Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast *Pichia pastoris*: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase

Annamari Vuorela, Johanna Myllyharju, Ritva Nissi, Taina Pihlajaniemi and Kari I.Kivirikko¹

Collagen Research Unit, Biocenter and Department of Medical Biochemistry, University of Oulu, Kajaanintie 52 A, FIN-90220 Oulu, Finland

¹Corresponding author

Prolyl 4-hydroxylase, the key enzyme of collagen synthesis, is an $\alpha_2\beta_2$ tetramer, the β subunit of which is protein disulfide isomerase (PDI). Coexpression of the human α subunit and PDI in *Pichia* produced trace amounts of an active tetramer. A much higher, although still low, assembly level was obtained using a *Saccharomyces* pre-pro sequence in PDI. Coexpression with human type III procollagen unexpectedly increased the assembly level 10-fold, with no increase in the total amounts of the subunits. The recombinant enzyme was active not only in *Pichia* extracts but also inside the yeast cell, indicating that *Pichia* must have a system for transporting all the cosubstrates needed by the enzyme into the lumen of the endoplasmic reticulum. The 4-hydroxyproline-containing procollagen polypeptide chains were of full length and formed molecules with stable triple helices even though *Pichia* probably has no Hsp47-like protein. The data indicate that collagen synthesis in *Pichia*, and probably also in other cells, involves a highly unusual control mechanism, in that production of a stable prolyl 4-hydroxylase requires collagen expression while assembly of a stable collagen requires enzyme expression. This *Pichia* system seems ideal for the high-level production of various recombinant collagens for numerous scientific and medical purposes.

Keywords: chaperone/collagen/expression/prolyl 4-hydroxylase/yeast

Introduction

The collagens are a family of extracellular matrix proteins that contain the repeating triplet sequence -Gly-X-Y- in which the Y position amino acid is often 4-hydroxyproline, and have the potential for three polypeptide chains with such sequences to fold into triple-helical molecules. At least 19 proteins representing more than 30 gene products have now been defined as collagens, and more than 10 additional proteins have collagen-like domains. Many of the recently discovered collagens are present in tissues in such small amounts that they have so far been characterized mainly or only at the cDNA rather than the protein level (for reviews on collagens, see Kielty *et al.*, 1993; Mayne

and Brewton, 1993; Pihlajaniemi and Rehn, 1995; Prockop and Kivirikko, 1995). It is thus obvious that an efficient expression system for recombinant collagens would have numerous applications. Nevertheless, most recombinant systems now available for the production of proteins cannot be used as such for the expression of recombinant collagens, as bacteria and yeast have no prolyl 4-hydroxylase activity and insect cells have insufficient levels of it (Lamberg *et al.*, 1996).

Prolyl 4-hydroxylase (EC 1.14.11.2) plays a central role in the synthesis of all collagens, as 4-hydroxyproline residues are essential for the formation of molecules with stable triple helices. The vertebrate enzyme is an $\alpha_2\beta_2$ tetramer (see Kivirikko *et al.*, 1989, 1992; Helaakoski *et al.*, 1995), the β subunit of which is identical to another enzyme, protein disulfide isomerase (PDI, EC 5.3.4.1) (Koivu *et al.*, 1987; Pihlajaniemi *et al.*, 1987). The role of PDI in the enzyme is not related to its disulfide isomerase activity, as a double mutant PDI polypeptide in which both -Cys-Gly-His-Cys- catalytic sites (Vuori *et al.*, 1992a; Freedman *et al.*, 1994) have been inactivated to -Ser-Gly-His-Cys- forms a fully active prolyl 4-hydroxylase tetramer (Vuori *et al.*, 1992c). The main function of the PDI polypeptide in prolyl 4-hydroxylase appears to be to keep the highly insoluble α subunits in a catalytically active, non-aggregated conformation, a function that resembles the role of certain chaperones such as Hsp90 in other proteins (Vuori *et al.*, 1992b; John *et al.*, 1993; Freedman *et al.*, 1994). All attempts to construct an active prolyl 4-hydroxylase tetramer from its subunits *in vitro* have been unsuccessful, but an active recombinant tetramer has been obtained by coinfection with two baculoviruses in insect cells (Vuori *et al.*, 1992b) or cotransfection in COS-1 cells (John and Bulleid, 1996). Tetramer assembly probably requires molecular chaperones, one such chaperone being BiP (John and Bulleid, 1996; Veijola *et al.*, 1996b). Prolyl 4-hydroxylase requires Fe^{2+} , 2-oxoglutarate, O_2 and ascorbate, and an active system appears to exist in vertebrate cells for the transport of 2-oxoglutarate and ascorbate into the lumen of the endoplasmic reticulum (see Kivirikko *et al.*, 1989, 1992).

Yeast would seem to be an ideal system for studying the various assembly steps in collagen synthesis and for the recombinant production of collagens for various scientific and medical purposes. Yeast has no prolyl 4-hydroxylase or collagen synthesis, and thus provides a system in which endogenous prolyl 4-hydroxylase plays no role and in which the recombinant collagen is not contaminated by any non-recombinant collagen. Nevertheless, studies were required to determine whether yeast would contain the appropriate chaperones and other conditions required for the assembly of an active recombinant

Table I. Expression vectors used to generate the various *Pichia* strains

Expression vectors	Strain	Selection	Polypeptides expressed
pARG815 α	α	Arg ⁺	4-PH ^a α subunit
pAO815PDI	PDI	His ⁺	PDI
pARG815 α , pAO815PDI	α /PDI	Arg ⁺ , His ⁺	4-PH α subunit, PDI
pARG815 α PDI, pAO815PDI	α PDI/PDI	Arg ⁺ , His ⁺	4-PH α subunit, PDI
pARG815 α , pHIL-S1PDI	α /PDI-PHO1	Arg ⁺ , His ⁺	4-PH α subunit, PDI with PHO1 ^a signal sequence
pARG815 α , pPIC9PDI	α /PDI- α MF	Arg ⁺ , His ⁺	4-PH α subunit, PDI with α MF ^a prepropeptide
pARG815 α , pPIC9PDIHDEL	α /PDI- α MF-HDEL	Arg ⁺ , His ⁺	4-PH α subunit, PDI with α MF prepropeptide and HDEL C-terminus
pARG815 α PHO1, pPIC9PDI	α -PHO1/PDI- α MF	Arg ⁺ , His ⁺	4-PH α subunit with PHO1 signal sequence, PDI with α MF prepropeptide
pHIL-D2pro α 1(III)	Pro α 1(III)	His ⁺	pro α 1 chain of type III procollagen
pARG815 α , pPIC9PDI, pPICZ Bpro α 1(III)	α /PDI- α MF/pro α 1(III)	Arg ⁺ , His ⁺ , Zeocin resistance	4-PH α subunit, PDI with α MF prepropeptide, pro α 1 chain of type III procollagen

^a4-PH, prolyl 4-hydroxylase; PHO1, acid phosphatase 1; α MF, α mating factor.

prolyl 4-hydroxylase tetramer and whether a system of transport into the lumen of the yeast endoplasmic reticulum would be available for all the cosubstrates required by the enzyme, and thus whether the enzyme would be active inside the yeast cell. Studies were also required to confirm whether yeast could synthesize the long collagen polypeptide chains with the repetitive -Gly-X-Y- sequences and whether such chains would form triple-helical molecules in yeast. This aspect was of particular interest as Hsp47 is regarded as a collagen-specific chaperone (Nagata, 1996; Satoh *et al.*, 1996) and as the yeast sequencing project (Goffeau *et al.*, 1996) has indicated that *Saccharomyces cerevisiae* has no Hsp47-like sequences. We report here on the expression and assembly of the subunits of human prolyl 4-hydroxylase and the pro α 1 chains of human type III procollagen in the yeast *Pichia pastoris*. In the course of this work we made the unexpected finding that the production of a stable prolyl 4-hydroxylase tetramer requires the expression of collagen polypeptide chains. This appears to constitute a novel control mechanism that may not be limited to recombinant expression in *Pichia* but may also exist in a number of other cell types.

Results

Coexpression of the human prolyl 4-hydroxylase α subunit and PDI polypeptide in *P.pastoris* produces a small amount of an active enzyme tetramer

The *his4*, *arg4* *P.pastoris* host strain, which has defects in enzymes required for the synthesis of histidine and arginine, was used in these studies. In order to study whether the recombinant human prolyl 4-hydroxylase α subunit and PDI polypeptide are able to form an active enzyme tetramer in yeast cells, cDNAs for the human prolyl 4-hydroxylase α subunit and PDI polypeptide were cloned into the *Pichia* expression vectors pARG815 (complementing for *arg4* in the host) and pAO815 (complementing for *his4* in the host), respectively (Table I). An additional expression vector, pARG815 α PDI, coding for both the α subunit and PDI polypeptide was also generated. The *his4*, *arg4* strain was then transformed separately with pARG815 α and pAO815PDI, or cotransformed with either pARG815 α and pAO815PDI or pARG815 α PDI and pAO815PDI. Arg⁺, methanol utilizing (Mut⁺) transformants of the α subunit strain, His⁺,

Mut⁺ transformants of the PDI strain and Arg⁺, His⁺, Mut⁺ cotransformants of the α /PDI and α PDI/PDI strains were selected. The presence of one copy of each transformed DNA in the α /PDI and α PDI/PDI strains was verified by dot blot analysis (data not shown). Cells were cultured in BMGY medium supplemented with yeast extract and peptone, and expression was induced in BMM medium, methanol being added every 24 h to a final concentration of 0.5%. The cells were harvested 60 h after induction, broken in a buffer containing Triton X-100 and centrifuged. Samples of the soluble and insoluble fractions of the cell lysates were then analysed by SDS-PAGE and non-denaturing PAGE followed by Western blotting with a polyclonal antibody to the human PDI polypeptide, or a monoclonal or polyclonal antibody to the human prolyl 4-hydroxylase α subunit (Figure 1). A band corresponding to the PDI polypeptide was found in the soluble fraction of cell lysates from the α /PDI and α PDI/PDI strains (Figure 1A, lanes 2 and 3). The prolyl 4-hydroxylase α subunit could not be detected in the soluble fractions (Figure 1B, lanes 2 and 3), whereas bands corresponding to it and its degradation products were found after treating the insoluble fractions with 1% SDS (Figure 1C, lanes 2 and 3). In non-denaturing PAGE a faint band corresponding to the enzyme tetramer was seen only in the soluble extract from the α PDI/PDI strain, which has two copies of the PDI DNA (Figure 2, lanes 2 and 3).

Prolyl 4-hydroxylase activities of the Triton X-100 soluble extracts were assayed by two methods, one based on the hydroxylation-coupled decarboxylation of 2-oxo[1-¹⁴C]glutarate with a synthetic peptide substrate and the other on the formation of 4-hydroxy[¹⁴C]proline in a [¹⁴C]proline-labelled procollagen substrate. The enzyme activities in the soluble extracts from the untransformed *his4*, *arg4* host strain before and after methanol induction and from the strains expressing only the α subunit or the PDI polypeptide never exceeded the background values. The enzyme activity level in the extract from the α /PDI strain was very low and could only be detected by the more sensitive method using the [¹⁴C]proline-labelled procollagen as a substrate (Table II). Prolyl 4-hydroxylase activity was four times higher in the extract from the α PDI/PDI strain than in that from the α /PDI strain, but was still less than 200 d.p.m./200 μ g soluble protein when the method based on the hydroxylation-

Fig. 1. Expression of the human prolyl 4-hydroxylase α subunit and PDI polypeptide in *P.pastoris*. (A) The *his, arg* *Pichia* host strain (lane 1) and its transformant strains α /PDI (lane 2), α PDI/PDI (lane 3), α /PDI-PHO1 (lane 4), α /PDI- α MF (lane 5) and α -PHO1/PDI- α MF (lane 6) were cultured in BMGY medium, and expression of recombinant proteins was induced by the addition of methanol. Cells were harvested 60 h after induction and broken in a buffer containing Triton X-100, and samples of the soluble fraction of the lysates were electrophoresed on 10% SDS-PAGE under reducing conditions, followed by Western blotting with a polyclonal antibody K38 to human PDI. The *Pichia* strains are identified at the top, and the position of the PDI polypeptide is indicated by an arrow. It may be noted that the PDI- α MF gives several bands as shown in detail in Figure 4. (B) As (A), but analysed with a monoclonal antibody M95L to the human α subunit. The positions of the variably glycosylated forms of the α subunit are indicated by an arrow. (C) As (A), but insoluble fractions of the lysates were analysed with a polyclonal antibody K17 to the human α subunit. The positions of the variably glycosylated forms of the α subunit are indicated by an arrow. Bands corresponding to various degradation products of the α subunit are also seen.

coupled decarboxylation of 2-oxo[1- 14 C]glutarate was used (Table II). This value is at least two orders of magnitude lower than the activity level obtained in insect cells by coinfection with two separate baculoviruses (Vuori et al., 1992b; Veijola et al., 1994).

Replacement of the signal peptide of human PDI with a yeast pre-pro sequence markedly increases both prolyl 4-hydroxylase tetramer assembly and PDI secretion

As the assembly of an active prolyl 4-hydroxylase tetramer from the inactive monomers was poor in *Pichia*, the effect of replacing the signal sequences of the α subunit and PDI polypeptide with yeast signal sequences was studied. The *Pichia* cotransformants (Table I) α /PDI-PHO1 (in which PHO1 is the *P.pastoris* acid phosphatase 1 signal sequence), α /PDI- α MF and α -PHO1/PDI- α MF (α MF is the *S.cerevisiae* α mating factor pre-pro sequence) were cultured, induced and harvested as above. The presence of only one copy of the α subunit and PDI DNAs was verified by dot blot analysis. Samples of the soluble and insoluble fractions of cell lysates were analysed by SDS-PAGE as above. Bands corresponding to the PDI polypeptide were again seen in the extracts from all three strains (Figure 1A, lanes 4–6). Faint bands corresponding to the α subunit were seen in the soluble fractions from the α /PDI-PHO1 and α -PHO1/PDI- α MF strains (Figure 1B, lanes 4 and 6), whereas slightly stronger α subunit bands were seen in the soluble fraction from the α /PDI- α MF strain (Figure 1B, lane 5). Nevertheless, strong bands corresponding to the α subunit and its degradation products were seen in the 1% SDS extract from the insoluble fraction in all three strains (Figure 1C, lanes 4–6). Non-denaturing PAGE analysis of the soluble extracts from all three strains gave a distinct band that could be immunostained with the polyclonal α subunit antibody and corresponded to the prolyl 4-hydroxylase tetramer (as shown for the α /PDI- α MF strain in Figure 2, lane 4).

Fig. 2. Formation of a prolyl 4-hydroxylase tetramer in *P.pastoris*. The *his, arg* host strain (lane 1) and the α /PDI (lane 2), α PDI/PDI (lane 3) and α /PDI- α MF (lane 4) strains were cultured, induced and harvested as in Figure 1. Samples of the soluble fraction of the lysates were analysed by 8% PAGE under non-denaturing conditions followed by Western blotting with a polyclonal antibody K17 to the human α subunit. Strains are identified at the top, and the prolyl 4-hydroxylase tetramer is indicated by the arrow $\alpha_2\beta_2$.

Soluble extracts from the α /PDI-PHO1, α /PDI- α MF and α -PHO1/PDI- α MF strains were analysed for prolyl 4-hydroxylase activity with the assay based on the hydroxylation-coupled decarboxylation of 2-oxo[1- 14 C]glutarate. In accordance with the strongest prolyl 4-hydroxylase tetramer band observed in non-denaturing PAGE, the highest prolyl 4-hydroxylase activity level, ~2600 d.p.m./200 μ g, was found in the extract from the α /PDI- α MF strain, while the levels in the extracts from the α /PDI-PHO1 and α -PHO1/PDI- α MF strains were distinctly lower, ~1100 d.p.m./200 μ g and 300 d.p.m./200 μ g, respectively (Table II). Thus the best signal sequences for prolyl 4-hydroxylase tetramer production among those studied here were the authentic one in the case of the α subunit and the *S.cerevisiae* α MF pre-pro sequence in the case of the PDI polypeptide.

Table II. Enzyme activity levels in various *Pichia* strains expressing prolyl 4-hydroxylase

Strain	Prolyl 4-hydroxylase activity (d.p.m./200 µg) ^a	
	Protocollagen assay	2-Oxoglutarate assay ^b
<i>his, arg</i> (uninduced)	<20	<50
<i>his, arg</i> (induced)	<20	<50
α	<20	<50
PDI	<20	<50
α /PDI	160	<50
α PDI/PDI	650	~200
α /PDI-PHO1	n.d. ^c	1060 $\pm$ 180
α /PDI- α MF	n.d.	2600 $\pm$ 350
α -PHO1/PDI- α MF	n.d.	320 $\pm$ 70

^aAbout 200 µg of extractable cell protein was used in each assay, and therefore the values are given as d.p.m./200 µg of protein. The values obtained with the 2-oxoglutarate assay are given as mean $\pm$ SD of three to seven independent experiments.

^bThe assay based on the formation of hydroxy[¹⁴C]proline in a [¹⁴C]proline-labelled protocollagen substrate is more sensitive than that based on the hydroxylation-coupled decarboxylation of 2-oxo[1-¹⁴C]glutarate.

^cn.d. = not determined.

Fig. 3. Secretion of PDI in *P.pastoris*. The α PDI/PDI (lane 1), α PDI-PHO1 (lanes 2 and 3), α PDI- α MF (lane 4) and α PDI- α MF-HDEL (lane 5) strains were cultured and induced as in Figure 1. The cell culture medium was harvested 60 h after induction, and unconcentrated (lanes 1, 2, 4 and 5) or concentrated (lane 3) samples were analysed by 10% SDS-PAGE followed by Western blotting using a polyclonal antibody K38 to human PDI. Strains are identified at the top, and the PDI polypeptide is indicated by an arrow.

Secretion of prolyl 4-hydroxylase and its subunits into the culture medium was studied by analysing samples of the minimal medium by SDS-PAGE followed by Coomassie staining or Western blotting using antibodies to the α subunit and the PDI polypeptide. The α subunit could never be detected in any cell culture medium sample (data not shown). The PDI polypeptide band was seen in unconcentrated medium samples only in the case of strains expressing PDI with the α MF pre-pro sequence (Figure 3, lanes 1, 2 and 4). The amount of PDI in the culture medium, based on Coomassie staining, was ~10 mg/l (data not shown). No differences were found in the cellular PDI levels or PDI secretion levels between strains having a KDEL or HDEL endoplasmic reticulum retention signal (Pelham, 1990) in the PDI- α MF polypeptide (Figure 3, lanes 4 and 5). PDI with the PHO1 signal sequence was much less efficiently secreted and could only be detected

Fig. 4. Analysis of purified recombinant human prolyl 4-hydroxylase tetramers. (A) Recombinant human prolyl 4-hydroxylases from *Sf9* insect cells (lane 1) and the α PDI-PHO1 (lane 2) and α PDI- α MF (lane 3) *Pichia* strains were affinity-column purified, and samples analysed by 8% PAGE under non-denaturing conditions followed by Coomassie staining. The enzymes are identified at the top, and the prolyl 4-hydroxylase tetramer is indicated by the arrow $\alpha_2\beta_2$. (B) Purified recombinant human prolyl 4-hydroxylases from *Sf9* insect cells (lane 1) and the α PDI- α MF (lane 2) *Pichia* strain were analysed by 8% SDS-PAGE under reducing conditions followed by Coomassie staining. The enzymes are identified at the top, and the locations of the PDI polypeptide, the PDI polypeptide containing the α MF propeptide (α MFpro) and the various glycosylated forms of the α subunit are indicated by arrows.

in Western blots of medium samples concentrated 30 times (Figure 3, lane 3), while PDI with its own signal sequence was never detected even in the concentrated samples. Determination of the N-terminal amino acid sequence of PDI partially purified from the culture medium of the α PDI- α MF strain demonstrated that the α MF pre-pro sequence had been cleaved as expected, resulting in an N-terminal sequence Tyr-Val-Glu-Phe-Asp-Ala-, in which the first four amino acids originate from the pPIC9 plasmid and the next two correspond to the N terminus of human PDI.

Characterization of human prolyl 4-hydroxylase tetramers expressed in *Pichia*

The α PDI-PHO1 and α PDI- α MF strains were cultured, induced and broken as above, and prolyl 4-hydroxylase was purified using a poly(L-proline) affinity column and gel filtration. Non-denaturing PAGE analysed by Coomassie staining showed no differences in migration between the recombinant human prolyl 4-hydroxylases purified from the α PDI-PHO1 *Pichia* strain and from the *Spodoptera frugiperda* (*Sf9*) insect cells (Figure 4A, lanes 1 and 2). N-terminal sequencing of the PDI polypeptide of the purified enzyme tetramer showed that the PHO1 signal peptide had been correctly cleaved, resulting in an N-terminal sequence Arg-Glu-Phe-Asp-Ala-Pro-Glu-Glu-, in which the first three amino acids originate from the pHIL-S1 plasmid.

Prolyl 4-hydroxylase purified from the α PDI- α MF strain migrated in non-denaturing PAGE as a broad band with a slightly slower mobility than the enzyme from the α PDI-PHO1 strain (Figure 4A, lane 3). In Coomassie stained SDS-PAGE several bands corresponding to polypeptides of ~70 kDa were seen in addition to the α subunit and PDI polypeptide bands (Figure 4B, lane 2). These were stained by the PDI antibody in Western blotting in addition to the 60 kDa PDI polypeptide band (Figure 5,

Fig. 5. Analysis of carbohydrate moieties of the prolyl 4-hydroxylase α subunit and the PDI- α MFpro polypeptide. Samples of prolyl 4-hydroxylase purified from the α /PDI- α MF strain were incubated in the absence (lanes 1 and 3) or presence (lanes 2 and 4) of *N*-glycosidase F and analysed by 10% SDS-PAGE followed by Western blotting with a polyclonal antibody K38 to the human PDI polypeptide (lanes 1 and 2) or a polyclonal antibody K17 to the α subunit (lanes 3 and 4). The arrows indicate the glycosylated forms of the PDI- α MFpro and PDI polypeptides, and the various glycosylated forms of the α subunit.

lane 1), suggesting that the PDI polypeptides concerned still contained the α MF propeptide. All these additional bands gave a homogeneous N-terminal sequence Ala-Pro-Val-Asn-Thr-, indicating that the prepeptide of the α MF pre-propeptide had been correctly cleaved while the propeptide was still attached to PDI. A single additional band of ~70 kDa remained after *N*-glycosidase F treatment (Figure 5, lane 2), indicating that the multiple bands were due to differentially glycosylated forms of the α MF propeptide. The human PDI polypeptide itself has no *N*-glycosylation sites (Pihlajaniemi *et al.*, 1987), and no heterogeneity was seen in the PDI polypeptide of the enzyme purified from the α /PDI-PHO1 strain (data not shown).

Western blotting using the polyclonal α subunit antibody indicated that three forms of the α subunit are present in prolyl 4-hydroxylases purified from the α /PDI-PHO1 (data not shown) and α /PDI- α MF strains (Figure 5, lane 3). *N*-glycosidase F treatment produced a single polypeptide with a mobility similar to that of the lowest of the three bands present in the untreated sample (Figure 5, lane 4), indicating the presence of two differentially glycosylated forms of the α subunit. This finding agrees with the presence of two potential *N*-glycosylation sites in the human α subunit (Helaakoski *et al.*, 1989). N-terminal sequencing of the α subunit of the enzyme purified from the α /PDI- α MF strain showed that the signal peptide had been correctly cleaved.

The specific activities of the recombinant enzymes purified from the α /PDI-PHO1 and α /PDI- α MF strains and from insect cells were essentially identical. The K_m values for Fe^{2+} , 2-oxoglutarate, ascorbate and the peptide substrate were likewise identical (data not shown).

Coexpression of collagen polypeptide chains markedly increases the amount of prolyl 4-hydroxylase tetramer

In order to study whether it is possible to produce recombinant human type III procollagen with a stable triple helix in *Pichia*, a construct termed pPICZ Bpro α 1(III) was transformed into the α /PDI- α MF strain (Table I), which

Table III. Prolyl 4-hydroxylase activity and expression of type III procollagen in various α /PDI- α MF/pro α 1(III) *Pichia* transformants

Strain	Prolyl 4-hydroxylase activity (d.p.m./200 μ g) ^b	Expression of type III procollagen ^a (ng/200 μ g)
α /PDI- α MF	2 300	0
α /PDI- α MF/pro α 1(III) 1	20 920	380
α /PDI- α MF/pro α 1(III) 2	19 560	430
α /PDI- α MF/pro α 1(III) 3	2 690	0
α /PDI- α MF/pro α 1(III) 4	22 100	340
α /PDI- α MF/pro α 1(III) 5	22 410	410

^aExpression of type III procollagen was measured by a radioimmunoassay for the *N*-propeptide of human type III procollagen.

^bValues are given as d.p.m./200 μ g of extractable cell protein.

had expressed the highest level of prolyl 4-hydroxylase activity. Cells were cultured, induced, harvested and broken as above. To make sure that the cells still expressed active prolyl 4-hydroxylase, the enzyme activity was measured in the soluble fraction of the cell lysates using the assay based on the hydroxylation-coupled decarboxylation of 2-oxo[1-¹⁴C]glutarate. Other aliquots of this fraction and samples of the cell culture medium were analysed with a radioimmunoassay for the *N*-propeptide of human type III procollagen.

The prolyl 4-hydroxylase activity level in all the transformants but one was ~10 times that in the original α /PDI- α MF strain (Table III). Subsequent radioimmunoassays for the *N*-propeptide of type III procollagen indicated that all the transformants but one expressed essentially identical amounts of type III procollagen (Table III). The transformant that expressed no type III procollagen was the same one that showed no increase in prolyl 4-hydroxylase activity level. This highly surprising finding has now been confirmed in four subsequent experiments.

The presence of only one copy each of the prolyl 4-hydroxylase α subunit, PDI and pro α 1(III) DNAs in all the α /PDI- α MF/pro α 1(III) transformants was verified by dot blot analysis (data not shown). No differences were found in the mRNA levels for the α subunit or the PDI polypeptide between the original α /PDI- α MF strain and any of the α /PDI- α MF/pro α 1(III) transformants when studied by Northern blotting (data not shown).

The soluble fraction and medium samples from the α /PDI- α MF/pro α 1(III) strain were compared with those of the α /PDI- α MF strain in SDS-PAGE followed by Western blotting with the polyclonal PDI antibody or the monoclonal α subunit antibody. Essentially identical amounts of the PDI polypeptide were seen in the soluble fractions and medium samples from both strains (Figure 6, lanes 1, 2, 7 and 8), but the amount of the α subunit was distinctly larger in the soluble fraction from the α /PDI- α MF/pro α 1(III) strain than in that from the α /PDI- α MF strain (Figure 6, lanes 3 and 4). Likewise a much stronger band corresponding to the enzyme tetramer was seen in the soluble fraction from the α /PDI- α MF/pro α 1(III) in non-denaturing PAGE analysis followed by Western blotting with the polyclonal α subunit antibody than in that from the α /PDI- α MF strain (Figure 6, lanes 9 and 10). The amount of the α subunit and its degradation

Fig. 6. Coexpression of polypeptide chains of type III procollagen increases the amount of enzyme tetramer. The α /PDI- α MF (lanes 1, 3, 5, 7 and 9) and α /PDI- α MF/pro α 1(III) (lanes 2, 4, 6, 8 and 10) strains were cultured, induced and harvested as in Figure 1. Samples of the soluble (lanes 1–4) and insoluble (lanes 5 and 6) fractions of the cell lysates and the cell culture medium (lanes 7 and 8) were analysed by 10% SDS–PAGE followed by Western blotting with a polyclonal PDI antibody K38 (lanes 1, 2, 7 and 8) or a monoclonal α subunit antibody M95L (lanes 3–6). Other aliquots of the soluble fractions were analysed by 8% non-denaturing PAGE followed by Western blotting with a polyclonal α subunit antibody K17 (lanes 9 and 10). Locations of the PDI- α MFpro and the PDI polypeptides, the α subunit and the prolyl 4-hydroxylase tetramer ($\alpha_2\beta_2$) are indicated by arrows.

products in the insoluble fraction from the α /PDI- α MF/pro α 1(III) strain was smaller than that in the insoluble fraction from the α /PDI- α MF strain (Figure 6, lanes 5 and 6).

Collagen polypeptide chains coexpressed in *Pichia* with prolyl 4-hydroxylase form molecules with stable triple helices

α /PDI- α MF/pro α 1(III) and Pro α 1(III) strains (Table I) were cultured and induced as above, ascorbate being added to the culture medium every 12 h during the induction. Cells were harvested 60 h after induction and broken in a 5% glycerol, 1 mM EDTA and 50 mM sodium phosphate buffer, pH 7.4. Samples of the soluble fraction of the cell lysates and the cell culture medium were analysed by means of a radioimmunoassay for the N-propeptide of type III procollagen. In both strains the majority of the recombinant human type III procollagen was present in the soluble fraction of the cell lysates, only 10–20% being found in the culture medium (data not shown).

Samples of the soluble fraction of the cell lysates were analysed by SDS–PAGE under reducing and non-reducing conditions followed by Coomassie staining or Western blotting with an antibody to the N-propeptide of type III procollagen, and aliquots of the samples were also studied after digestion with pepsin for 2 h at 22°C. The triple helix of collagens is resistant to pepsin while non-triple-helical pro α 1(III) chains and the propeptides of triple-helical procollagen molecules are digested. Only one band corresponding to full-length pro α 1(III) chains was seen in immunoblots of non-pepsinized samples from both strains when analysed by SDS–PAGE under reducing conditions (as shown for the Pro α 1(III) strain in Figure 7, lane 1). Essentially all the pro α 1(III) chains were found as disulfide-bonded trimers when analysed by SDS–PAGE under non-reducing conditions (as shown for the Pro α 1(III) strain in Figure 7, lane 2). Pepsin-resistant α 1 chains of type III collagen were seen in the digested samples from both strains when analysed by SDS–PAGE under reducing conditions (Figure 7, lanes 3 and 5). However, a substantial fraction of the α 1(III) chains from the Pro α 1(III) strain was digested by pepsin to smaller products and these α 1(III) chains could only be detected by Western blotting

with a monoclonal antibody that recognizes the collagenous regions of various collagen chains (Figure 7, lane 3), whereas essentially all the α 1(III) chains from the α /PDI- α MF/pro α 1(III) strain were of full length and could readily be seen by Coomassie staining (Figure 7, lane 5). Moreover, even the full-length α 1(III) chains from the Pro α 1(III) strain were not disulfide-bonded (Figure 7, lane 4), whereas all the α 1(III) chains from the α /PDI- α MF/pro α 1(III) strain were present as disulfide-bonded trimers when analysed by SDS–PAGE under non-reducing conditions (Figure 7, lane 6).

The thermal stability of the pepsin-resistant type III collagens was studied using digestion with a mixture of trypsin and chymotrypsin after heating to various temperatures (Bruckner and Prockop, 1981). The recombinant type III collagen produced in the Pro α 1(III) strain was completely digested at 27°C (Figure 7, lane 7). In agreement with this result, no 4-hydroxyproline was detected in amino acid analysis of the recombinant collagen purified from this strain (data not shown). The thermal stability of the recombinant type III collagen produced in the α /PDI- α MF/pro α 1(III) strain was much higher, the T_m being ~39°C (Figure 8). The amino acid composition of the recombinant collagen purified from this strain was essentially identical to that of the non-recombinant human type III collagen (data not shown), except that the degree of 4-hydroxylation of the incorporated proline residues (i.e. the percentage of 4-Hyp of the sum of 4-Hyp + Pro) was 44.2%, while the corresponding value for the recombinant human type III collagen produced in insect cells, which had a T_m identical to that of the non-recombinant human type III collagen, was 50.1% (Lamberg *et al.*, 1996) and that for the non-recombinant human type III collagen determined here (Chemicon) was 51.6%. The only other significant difference between the recombinant collagen and the corresponding non-recombinant protein was that the *Pichia*-derived collagen contained no hydroxylysine while the latter has five hydroxylysine residues per 1000 amino acids (Chung and Miller, 1974).

The level of expression of the recombinant type III procollagen in the α /PDI- α MF/pro α 1(III) strain, as estimated from the amount of purified type III collagen obtained

Fig. 7. Expression of recombinant human type III procollagen in *P.pastoris*. The transformant strains Pro α 1(III) and α /PDI- α MF/pro α 1(III) were cultured and induced as in Figure 1, ascorbate being added every 12 h during the induction. Cells were harvested 60 h after induction and broken in a glycerol-phosphate buffer. Aliquots of the soluble fraction of the cell lysates were digested first with pepsin and then with a mixture of trypsin and chymotrypsin. Samples were electrophoresed on 5% SDS-PAGE under reducing (lanes 1, 3, 5 and 7) or non-reducing (lanes 2, 4 and 6) conditions. Undigested samples were analysed by Western blotting with an antibody to the N-propeptide of type III procollagen (lanes 1 and 2). The pepsin-digested samples from the Pro α 1(III) strain were analysed by Western blotting with a monoclonal collagen antibody 95D1A (lanes 3 and 4). The pepsin-digested samples from the α /PDI- α MF/pro α 1(III) strain (lanes 5 and 6) and the sample from the Pro α 1(III) strain digested with a mixture of trypsin and chymotrypsin at 27°C (lane 7) were analysed by Coomassie staining. Strains are identified at the top, and the positions of the trimeric [pro α 1(III)]₃ and [α 1(III)]₃ and the monomeric pro α 1(III) and α 1(III) chains are shown by arrows. The mobilities of these chains were identical to those of the corresponding chains expressed in insect cells.

and the recovery of the purification, was ~15 mg/l (data not shown).

Discussion

The present data indicate that coexpression of the α subunit of human prolyl 4-hydroxylase and the human PDI polypeptide in *P.pastoris* produces a fully active recombinant enzyme tetramer. It is thus obvious that *Pichia* contains appropriate chaperones and other conditions required for the assembly of an active prolyl 4-hydroxylase. The recombinant enzyme was active not only in *Pichia* extracts but also inside the yeast cell, indicating that *Pichia* must have an active or passive system for transporting all the cosubstrates needed by the enzyme into the lumen of the endoplasmic reticulum. No prolyl 4-hydroxylase activity was produced when the α subunit was expressed alone, indicating that the human α subunit does not form an active tetramer with the *Pichia* PDI. This result differs from the situation in insect cells, in which the α subunit, when expressed alone, produces significant amounts of enzyme activity due to tetramer formation with the endogenous insect cell PDI (Vuori *et al.*, 1992b; Veijola *et al.*, 1994). *Pichia* would thus seem to be more suitable than insect cells for detailed studies of the mechanisms involved in prolyl 4-hydroxylase assembly, as tetramer formation with the endogenous PDI polypeptide complicates interpretation of some of the data obtained in insect cells. The *Pichia* PDI thus appears to differ distinctly from the insect cell and *Caenorhabditis elegans* (Veijola *et al.*, 1996a) PDI polypeptides, which do form an active prolyl 4-hydroxylase tetramer with the human α subunit.

Fig. 8. Analysis of the thermal stability of the recombinant human type III collagen expressed in the α /PDI- α MF/pro α 1(III) *Pichia* strain. Pepsin-digested samples were treated with a mixture of trypsin and chymotrypsin at the temperatures indicated at the bottom. Samples were electrophoresed on 8% SDS-PAGE under reducing conditions and analysed by Coomassie staining. The arrow shows the position of the α 1(III) chain.

The signal sequence of PDI was found to play a major role in prolyl 4-hydroxylase assembly in *Pichia*. The authentic human signal sequence appeared to be particularly ineffective for the transport of PDI into the lumen of the endoplasmic reticulum, as only trace amounts of tetramer were produced even when the α subunit DNA was used together with two copies of the PDI DNA. A much higher tetramer assembly level was obtained by using the *P.pastoris* acid phosphatase 1 signal sequence, but even this level was only ~40% of that obtained with the *S.cerevisiae* α mating factor pre-pro sequence. The latter gave the highest amount of tetramer among the various constructs studied here even though this sequence also markedly increased the secretion of PDI into the culture medium. The tetramer produced with this construct contained both molecules in which the propeptide was retained and others in which the propeptide had been

cleaved at the correct site. The presence of the latter molecules suggests that either the PDI polypeptide which subsequently associated with the α subunit or the tetramer had been retrieved from the secretory pathway after the point at which the propeptide had been cleaved. The presence in the tetramer of a PDI polypeptide containing the propeptide agrees with recent data indicating that the N terminus of PDI is not involved in prolyl 4-hydroxylase assembly, as a PDI polypeptide containing a poly-histidine tag in its N terminus forms an enzyme tetramer as efficiently as the wild-type polypeptide (Annunen *et al.*, 1997).

A highly surprising finding to emerge here was that coexpression of collagen polypeptide chains markedly increased the amount of active prolyl 4-hydroxylase tetramer. Northern analyses indicated that this increase was not due to increased transcription of the respective genes or to increased stability of the mRNAs, as the levels of the α subunit and PDI mRNAs were not increased. The amount of PDI polypeptide was likewise not increased. The amount of α subunit in the soluble fraction, i.e. in the soluble enzyme tetramer, was markedly increased, whereas the amount in the insoluble fraction was decreased. The latter decrease cannot be quantified exactly, as more than 95% of the total α subunit was present in the insoluble fraction in the absence of collagen synthesis and more than 50% was insoluble even in its presence. Nevertheless, the data appear to exclude any increase in the total amount of α subunit by more than ~20–30%, a value that is insignificant compared with the ~1000% increase in the amount of enzyme tetramer. It thus seems that the main reason, and probably the only reason, for the ~10-fold increase in the amount of active enzyme tetramer was an ~10-fold increase in the level of association of enzyme subunits.

It has been demonstrated previously that an α subunit present without the PDI polypeptide forms insoluble, inactive aggregates (Kivirikko *et al.*, 1989, 1992; Vuori *et al.*, 1992b). The α subunit may form soluble complexes with BiP (John and Bulleid, 1996; Veijola *et al.*, 1996b), but they have no prolyl 4-hydroxylase activity (Veijola *et al.*, 1996b). It thus seems unlikely that an unassembled α subunit would recognize collagen expression or that collagen expression would influence subunit assembly. A much more likely mechanism is that the α subunit and PDI polypeptide may effectively assemble to form the tetramer, but in the absence of collagen synthesis this tetramer, which has no covalent interchain bonds (Kivirikko *et al.*, 1992), will rapidly dissociate back to its subunits and the α subunit will form insoluble aggregates. The enzyme tetramer is quite stable in the presence of collagen synthesis, and has a half-life of ~2–3 days (Chichester *et al.*, 1976; Majamaa *et al.*, 1979; Berg *et al.*, 1980). Interestingly, an early study suggested that when L-929 fibroblasts began to divide and ceased synthesizing collagen their prolyl 4-hydroxylase activity was lost with an apparent half-life of ~1–2 h (Hebda *et al.*, 1983), but no further data had been available to suggest what mechanism was involved. It is also well known that the enzyme tetramer easily becomes dissociated into its subunits both in cell extracts and in a purified form.

Prolyl 4-hydroxylase catalyses uncoupled decarboxylation of 2-oxoglutarate in the absence of its polypep-

tide substrate (Kivirikko *et al.*, 1989, 1992), and uncoupled decarboxylation, unlike the complete reaction, consumes ascorbate stoichiometrically with 2-oxoglutarate decarboxylation (de Jong and Kemp, 1984; Myllylä *et al.*, 1984). Dissociation of the tetramer in the absence of collagen synthesis would thus prevent the enzyme from performing a reaction that consumes 2-oxoglutarate, O₂ and ascorbate within the lumen of the endoplasmic reticulum. This mechanism would seem to be particularly important in situations in which collagen synthesis is rapidly downregulated, as an active tetramer would otherwise continue to exist for a long period of time. Our re-examination of insect cell expression data and preliminary additional experiments (J. Myllyharju, A. Lamberg and K. I. Kivirikko, unpublished) indicate that the same phenomenon is also seen in insect cells in the absence of recombinant collagen synthesis, although the endogenous collagen synthesis that takes place in these cells partly prevents dissociation, especially when the enzyme is expressed in small amounts. It thus seems likely that this property of prolyl 4-hydroxylase may be found in all cell types, although it has not been recognized previously.

The pro α 1 chains of type III procollagen produced either in the presence or absence of prolyl 4-hydroxylase were of full length, with no evidence of degradation and formed disulfide-bonded trimers. Both types of pro α 1(III) chain formed molecules with collagen domains that were resistant to pepsinization at 22°C, but a considerable fraction of the 4-hydroxyproline-free chains was digested with pepsin, and it seems possible that the pepsin-resistant 4-hydroxyproline-free molecules may have been formed only upon cooling of the cell extracts. The α 1(III) chains produced by pepsin digestion were disulfide-bonded only provided that they had been synthesized in the presence of prolyl 4-hydroxylase. This agrees with recent data indicating that the two cysteine residues present at the very end of the collagen domain in the α 1(III) chains form interchain disulfide bonds only after proline 4-hydroxylation and triple helix formation (Bulleid *et al.*, 1996). This result also indicates that the 4-hydroxyproline-containing pro α 1(III) chains formed triple-helical procollagen molecules inside the yeast cells even though *Pichia*, like *Saccharomyces* (Goffeau *et al.*, 1996), probably has no Hsp47-like protein. Our data thus indicate that Hsp47 (Nagata *et al.*, 1996; Satoh *et al.*, 1996) is probably not needed for the assembly of triple-helical procollagen molecules.

The degree of 4-hydroxylation of proline residues and T_m of the recombinant type III collagen produced in *Pichia* were almost the same as those of the corresponding non-recombinant protein. The expression level obtained under shaker-flask conditions and with only single copies of each DNA was ~15 mg type III procollagen/litre. Many previous studies have demonstrated that shaker-flask conditions are suboptimal for protein production in *Pichia* due to the lack of sufficient O₂, and about a 5- to 10-fold increase in expression levels is usually obtained in bioreactors (Romanos *et al.*, 1992; Cregg *et al.*, 1993). As the K_m of O₂ in the prolyl 4-hydroxylase reaction is as high as ~40 μ M (Tuderman *et al.*, 1977), the O₂ concentration within the lumen of the endoplasmic reticulum is also likely to be rate-limiting for hydroxylation under shaker-flask conditions. Thus the minor differences

in 4-hydroxyproline content and T_m between the recombinant and non-recombinant collagens are likely to disappear when the protein is produced in a bioreactor. It has further been demonstrated previously that the levels of expression of various proteins in *Pichia* increase markedly with the number of DNA copies at least up to 30–50 copies (Buckholz and Gleeson, 1991; Romanos *et al.*, 1992; Scorer *et al.*, 1994). It would thus seem to be easy to optimize the present system under bioreactor conditions and by using high-copy integrants for the production of very large amounts of various collagens.

In conclusion, the present data indicate that it is possible to produce a fully active recombinant human prolyl 4-hydroxylase tetramer and collagen molecules with stable triple helices in *P.pastoris*. *Pichia* would thus seem to provide an ideal system for detailed studies of the mechanisms involved in the assembly of the enzyme tetramer and the collagen triple helix, and of the possible role of Hsp47 in the latter. In addition, *Pichia* would seem to be an ideal system for the high-level production of various recombinant collagen types for numerous scientific and medical purposes. The data further indicate that a highly unusual control system exists in collagen synthesis in *Pichia*, and probably also in other cell types, in that production of a stable prolyl 4-hydroxylase tetramer requires expression of collagen polypeptide chains while the production of collagen molecules with stable triple helices requires expression of active prolyl 4-hydroxylase.

Materials and methods

Construction of expression vectors

A cDNA for the human PDI polypeptide (Pihlajaniemi *et al.*, 1987) extending from the translation initiation codon to the stop codon and flanked by *EcoRI* restriction sites was synthesized using PCR and ligated into the *EcoRI* site of pAO815 (Invitrogen), generating pAO815PDI.

A modified *Pichia* expression vector pARG815 was constructed by replacing the *HIS4* selection marker with an *S.cerevisiae* *ARG4* selection marker. pYM25 (a gift from Dr James Cregg) was digested with *HpaI*, and the *ARG4* *HpaI* fragment was ligated into *EcoRV*-digested pAO815.

The 5' end of human prolyl 4-hydroxylase α subunit cDNA (Helaakoski *et al.*, 1989) extending from the translation initiation codon to an internal *HindIII* site, with *HindIII* and *SmaI* sites preceding the initiation codon, and the 3' end of the same cDNA extending from an internal *PstI* site to the translation stop codon, with *SmaI* and *BamHI* sites following the stop codon, were synthesized by PCR. These PCR fragments were used to generate a modified cDNA corresponding to the human α subunit cDNA clone PA59 (Helaakoski *et al.*, 1989) but without the 5' and 3' untranslated regions. The *SmaI*–*SmaI* α subunit cDNA was ligated into the *EcoRI* site of pARG815, generating pARG815 α .

A prolyl 4-hydroxylase double expression vector was constructed by cloning the *BglII*–*BamHI* PDI expression cassette obtained from pAO815PDI into *BamHI*-digested pARG815 α , generating pARG-815 α PDI.

The sequence coding for the signal peptide of human PDI was replaced with that coding for either the *P.pastoris* acid phosphatase 1 signal sequence (PHO1) or the *S.cerevisiae* α mating factor (α MF) pre-pro sequence. A cDNA for human PDI extending from the codon for the first amino acid after the signal peptide cleavage site to the stop codon and flanked by *EcoRI* restriction sites was synthesized using PCR and ligated into the *EcoRI* site following the PHO1 signal sequence of pHIL-S1 (Invitrogen) and the α MF signal sequence of pPIC9 (Invitrogen), generating pHIL-S1PDI and pPIC9PDI, respectively. In the pPIC9PDI construct the cDNA sequence coding for the C-terminal endoplasmic reticulum retention signal -KDEL of PDI was modified to an -HDEL signal by PCR, generating pPIC9PDI-HDEL. The sequence coding for the signal peptide of the human prolyl 4-hydroxylase α subunit was replaced with that coding for the PHO1 signal sequence by PCR, generating pARG815 α -PHO1.

The 3' end of the cDNA for the pro α 1 chain of human type III procollagen, extending from an internal *EcoRI* site to the translation stop codon, with a *XbaI* site following the stop codon, was synthesized by PCR and used to replace the 3' end containing 713 bp of 3' untranslated sequence of a full-length pro α 1(III) cDNA (Tromp *et al.*, 1989) with an artificial *BglIII* site created 6 bp upstream from the translation initiation codon (Lamberg *et al.*, 1996). The *BglIII*–*XbaI* pro α 1(III) cDNA was ligated into the *EcoRI* site of pHIL-D2 (Invitrogen) and the *EcoRI*–*XbaI* site of pPICZ B (Invitrogen), generating pHIL-D2pro α 1(III) and pPICZ Bpro α 1(III), respectively.

Transformation, growth and induction of *P.pastoris*

The *his4*, *arg4* *P.pastoris* host strain was a gift from Dr James Cregg. pARG815 α and pAO815PDI were separately transformed into the host strain by the electroporation method (Manual Version 3.0 of the *Pichia* Expression Kit, Invitrogen) to generate the α and PDI strains, expressing the prolyl 4-hydroxylase α subunit and PDI, respectively. Six *Pichia* strains expressing prolyl 4-hydroxylase: α /PDI, α PDI/PDI, α /PDI-PHO1, α /PDI- α MF, α /PDI- α MF-HDEL and α -PHO1/PDI- α MF (Table I), were generated by cotransforming the host strain with pARG815 α and pAO815PDI, pARG815 α PDI and pAO815PDI, pARG815 α and pHIL-S1PDI, pARG815 α and pPIC9PDI, pARG815 α and pPIC9PDI-HDEL, and pARG815 α -PHO1 and pPIC9PDI, respectively. In all the transformations the α subunit expression vectors and the pARG815 α PDI vector were linearized with *DraIII*, while the PDI expression vectors were linearized with *StuI*.

A *Pichia* strain expressing the pro α 1 chains of human type III procollagen, Pro α 1(III), was generated by transforming the host strain with *StuI*-linearized pHIL-D2pro α 1(III), and a strain α /PDI- α MF/pro α 1(III) coexpressing human prolyl 4-hydroxylase and pro α 1(III) chains by transforming *PmeI*-linearized pPICZ Bpro α 1(III) into the prolyl 4-hydroxylase-expressing strain α /PDI- α MF (Table I).

Arg⁺, Mut⁺ transformants of the α strain, His⁺, Mut⁺ transformants of the PDI strain, Arg⁺, His⁺, Mut⁺ transformants of the six prolyl 4-hydroxylase-expressing strains, and His⁺, Mut⁺ transformants of the pro α 1(III) strain were selected. Transformants of the α /PDI- α MF/pro α 1(III) strain were selected in YPD (+ Zeocin) plates (Invitrogen), followed by Arg⁺, His⁺ selection. The cells were cultured according to the methods described in the Manual Version 3.0 of the *Pichia* Expression Kit (Invitrogen) with slight modifications. Shaker-flask cultures were grown in a buffered glycerol complex medium (BMGY, pH 6.0) with 1 g/l yeast extract and 2 g/l peptone. Expression was induced in a buffered minimal methanol medium (BMM, pH 6.0), and methanol was added every 24 h to a final concentration of 0.5%. Amino acids were added up to 50 μ g/l as required. L-ascorbic acid (sodium salt), 80 μ g/ml, was added every 12 h during the induction of collagen expression.

Analysis of recombinant proteins produced in *P.pastoris*

Cells were harvested after a 60 h methanol induction at 30°C, washed once and suspended in a cold (4°C) 0.1 M NaCl, 0.1 M glycine, 0.1% Triton X-100, 10 μ M dithiothreitol (DTT), 1 mM phenylmethylsulfonyl-fluoride (PMSF) and 10 mM Tris buffer, pH 7.8, or in a 5% glycerol, 1 mM PMSF, 1 mM EDTA and 50 mM sodium phosphate buffer, pH 7.4. Cells harvested from small-scale cultures were broken by vortexing with glass beads (0.5 mm diameter) and those from large-scale cultures by agitation with glass beads using a Minibead-Beater (Biospec Products). The lysate was centrifuged at 10 000 g for 30 min, and the insoluble fraction was further solubilized in 1% SDS. Protein concentrations were determined using the Bio-Rad Protein Assay (Bio-Rad). The cell culture medium was concentrated in Centricon-30 (Amicon). Aliquots of the soluble and insoluble fractions of the cell lysates and the cell culture medium were analysed by SDS-PAGE and non-denaturing PAGE followed by staining with Coomassie Blue or by Western blotting with a polyclonal antibody K38 (Veijola *et al.*, 1996b) to the human PDI polypeptide, a monoclonal antibody M95L or a polyclonal antibody K17 (Veijola *et al.*, 1996b) to the human prolyl 4-hydroxylase α subunit, a polyclonal antibody to the N-propeptide of human type III procollagen (Farnos Diagnostica) or a monoclonal antibody 95D1A recognizing the collagenous regions of various collagen chains (A. Snellman and T. Pihlajaniemi, unpublished observations). Prolyl 4-hydroxylase activity was assayed by a method based on the hydroxylation-coupled decarboxylation of 2-oxo[1-¹⁴C]glutarate or on the formation of 4-hydroxy[¹⁴C]proline using a [¹⁴C]proline-labelled procollagen substrate (Kivirikko and Myllylä, 1982). K_m values were determined as described previously (Myllylä *et al.*, 1977). Further aliquots of the soluble fraction and the cell culture medium were analysed by a radioimmunoassay for the trimeric N-propeptide of human type III

procollagen (Farnos Diagnostica), and others were digested with pepsin for 2 h at 22°C. The thermal stability of the pepsin-resistant recombinant type III collagen was studied by means of a brief digestion with a mixture of trypsin and chymotrypsin at various temperatures (Bruckner and Prockop, 1981).

Dot blot and Northern blot analyses

Chromosomal DNA, 1–5 µg, was analysed by a DNA dot blot method (Ausubel *et al.*, 1995). The DNA was isolated using the Easy-DNA Kit (Invitrogen), with *P.pastoris* PER3 (obtained from Dr James Cregg) as a single copy control.

Total RNA was isolated from transformed cells 12 h after induction with methanol (0.5%) (Schmitt, 1990). RNA was analysed by Northern blotting using a ³²P-labelled 1099 bp fragment of cDNA for the human prolyl 4-hydroxylase α subunit, a 828 bp fragment of cDNA for human PDI and a 862-bp fragment of cDNA for human pro α 1(III) chains as probes. A *P.pastoris* AOX1 fragment was used as an internal control.

Purification of recombinant proteins

Prolyl 4-hydroxylase was purified by breaking the cells in a 0.1 M NaCl, 0.1 M glycine, 0.1% Triton X-100, 10 µM DTT, 1 mM PMSF and 10 mM Tris buffer, pH 7.8. The cell lysate was centrifuged at 10 000 g for 30 min, and the supernatant was chromatographed on a poly(L-proline) affinity column (Kivirikko and Myllylä, 1987) followed by a Superdex 200 gel filtration column (Pharmacia). The PDI polypeptide secreted into the minimal medium was partly purified in a Superdex 200 column (Pharmacia).

Type III collagen was purified by breaking the cells in a 5% glycerol, 1 mM EDTA and 50 mM sodium phosphate buffer, pH 7.4. The cell lysate was centrifuged at 10 000 g for 30 min, the supernatant was digested with a final concentration of 150 µg/ml of pepsin for 2 h at 22°C, and the pepsin was inactivated by neutralization. The sample was centrifuged at 100 000 g for 40 min, concentrated in Centricon-50, chromatographed on a Superdex 200 column (Pharmacia) with a solution of 200 mM NaCl and 25 mM sodium phosphate, pH 7.4, dialysed against 0.1 M acetic acid, and lyophilized.

Other assays

N-glycosidase F treatment was performed according to the instructions provided by the manufacturer (Boehringer Mannheim). The N-terminal sequences of the subunits of the purified recombinant human prolyl 4-hydroxylases were determined on an Applied Biosystems model 477A on-line 120 liquid pulse protein sequencer (Department of Biochemistry and Biotechnology, University of Kuopio). Amino acid analysis of the purified type III collagen was performed in an Applied Biosystems 421 amino acid analyzer.

Acknowledgements

We thank Dr James Cregg, Oregon Graduate Institute of Science and Technology, for the highly valuable advice and suggestions, the *his4*, *arg4* *P.pastoris* host strain, the plasmid pYM25 and the *P.pastoris* PER3 DNA; Kari Juntunen, Department of Clinical Chemistry, University of Oulu, for help with protein purification; and Minna Siurua, Anu Myllymäki, Eeva Lehtimäki and Anne Kokko for their expert technical assistance. This work was supported by grants from the Health Sciences Council of the Academy of Finland, from the European Commission (BIO4-CT96-0537) and from FibroGen Inc. (South San Francisco, CA).

References

- Annunen, P., Helaakoski, T., Myllyharju, J., Veijola, J., Pihlajaniemi, T. and Kivirikko, K.I. (1997) Cloning of the human prolyl 4-hydroxylase α subunit isoform α (II) and characterization of the type II enzyme tetramer. The α (I) and α (II) subunits do not form a mixed α (I) α (II) β ₂ tetramer. *J. Biol. Chem.*, **272**, 17342–17348.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) *Current Protocols in Molecular Biology*. John Wiley and Sons, New York.
- Berg, R.A., Kao, W.W.-Y. and Kedesha, N.L. (1980) The assembly of tetrameric prolyl hydroxylase in tendon fibroblasts from newly synthesized α -subunits and from preformed cross-reacting protein. *Biochem. J.*, **189**, 491–499.
- Bruckner, P. and Prockop, D.J. (1981) Proteolytic enzymes as probes for the triple-helical conformation of procollagen. *Anal. Biochem.*, **110**, 360–368.

- Buckholz, R.G. and Gleeson, M.A. (1991) Yeast systems for the commercial production of heterologous proteins. *Biotechnology*, **9**, 1067–1072.
- Bulleid, N.J., Wilson, R. and Lees, J.F. (1996) Type-III procollagen assembly in semi-intact cells: chain association, nucleation and triple-helix folding do not require formation of inter-chain disulphide bonds but triple-helix nucleation does require hydroxylation. *Biochem. J.*, **317**, 195–202.
- Chichester III, C.O., Fuller, G.C. and Cardinale, G.J. (1976) *In vivo* labeling and turnover of prolyl hydroxylase and a related immunoreactive protein. *Biochem. Biophys. Res. Commun.*, **73**, 1056–1062.
- Chung, E. and Miller, E.J. (1974) Collagen polymorphism: characterization of molecules with the chain composition α 1(III)₃. *Science*, **183**, 1200–1201.
- Cregg, J.M., Vedvick, T.S. and Raschke, W.C. (1993) Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology*, **11**, 905–910.
- de Jong, L. and Kemp, A. (1984) Stoichiometry and kinetics of the prolyl 4-hydroxylase partial reaction. *Biochim. Biophys. Acta*, **787**, 105–111.
- Freedman, R.B., Hirst, T.R. and Tuite, M.F. (1994) Protein disulphide-isomerase: building bridges in protein folding. *Trends Biochem. Sci.*, **19**, 331–336.
- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S.G. (1996) Life with 6000 genes. *Science*, **274**, 546–567.
- Hebda, P.A., Ebert, J., Chou, K.-L.L., Shields, M. and Kao, W.W.-Y. (1983) The association between prolyl hydroxylase metabolism and cell growth in cultured L-929 fibroblasts. *Biochim. Biophys. Acta*, **758**, 128–134.
- Helaakoski, T., Vuori, K., Myllylä, R., Kivirikko, K.I. and Pihlajaniemi, T. (1989) Molecular cloning of the α -subunit of human prolyl 4-hydroxylase: The complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. *Proc. Natl Acad. Sci. USA*, **86**, 4392–4396.
- Helaakoski, T., Annunen, P., Vuori, K., MacNeil, I.A., Pihlajaniemi, T. and Kivirikko, K.I. (1995) Cloning, baculovirus expression, and characterization of a second mouse prolyl 4-hydroxylase α -subunit isoform: Formation of an α ₂ β ₂ tetramer with the protein disulfide-isomerase/ β subunit. *Proc. Natl Acad. Sci. USA*, **92**, 4427–4431.
- John, D.C.A. and Bulleid, N.J. (1996) Intracellular association and reassembly of prolyl 4-hydroxylase: the α -subunits associate with the immunoglobulin-heavy-chain binding protein (BiP) allowing reassembly with the β -subunit. *Biochem. J.*, **317**, 659–665.
- John, D.C.A., Grant, M.E. and Bulleid, N.J. (1993) Cell-free synthesis and assembly of prolyl 4-hydroxylase: the role of the β -subunit (PDI) in preventing misfolding and aggregation of the α -subunit. *EMBO J.*, **12**, 1587–1595.
- Kiely, C.M., Hopkinson, I. and Grant, M.E. (1993) The collagen family: Structure, assembly and organization in the extracellular matrix. In Royce, P.M. and Steinman, B. (eds), *Connective Tissue and its Heritable Disorders. Molecular, Genetic and Medical Aspects*. Wiley-Liss, New York, pp. 103–147.
- Kivirikko, K.I. and Myllylä, R. (1982) Posttranslational enzymes in the biosynthesis of collagen: Intracellular enzymes. *Methods Enzymol.*, **82**, 245–304.
- Kivirikko, K.I. and Myllylä, R. (1987) Recent developments in posttranslational modification: Intracellular processing. *Methods Enzymol.*, **144**, 96–114.
- Kivirikko, K.I., Myllylä, R. and Pihlajaniemi, T. (1989) Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit. *FASEB J.*, **3**, 1609–1617.
- Kivirikko, K.I., Myllylä, R. and Pihlajaniemi, T. (1992) Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In Harding, J.J. and Crabbe, M.J.C. (eds), *Post-translational Modifications of Proteins*. CRC Press, Boca Raton, FL, pp. 1–51.
- Koivu, J., Myllylä, R., Helaakoski, T., Pihlajaniemi, T., Tasanen, K. and Kivirikko, K.I. (1987) A single polypeptide acts both as the β subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase. *J. Biol. Chem.*, **262**, 6447–6449.
- Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T. and Kivirikko, K.I. (1996) Characterization of human type III collagen expressed in a baculovirus system. Production of a protein with a stable triple helix requires coexpression with the two types of recombinant prolyl 4-hydroxylase subunit. *J. Biol. Chem.*, **271**, 11988–11995.

- Majamaa,K., Kuutti-Savolainen,E.R., Tuderman,L. and Kivirikko,K.I. (1979) Turnover of prolyl hydroxylase tetramers and the monomer-size protein in chick-embryo cartilaginous bone and lung *in vivo*. *Biochem. J.*, **178**, 313–322.
- Mayne,R. and Brewton,R.G. (1993) New members of collagen superfamily. *Curr. Opin. Cell Biol.*, **5**, 883–890.
- Myllylä,R., Tuderman,L. and Kivirikko,K.I. (1977) Mechanism of the prolyl hydroxylase reaction. 2. Kinetic analysis of the reaction sequence. *Eur. J. Biochem.*, **80**, 349–357.
- Myllylä,R., Majamaa,K., Günzler,V., Hanauske-Abel,H.M. and Kivirikko,K.I. (1984) Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *J. Biol. Chem.*, **259**, 5403–5405.
- Nagata,K. (1996) Hsp47: a collagen-specific molecular chaperone. *Trends Biochem. Sci.*, **21**, 23–26.
- Pelham,H.R.B. (1990) The retention signal for soluble proteins of the endoplasmic reticulum. *Trends Biochem. Sci.*, **15**, 483–486.
- Pihlajaniemi,T. and Rehn,M. (1995) Two new collagen subgroups: Membrane-associated collagens and types XV and XVIII. *Progr. Nucl. Acid Res. Mol. Biol.*, **50**, 225–262.
- Pihlajaniemi,T., Helaakoski,T., Tasanen,K., Myllylä,R., Huhtala,M.-L., Koivu,J. and Kivirikko,K.I. (1987) Molecular cloning of the β -subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. *EMBO J.*, **6**, 643–649.
- Prockop,D.J. and Kivirikko,K.I. (1995) Collagens: Molecular biology, diseases, and potentials for therapy. *Annu. Rev. Biochem.*, **64**, 403–434.
- Romanos,M.A., Scorer,C.A. and Clare,J.J. (1992) Foreign gene expression in yeast: a review. *Yeast*, **8**, 423–488.
- Satoh,M., Hirayoshi,K., Yokota,S., Hosokawa,N. and Nagata,K. (1996) Intracellular interaction of collagen-specific stress protein Hsp47 with newly synthesized procollagen. *J. Cell Biol.*, **133**, 469–483.
- Schmitt,M.E. (1990) A rapid and simple method for preparation of RNA from *S.cerevisiae*. *Nucleic Acids Res.*, **18**, 3091–3092.
- Scorer,C.A., Clare,J.J., McCombie,W.R., Romanos,M.A. and Sreekrishna,K. (1994) Rapid selection using G418 of high copy number transformants of *Pichia pastoris* for high-level foreign gene expression. *Biotechnology*, **12**, 181–184.
- Tromp,G., Kuivaniemi,H., Shikata,H. and Prockop,D.J. (1989) A single base mutation that substitutes serine for glycine 790 of the $\alpha 1(\text{III})$ chain of type III procollagen exposes an arginine and causes Ehlers–Danlos syndrome IV. *J. Biol. Chem.*, **264**, 1349–1352.
- Tuderman,L., Myllylä,R. and Kivirikko,K.I. (1977) Mechanism of the prolyl hydroxylase reaction. 1. Role of co-substrates. *Eur. J. Biochem.*, **80**, 341–348.
- Veijola,J., Koivunen,P., Annunen,P., Pihlajaniemi,T. and Kivirikko,K.I. (1994) Cloning, baculovirus expression, and characterization of the α subunit of prolyl 4-hydroxylase from the nematode *Caenorhabditis elegans*. This α subunit forms an active $\alpha\beta$ dimer with the human protein disulfide isomerase/ β subunit. *J. Biol. Chem.*, **269**, 26746–26753.
- Veijola,J., Annunen,P., Koivunen,P., Page,A.P., Pihlajaniemi,T. and Kivirikko,K.I. (1996a) Baculovirus expression of two protein disulphide isomerase isoforms from *Caenorhabditis elegans* and characterization of prolyl 4-hydroxylases containing one of these polypeptides as their β subunit. *Biochem. J.*, **317**, 721–729.
- Veijola,J., Pihlajaniemi,T. and Kivirikko,K.I. (1996b) Co-expression of the α subunit of human prolyl 4-hydroxylase with BiP polypeptide in insect cells leads to the formation of soluble and insoluble complexes. *Biochem. J.*, **315**, 613–618.
- Vuori,K., Myllylä,R., Pihlajaniemi,T. and Kivirikko,K.I. (1992a) Expression and site-directed mutagenesis of human protein disulfide isomerase in *Escherichia coli*. This multifunctional polypeptide has two independently acting catalytic sites for the isomerase activity. *J. Biol. Chem.*, **267**, 7211–7214.
- Vuori,K., Pihlajaniemi,T., Marttila,M. and Kivirikko,K.I. (1992b) Characterization of the human prolyl 4-hydroxylase tetramer and its multifunctional protein disulfide-isomerase subunit synthesized in a baculovirus expression system. *Proc. Natl Acad. Sci. USA*, **89**, 7467–7470.
- Vuori,K., Pihlajaniemi,T., Myllylä,R. and Kivirikko,K.I. (1992c) Site-directed mutagenesis of human protein disulfide isomerase: effect on the assembly, activity and endoplasmic reticulum retention of human prolyl 4-hydroxylase in *Spodoptera frugiperda* insect cells. *EMBO J.*, **11**, 4213–4217.

Received on August 8, 1997; revised on September 8, 1997