter, is about 5–10 times that obtained in insect cells. The enzyme expressed in *E. coli* differed from those present in vivo and those produced in other hosts in that it lacked the N glycosylation of its α subunits, which may be advantageous in crystallization experiments.

Keywords: Collagen; Prolyl 4-hydroxylase; *Escherichia coli*

1. Introduction

The collagen prolyl 4-hydroxylases (C-P4Hs, EC 1.14.11.2), enzymes residing within the lumen of the endoplasmic reticulum (ER), catalyze the formation of 4-hydroxyproline in collagens and more than 20 other proteins with collagen-like sequences. The hydroxylation of proline residues in the Y positions of the collagenous -Gly-X-Y- sequences is essential for the generation of stable collagen triple helices at physiological temperatures (Myllyharju and Kivirikko, 2004). The vertebrate C-P4Hs are $\alpha_2\beta_2$ tetramers in which the two catalytic sites are located in the α subunits and the β subunit is identical to the multifunctional enzyme and chaperone protein disulfide isomerase (PDI; Kivirikko and Myllyharju, 1998; Kivirikko and Pihlajaniemi, 1998; Myllyharju, 2003). Three α subunit isoforms characterized in vertebrates have been shown to form $[\alpha(I)]_2\beta_2$, $[\alpha(II)]_2\beta_2$ and $[\alpha(III)]_2\beta_2$ tetramers called type I, II and III C-P4Hs, respectively (Helaakoski et al., 1989, 1995; Annunen et al., 1997; Kukkola et al., 2003; Van Den Diepstraten et al., 2003). The type I enzyme is the major form in most cell types and tissues, but type II is the main form in chondrocytes and endothelial cells, while type III is

Abbreviations: C-P4H, collagen prolyl 4-hydroxylase; PDI, protein disulfide isomerase; BiP, immunoglobulin heavy chain binding protein; ER, endoplasmic reticulum; aTc, anhydrotetracycline; IPTG, isopropyl-β-D-thiogalactoside; DTT, dithiothreitol.

* Corresponding author. Tel.: +358 8 537 5740; fax: +358 8 537 5811.
E-mail address: johanna.myllyharju@oulu.fi (J. Myllyharju).

0945-053X/$ - see front matter © 2004 Elsevier B.V./International Society of Matrix Biology. All rights reserved.
doi:10.1016/j.matbio.2004.11.004

expressed in several tissues, but at much lower levels (Annunen et al., 1998; Nissi et al., 2001; Kukkola et al., 2003; Van Den Diepstraten et al., 2003).

All attempts to assemble an active C-P4H tetramer from its subunits in in vitro cell-free systems have been unsuccessful (Kivirikko and Myllyharju, 1998; Kivirikko and Pihlajaniemi, 1998), but assembly of an active recombinant human C-P4H tetramer has been reported in various cell types by coexpression of the two types of subunit (Vuori et al., 1992a; Vuorela et al., 1997; John et al., 1999; Toman et al., 2000; Merle et al., 2002). This has enabled detailed studies to be made of the catalytic properties of the recombinant enzyme (Lamberg et al., 1995; Myllyharju and Kivirikko, 1997) and identification and characterization of its peptide substrate-binding domain (Myllyharju and Kivirikko, 1999; Hieta et al., 2003; Pekkala et al., 2004).

Site-directed mutagenesis has shown that two intrachain disulfide bonds in the catalytic α subunit are essential for assembly of the C-P4H tetramer (John and Bulleid, 1994; Lamberg et al., 1995), whereas the functioning of PDI as the β subunit is not dependent on its disulfide isomerase activity, as a mutant PDI polypeptide in which both -Cys-Gly-His-Cys- catalytic sites have been inactivated by mutation to -Ser-Gly-His-Cys- forms a fully active enzyme tetramer (Vuori et al., 1992b). The main function of PDI in the C-P4H tetramer is to keep the highly insoluble α subunits in a catalytically active, nonaggregated conformation (Kivirikko and Myllyharju, 1998; Kivirikko and Pihlajaniemi, 1998; Myllyharju, 2003). PDI has a similar role in the microsomal triglyceride transfer protein dimer, so that this property may be related to its chaperone function (Wetterau et al., 1990, 1991).

Many recombinant proteins have been shown to fold correctly in the periplasm of *Escherichia coli* (Bessette et al., 1999; Swartz, 2001). Novel mutant *E. coli* strains with a more oxidizing cytoplasm have recently been developed, and model proteins with up to 17 disulfide bonds have now been successfully expressed in such strains (Bessette et al., 1999). We report here that large amounts of a fully active recombinant human C-P4H tetramer can be assembled in the cytoplasm of a thioredoxin reductase and glutathione reductase double mutant *E. coli* strain by coexpression of the two types of subunit. Cytoplasmic expression resulted in higher amounts of active tetramer than were obtained in periplasmic expression. Coexpression of the chaperone immunoglobulin heavy chain binding protein (BiP), which is known to form soluble α subunit–BiP complexes as intermediates in the synthesis of the enzyme tetramer within the ER of mammalian cells (John and Bulleid, 1996), did not increase the amount of C-P4H tetramer in the *E. coli* periplasmic or cytoplasmic fractions.

2. Results and discussion

2.1. Expression of recombinant human type I C-P4H in the periplasm of E. coli

The C-P4H tetramer is a soluble ER luminal protein, the two intrachain disulfide bonds in the catalytic α subunit being essential for its assembly (John and Bulleid, 1994; Lamberg et al., 1995). Recombinant expression of active human type I C-P4H in *E. coli* was obtained using a plasmid encoding the α(I) and PDI polypeptides possessing periplasmic signal sequences. A cDNA encoding the human α(I) subunit was cloned into the vector pASK-IBA2 under control from the *tetA* promoter in-frame with the *ompA* signal sequence, and the PDI cDNA was cloned into the same vector downstream of the α subunit cDNA under control from the T*5lac* promoter in-frame with the *ompA* signal sequence to generate the expression construct pP4Hper (Fig. 1B). The recombinant pP4Hper plasmid was then transformed into the *E. coli* strains RB791 and BL21, and the cells were grown in shaker flasks at 25 °C in a modified LB medium supplemented with ampicillin. PDI

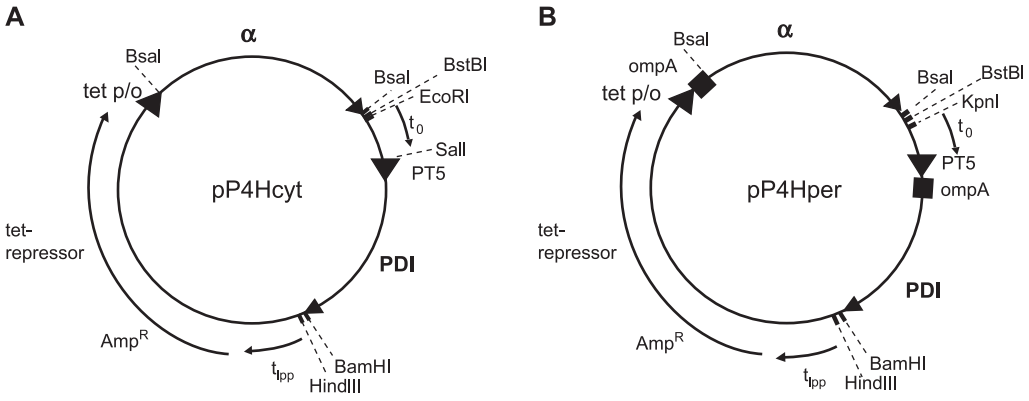

Fig. 1. Construction of the pP4Hcyt (A) and pP4Hper (B) expression plasmids. A human α subunit cDNA without the native signal sequence was cloned under control from the *tetA* promoter (tet p/o) into a pASK-IBA3 expression vector (A) or into a pASK-IBA2 expression vector with the *ompA* signal sequence (B), and the transcription terminator t_0 was cloned downstream of the α subunit cDNA. A PDI cDNA without the native signal sequence was cloned under control from the T*5lac* promoter (PT5) downstream from the t_0, the PDI cDNA in *B* having the *ompA* signal sequence. The PDI cDNAs are followed by the transcription terminator t_{lpp}.