Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Protein Expression and Purification 38 (2004) 279–291

Protein Expression & Purification

www.elsevier.com/locate/yprep

Production of human prolyl 4-hydroxylase in *Escherichia coli*

Elizabeth A. Kersteen[a], Joshua J. Higgin[b], Ronald T. Raines[a,b,*]

[a] *Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA*
[b] *Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA*

Received 6 August 2004, and in revised form 7 September 2004

Abstract

Prolyl 4-hydroxylase (P4H) catalyzes the post-translational hydroxylation of proline residues in collagen strands. The enzyme is an $\alpha_2\beta_2$ tetramer in which the α subunits contain the catalytic active sites and the β subunits (protein disulfide isomerase) maintain the α subunits in a soluble and active conformation. Heterologous production of the native $\alpha_2\beta_2$ tetramer is challenging and had not been reported previously in a prokaryotic system. Here, we describe the production of active human P4H tetramer in *Escherichia coli* from a single bicistronic vector. P4H production requires the relatively oxidizing cytosol of Origami B(DE3) cells. Induction of the wild-type α(I) cDNA in these cells leads to the production of a truncated α subunit (residues 235–534), which assembles with the β subunit. This truncated P4H is an active enzyme, but has a high K_m value for long substrates. Replacing the Met235 codon with one for leucine removes an alternative start codon and enables production of full-length α subunit and assembly of the native $\alpha_2\beta_2$ tetramer in *E. coli* cells to yield 2 mg of purified P4H per liter of culture (0.2 mg/g of cell paste). We also report a direct, automated assay of proline hydroxylation using high-performance liquid chromatography. We anticipate that these advances will facilitate structure–function analyses of P4H.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Collagen; Poly(proline)-affinity chromatography; Prolyl 4-hydroxylase; Protein disulfide isomerase; Translation initiation; Tetramer

Collagen, the most abundant protein in animals, is a right-handed triple helix composed of three polyproline type-II strands. Each strand contains a repeating sequence of Xaa-Yaa-Gly, where the Xaa residue is often (2*S*)-proline (Pro[1]) and the Yaa residue is often (2*S*,4*R*)-4-hydroxyproline (Hyp). The Hyp residues in collagen strands arise from the post-translational modification of Pro residues by prolyl 4-hydroxylase (P4H). This hydroxylation reaction is critical for the folding of the collagen triple helix [1–3]. Moreover, the conformational stability of a collagen triple helix correlates with its Hyp content [4,5]. P4H is known to be an essential enzyme for the nematode *Caenorhabditis elegans* [6,7].

P4H resides in the lumen of the endoplasmic reticulum (ER), where the folding and assembly of procollagen occurs. In vertebrates, P4H is found as an $\alpha_2\beta_2$ tetramer; in *C. elegans*, P4H can also exist as an αβ dimer [8]. Each α subunit contains a catalytic active site for proline hydroxylation [9]. The β subunit is protein disulfide isomerase (PDI) [10], a thiol–disulfide oxidoreductase with numerous roles in the cell [11–13]. The primary roles of PDI in the P4H tetramer appear to be to keep the α subunit in a soluble and active form [14,15] and to retain the enzyme in the ER through the KDEL sequence at its C terminus [16]. The active-site cysteine

* Corresponding author. Fax: +1 608 2623453.
E-mail address: raines@biochem.wisc.edu (R.T. Raines).

[1] *Abbreviations used*: ER, endoplasmic reticulum; Hyp, (2*S*,4*R*)-4-hydroxyproline; FPLC, fast performance liquid chromatography; HPLC, high-performance liquid chromatography; IPTG, isopropyl-1-thio-β-D-galactopyranoside; LB, Luria–Bertani; LC/MS, liquid chromatography/mass spectrometry; PAGE, poly(acrylamide) gel electrophoresis; Pro, (2*S*)-proline; P4H, prolyl 4-hydroxylase; P4Hα(I), isoform I of the α subunit of human prolyl 4-hydroxylase; PDI, protein disulfide isomerase; rbs, ribosome-binding site; SDS, sodium dodecyl sulfate; TB, Terrific Broth; Tris, tris(hydroxymethyl)aminomethane.

1046-5928/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2004.09.008

residues of PDI are not required for P4H activity or tetramer assembly [15]. Three human isoforms of the α subunit have been identified: α(I), α(II), and α(III) [17–19]. P4Hα(I) accounts for most of the prolyl 4-hydroxylase activity in a variety of cell types [17,18,20].

P4H is a member of a class of α-ketoglutarate-dependent non-heme Fe(II) dioxygenase enzymes [11,21,22]. These enzymes require α-ketoglutarate, Fe(II), and O_2 for catalysis. During proline hydroxylation (Fig. 1), α-ketoglutarate is decarboxylated oxidatively to form succinate [23]. One atom of O_2 is incorporated into the nascent hydroxyl group of Hyp, and the other into succinate [24]. Ascorbate is required to rescue inactivated enzyme that accumulates from decarboxylation of α-ketoglutarate that is uncoupled from substrate hydroxylation [25–27].

α Subunit residues that are critical for catalysis have been identified by site-directed mutagenesis. The iron-binding ligands are His412, Asp414, and His483 [8,9]. Lys493 binds to the C-5 carboxyl group of α-ketoglutarate, and His501 likely binds to its C-1 carboxyl group [8]. Additional insights into the catalytic mechanism have been gained by comparison of P4H to other α-ketoglutarate-dependent dioxygenases [28,29]. Substrate binding has been investigated with an assay that measures hydroxylation of $(\text{Pro-Pro-Gly})_n$ collagen mimics indirectly—by monitoring the release of $[^{14}C]CO_2$ during decarboxylation of $[1\text{-}^{14}C]$α-ketoglutarate [23,30]. Increasing the length of collagen-like substrates leads to tighter binding in the enzymic active site [31]. The lack of a known three-dimensional structure for P4H, however, has made it difficult to further decipher details of the mechanism of hydroxylation. The difficulties faced in the assembly of recombinant P4H tetramer have significantly hindered crystallographic study of the enzyme.

Production of the P4H tetramer is not trivial. Attempts to assemble an active P4H tetramer from its individual α and β subunits in vitro have been unsuccessful [32,33]. Although the P4H subunits dissociate easily, reassociation of the α subunit (which is insoluble) with PDI to form an active tetramer has not been possible. It has been suggested [34] that in vivo assembly of P4H requires chaperone proteins, present in the ER of eukaryotic cells, to maintain the α subunit in a soluble form until PDI can bind, and allow formation of native tetramer.

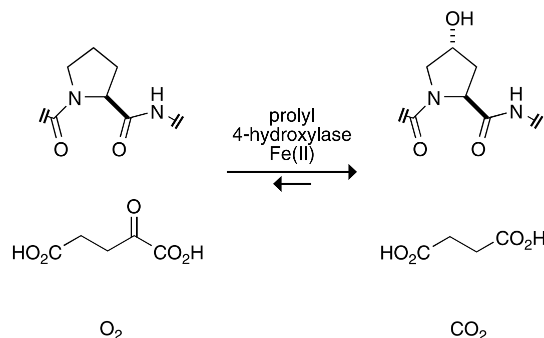

Fig. 1. Reaction catalyzed by prolyl 4-hydroxylase.

Human P4H has been produced heterologously in baculovirus [14]. Baculovirus expression vectors encoding the α and β subunits were used to infect *Spodoptera frugiperda* (Sf9) insect cells. Association of the subunits to form active tetramer (5 mg/L of culture) was observed upon cotransfection. Recently, assembly of active human P4H tetramer was also achieved in the yeast *Pichia pastoris* [35]. Production of the α subunit along with the β subunit, attached to the *Saccharomyces cerevisiae* α mating factor pre-pro signal sequence, yields active P4H tetramer in yeast. Higher levels of P4H tetramer (15 mg/L of culture) were obtained by coexpressing the α and β subunits with recombinant human type III procollagen in yeast [35].

We suspected that developing an *Escherichia coli* expression system for P4H could be of substantial benefit to its further study. Heterologous production of recombinant proteins has been successful in a number of host organisms [36–38]. Still, *E. coli* remains the most powerful [39,40]. For example, almost all proteins that have been crystallized (97%) and have had their structures determined by X-ray diffraction analysis (93%) were produced in *E. coli* [41]. Moreover, the generation of protein variants is comparatively facile in *E. coli* systems [42].

Here, we report the production of active human P4H in *E. coli*. The two P4H subunits were expressed from a bicistronic vector in cells that have an oxidizing cytosol (Origami B(DE3)). In addition to intact P4H, we also inadvertently produced a truncated variant of P4H (P4Hα(I)235–534/β) by the initiation of translation from an internal methionine residue. Finally, we report the development of a facile assay for proline hydroxylation. These advances could facilitate structure–function analyses of P4H.

Materials and methods

Materials

Escherichia coli strains BL21(DE3), Origami B(DE3), and Rosetta-gami 2(DE3), and the pET22b(+) expression vector were from Novagen (Madison, WI). Enzymes for DNA manipulation were from Promega (Madison, WI). DNA oligonucleotides for mutagenesis and sequencing were from Integrated DNA Technologies (Coralville, IA). DNA sequencing reactions were performed with the BigDye kit from Applied Biosystems (Foster City, CA) and CleanSeq magnetic beads were from Agencourt Bioscience (Beverly, MA). DNA sequences were determined by using capillary arrays on

an Applied Bioscience automated sequencing instrument at the University of Wisconsin Biotechnology Center.

Poly(proline) was from Sigma Chemical (St. Louis, MO). Prosieve protein markers for SDS–PAGE and PAGEr Gold Precast gels (7.5% Tris–glycine) for native PAGE were from Cambrex Bio Science Rockland (Rockland, ME). Native PAGE molecular mass standards were from Amersham Biosciences (Piscataway, NJ). Prestained SDS–PAGE molecular mass standards were from Bio-Rad (Hercules, CA). $(Pro\text{-}Pro\text{-}Gly)_{10}$ was from Peptides International (Louisville, KY). $[1\text{-}^{14}C]\alpha$-Ketoglutarate was from American Radiolabeled Chemicals (St. Louis, MO). All other chemicals were of reagent grade or better, and were used without further purification.

Luria–Bertani (LB) medium contained (in 1.0 L) tryptone (10 g), yeast extract (5 g), and NaCl (10 g). Terrific broth (TB) medium contained (in 1.0 L) tryptone (12 g), yeast extract (24 g), K_2HPO_4 (72 mM), KH_2PO_4 (17 mM), and glycerol (4 mL). All media were prepared in deionized, distilled water, and autoclaved.

Human P4H produced in a baculovirus system [14] was a generous gift of Johanna Myllyharju (University of Oulu) and was used as a standard for the experiments reported herein.

Instrumentation

UV absorbance measurements were made with a Cary 50 spectrophotometer from Varian (Palo Alto, CA). High-performance liquid chromatography (HPLC) was performed with a system from Waters (Milford, MA) that includes two 515 pumps, a 717plus autosampler, and a 996 photodiode array detector. The system was controlled with the manufacturer's Millennium32 (Version 3.20) software. Fast performance liquid chromatography (FPLC) was performed with an ÄKTA system from Amersham–Pharmacia (Piscataway, NJ) and analyzed with the UNICORN Control System. Scintillation counting was performed on a Wallac 1450 MicroBeta TriLux liquid scintillation counter from Perkin–Elmer (Wellesley, MA).

Design of P4H expression vector

Our strategy for expressing active P4H tetramer in *E. coli* was to clone cDNAs encoding the α and β subunit into the same plasmid. From this bicistronic vector, transcription of both cDNAs could be accomplished from the same T7 promoter, and each subunit would have its own ribosome binding site (rbs) for the initiation of translation.

cDNA encoding human PDI was a kind gift from Alan D. Attie (University of Wisconsin-Madison). That cDNA, without its signal sequence encoding region, was inserted between the *Nde*I and *Bam*HI restriction sites of a pET22b(+) expression vector to give plasmid pBK1.PDI1.

cDNA encoding P4Hα(I) (PA11 clone), which had been isolated from HeLa cells and inserted into a pBSKS-vector (pBS.LF17-1), was a kind gift from Lisa Friedman (University of Wisconsin-Madison). DNA encoding P4Hα(I) was isolated from pBS.LF17-1 by the PCR using primers that flank regions on the 5′ side of the translation initiation codon and the 3′ side of the stop codon, each including a *Bam*HI restriction site. The PCR fragment was cloned into the PCR4-TOPO vector (Invitrogen, Carlsbad, CA) and then digested with *Bam*HI. The resulting fragment was then ligated into pBK1.PDI1 that had been digested with *Bam*HI, yielding plasmid pBK1.PDI1.P4H1.

DNA encoding the signal sequence of P4Hα(I) was removed in a series of steps. QuikChange (Stratagene, La Jolla, CA) site-directed mutagenesis was used to: (1) add an rbs approximately 15 bp on the 5′ side of the first encoded amino acid of the processed protein, (2) replace the last codon for the signal sequence region with an ATG start codon, and (3) add an *Nde*I site immediately on the 5′ side of the engineered start codon. The resulting plasmid pBK1.PDI1.P4H5 was used to produce the P4Hα(I)235–534/β enzyme, which is a P4H oligomer with an α subunit of only 32 kDa.

A plasmid encoding the P4Hα(I) subunit alone was produced by digesting pBK1.PDI1.P4H5 with *Nde*I, removing the DNA fragment encoding PDI, and ligating the vector. The resulting construct (pBK1.P4H5) was subjected to QuikChange mutagenesis to add a *Bam*HI site on the 5′ side of the pET22b(+) rbs, yielding plasmid pBK1.P4H6. *Bam*HI digestion of pBK1.PDI1 and pBK1.P4H6, followed by ligation, resulted in a vector with the PDI cDNA preceding the P4Hα(I) cDNA, both having the polylinker of pET22b(+) on the 5′ side of their start codons (pBK1.PDI1.P4H6). Finally, the ATG codon of Met235 of the α subunit was replaced with the CTT codon of leucine by QuikChange mutagenesis to yield plasmid pBK1.PDI1.P4H7 (Fig. 2), which was used to produce the full-length P4H tetramer.

Production of P4H

Expression vectors were transformed into BL21(DE3) or Origami B(DE3) cells by electroporation. Transformations in BL21(DE3) were grown on LB agar containing ampicillin (100 μg/mL); transformations in Origami B(DE3) were grown on LB agar containing ampicillin (100 μg/mL), kanamycin (15 μg/mL), and tetracycline (12.5 μg/mL). A starter culture in LB medium (25 mL), supplemented with antibiotics as above, was inoculated with a fresh colony within 24 h of transformation. Cells were grown overnight for 10–12 h, and then harvested by centrifugation at 5000*g* for 10 min. The resulting cell pellet was resuspended in fresh LB

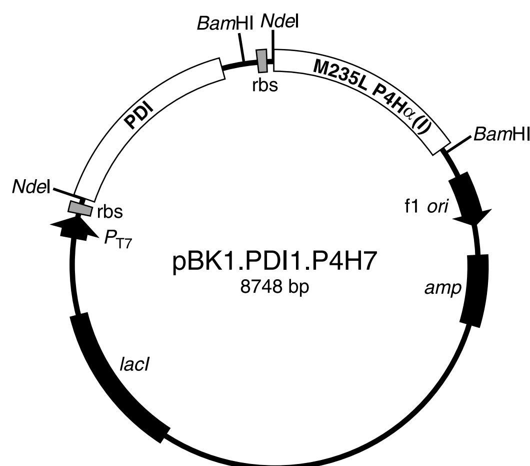

Fig. 2. Map of plasmid pBK1.PDI1.P4H7, which is a bicistronic expression vector in which cDNAs encoding human PDI and the M235L variant of human P4Hα(I) have been inserted into pET22b(+). Both cDNAs are transcribed from a single T7 promoter. Each resulting transcript contains an rbs for the initiation of translation.

medium (25 mL) and used to inoculate to $OD_{600} = 0.01$ either a 1 L culture of TB medium containing ampicillin (200 μg/mL) for the growth of BL21(DE3) cells, or TB medium containing ampicillin (200 μg/mL), kanamycin (15 μg/mL), and tetracycline (12.5 μg/mL) for the growth of Origami B(DE3) cells. The inoculated culture was shaken (210 rpm) at 37 °C until reaching $OD_{600} = 1.7$–1.8. cDNA expression was then induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG, 500 μM) and shaking at 37 °C for 4 h in BL21(DE3) cells or at 23 °C for 12–18 h in Origami B(DE3) cells. Cells were harvested by centrifugation and resuspended in 100 mM glycine, 10 mM Tris, and 100 mM NaCl, pH 7.8. Approximately, 10 g of cell paste was obtained from each liter of culture. Harvested and resuspended cells (10 g of cell paste in 20 mL) were lysed by passage three times through a French pressure cell. Insoluble material was removed by centrifugation at 50,000*g* for 1 h at 4 °C in an Optima XL-80K ultracentrifuge from Beckman Coulter (Fullerton, CA). Both subunits of P4H remain in the soluble lysate fraction.

Purification of P4H

Purification of P4H essentially follows procedures published previously [43]. The lysate was first purified by ammonium sulfate fractionation. Saturated ammonium sulfate was added to the lysate to 30% saturation, and the pellet was removed by centrifugation at 15,000*g* for 30 min. The supernatant was precipitated by adding ammonium sulfate to 65% saturation. The pellet was collected by centrifugation at 15,000*g* for 30 min and resuspended in 100 mM glycine, 10 mM Tris, and 100 mM NaCl, pH 7.8. The 30–65% ammonium sulfate fraction was dialyzed overnight against 4 L of the same buffer.

Poly(proline) resin was prepared by coupling poly(proline) (average molecular mass: 32 kDa) to cyanogen bromide-activated Sepharose four resin (Sigma Chemical), as described previously [44,45]. The ammonium sulfate fraction was diluted to 10 mg/mL (total volume: 180 mL) and loaded onto a column (7 × 7 cm^2) of poly(proline) resin at a flow rate of 1 mL/min. The column was washed with 100 mM glycine, 10 mM Tris, and 100 mM NaCl, pH 7.8 until the OD_{230} and OD_{280} of the eluate were <0.05 (approximately 250 mL). P4H was eluted with 20 mL of 5 mg/mL poly(proline) (average molecular mass: 7 kDa) in the same buffer, followed by buffer alone. The absorbance at 230 and 280 nm of each fraction (4 mL) was measured, and appropriate fractions were pooled. Pooled fractions were concentrated with a Vivaspin concentrator (20 mL, molecular mass cutoff: 10 kDa; Vivascience AG, Hanover, Germany) and dialyzed against 25 mM sodium phosphate, 10 mM glycine, and 50 mM NaCl, pH 7.8 in preparation for anion-exchange chromatography.

Cation-exchange chromatography was used as an alternative purification step for the truncated P4H, which did not bind to the poly(proline) resin. The 30–65% ammonium sulfate fraction was dialyzed against 25 mM sodium phosphate, 10 mM glycine, and 50 mM NaCl, pH 7.0, and then loaded onto a column (20 mL) of Mono S cation-exchange resin (HiLoad 16.60, Amersham Biosciences) that had been equilibrated with the same buffer. The anionic P4H protein did not bind to the resin and was washed through with 60 mL of buffer. P4H was concentrated with a Vivaspin concentrator and dialyzed against 25 mM sodium phosphate, 10 mM glycine, and 50 mM NaCl, pH 7.8 in preparation for anion-exchange chromatography.

In the next purification step, dialyzed protein from affinity chromatography (or cation-exchange chromatography) was injected at 1 mL/min onto a column (8 mL) of Resource Q anion-exchange resin (Amersham Biosciences) that had been equilibrated with 25 mM sodium phosphate, 10 mM glycine, and 50 mM NaCl, pH 7.8. After washing with equilibration buffer (16 mL), the column was eluted with a linear gradient (60 + 60 mL) of NaCl (50–430 mM) in equilibration buffer. Pooled fractions were concentrated with a Vivaspin concentrator and dialyzed overnight against 100 mM glycine, 10 mM Tris, and 100 mM NaCl, pH 7.8.

In the final purification step, dialyzed protein from anion-exchange chromatography was loaded onto a column (318 mL) of Superdex-200 gel-filtration resin (Amersham Biosciences) that had been preequilibrated with 100 mM glycine, 10 mM Tris, and 100 mM NaCl, pH 7.8. The column was eluted at a flow rate of 1.5 mL/min. Peak fractions were combined and concentrated. A correlation between protein molecular mass and column retention time was made by constructing a standard curve using a calibration kit (29–700 kDa) for gel-filtration chromatography (Sigma Chemical).

Analysis of P4H

During the purification of P4H, protein fractions were analyzed by SDS–PAGE in a gel containing 10% (w/v) poly(acrylamide) under reducing conditions, followed by staining with Coomassie brilliant blue R-250 or immunoblotting. Immunoblot analysis was performed with a polyclonal antibody to human PDI, which was a generous gift from Alan D. Attie, and anti-human P4Hα(I) monoclonal antibody (ICN Biomedicals, Costa Mesa, CA). Primary antibodies were detected with HRP-conjugated secondary antibodies and ECL Western blotting detection reagents (Amersham Biosciences).

Native PAGE was performed under non-reducing conditions to estimate the molecular mass of the truncated P4H. In this analysis, the truncated P4H and molecular mass standards were subjected to electrophoresis on native gels containing 6, 8, 10, and 12% (w/v) poly(acrylamide).

In-gel trypsin digestion of proteins followed by MALDI-MS was performed at the University of Wisconsin Biotechnology Center. Protein bands excised from SDS–PAGE gels were destained and digested with trypsin. MALDI-MS was used to detect peptides isolated from the digest. The Mascot search engine from Matrix Science (www.matrixscience.com) was used to perform a peptide mass fingerprint search, allowing carbamidomethyl cysteine and oxidation of methionine as variable modifications, a peptide mass tolerance of ±1 kDa, and a maximum of one missed cleavage.

Measurement of protein concentrations

Total protein concentration was determined for crude samples with the Micro BCA Protein Assay Reagent Kit (Pierce Biotechnology, Rockford, IL) according to the manufacturer's instructions, except that samples were incubated at 37 °C for 2 h before taking absorbance measurements. Extinction coefficients were determined at 280 nm for P4H by using the method of von Hippel and co-worker [46]. Values obtained ($\varepsilon_{280} = 290{,}000\ M^{-1}\,cm^{-1}$ for full-length P4H; $\varepsilon_{280} = 190{,}000\ M^{-1}\,cm^{-1}$ for P4Hα(I)235–534/β) were used to calculate the P4H concentration in purified samples.

Synthesis of P4H substrates

A fluorescent tetrapeptide substrate (dansyl-Gly-Phe-Pro-Gly-OEt) was synthesized by standard solution-phase methods. Boc-Phe-Pro-OH (870 mg, 2.4 mmol), glycine ethyl ester · HCl (341 mg, 2.4 mmol), dicyclohexylcarbodiimide (623 mg, 2.4 mmol), 1-hydroxybenzotriazole (330 mg, 2.4 mmol), and *N*-ethylmorpholine (309 μL, 2.4 mmol) were dissolved in tetrahydro furan (10 mL), and the resulting solution was stirred at room temperature overnight. The product was extracted into CH_2Cl_2 (50 mL) from saturated aqueous $NaHCO_3$ (50 mL). The organic layer was washed with aqueous citric acid (50 mL, 5% w/v) and brine (50 mL), and concentrated by rotary evaporation. The Boc protecting group was removed by stirring the crude residue in CH_2Cl_2 (20 mL) with trifluoroacetic acid (7.3 mL) at 0 °C for 3 h. The intermediate product (Phe-Pro-Gly-OEt) was purified by silica gel chromatography, eluting with methanol (8% v/v) in CH_2Cl_2. Solvent was removed from the fractions containing the desired compound by rotary evaporation, and the residue was dissolved in dimethylformamide (20 mL). 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (631 mg, 3.3 mmol), HOBt (349 mg, 2.6 mmol), *N*-ethylmorpholine (714 μL, 5.6 mmol), and dansylglycine (802 mg, 2.6 mmol) were added, and the reaction mixture was allowed to stir for 48 h. Water (10 mL) was added to quench the reaction. The product was extracted into ethylacetate (50 mL) and washed with aqueous HCl (5% v/v, 50 mL), saturated aqueous $NaHCO_3$ (50 mL), and brine (50 mL). The initial aqueous layer was extracted further into CH_2Cl_2 (50 mL), and the CH_2Cl_2 layer was washed with aqueous HCl (5% v/v, 50 mL), saturated aqueous $NaHCO_3$ (50 mL), and brine (50 mL). The organic layers were combined and dried over anhydrous $MgSO_4$(s), and the solvent was removed by rotary evaporation. The product was purified by silica gel chromatography, eluting with methanol (4% v/v) in CH_2Cl_2 to yield dansyl-Gly-Phe-Pro-Gly-OEt as a bright yellow oil (328 mg, 21.4%). ESI-MS *m*/*z*: $[M + Na]^+$ calcd, 660.2468; found, 660.2438.

The corresponding product peptide (dansyl-Gly-Phe-Hyp-Gly-OEt) was used as an HPLC standard. The Hyp-containing peptide was synthesized by using standard Fmoc-protection strategies with HATU activation on an Applied Biosystems Pioneer automated synthesizer. Dansyl-Gly-Phe-Hyp-Gly-OH was cleaved from the solid support with trifluoroacetic acid (1 mL) and dripped into ice-cold ethyl ether. The precipitated peptide was collected by centrifugation and the crude product (341 mg, 0.5 mmol) was dried under vacuum. The dried peptide was heated at reflux with $SOCl_2$ (80 mL, 1.1 mmol) in ethanol (20 mL) for 2 h. The solvent was removed by rotary evaporation, and the product was purified by silica gel chromatography, eluting with methanol (6% v/v) in CH_2Cl_2 to yield dansyl-Gly-Phe-Hyp-Gly-OEt as a yellow oil (77 mg, 16.0%). ESI-MS *m*/*z*: $[M + Na]^+$ calcd, 676.2417; found, 676.2422.

Assays of enzymatic activity

An HPLC-based assay was developed to monitor hydroxylation of proline by P4H. Assays were performed at 30 °C in 100 μL of 50 mM Tris–HCl buffer, pH 7.8 containing bovine serum albumin (1 mg/mL), catalase (100 μg/mL), dithiothreitol (100 μM), $FeSO_4$

(50 μM), P4H, and α-ketoglutarate (500 μM). The tetrapeptide substrate (dansyl-Gly-Phe-Pro-Gly-OEt, 92 mM stock in methanol) was added to initiate the reaction. Aliquots (30 μL) were withdrawn at known times and quenched by boiling for 30 s. All assays were performed in triplicate.

Conversion to hydroxylated product (dansyl-Gly-Phe-Hyp-Gly-OEt) was monitored by HPLC using a reversed-phase analytical C18 column (4.6 × 250 mm) from Varian (Lake Forest, CA) equipped with an Adsorbosphere guard cartridge from Alltech (Deerfield, IL). Peptides were separated by isocratic elution with acetonitrile:water (1:1) at 1.0 mL/min and quantified by integration of the $A_{337.5\,nm}$ according to a processing method developed with the Millennium32 software. Calibration curves were developed for the substrate and product peptides so as to correlate peak area with peptide concentration. The identity of HPLC peaks was confirmed by liquid chromatography/mass spectrometry (LC/MS) at the University of Wisconsin Biotechnology Center.

A more traditional P4H assay, which monitors the release of $[^{14}C]CO_2$ from $[1\text{-}^{14}C]$α-ketoglutarate, was used to confirm the validity of the new HPLC-based assay. Methods followed those described elsewhere [23,30]. Briefly, bovine serum albumin (1 mg/mL), catalase (100 μg/mL), dithiothreitol (100 μM), ascorbate (2 mM), $FeSO_4$ (50 μM), P4H, and a substrate were combined (500 μL total reaction volume) on ice. Substrates used in this assay include the tetrapeptide described above as well as $(Pro\text{-}Pro\text{-}Gly)_{10}$. The latter substrate was boiled for 5 min and then placed on ice, immediately before performing the reaction. A hanging well that contained filter paper soaked with aqueous NaOH (2 M) was inserted through a rubber stopper that was used to seal each reaction. $[1\text{-}^{14}C]$α-ketoglutarate (100 μM, 185,000 dpm) was added to initiate the reaction. After 15 min at 30 °C, reactions were quenched by the addition of 500 μL of aqueous HCl (2 M) and allowed to sit at room temperature for 60 min. The filter paper was removed and dried under a bright light for 1.5 h. Liquid scintillation cocktail (3 mL, Wallac Optiphase SuperMix from Perkin–Elmer, Wellesley, MA) was added to each piece of paper, and ^{14}C was quantified by scintillation counting (5 min). All reactions were performed in duplicate and corrected for the rate of decarboxylation in the absence of the peptide substrate.

Results

Production of P4H from pBK1.PDI1.P4H5

The bicistronic plasmid pBK1.PDI1.P4H5 was constructed for the production of human P4H by placing DNA encoding the α(I) and β subunits into pET22b(+), under control of the same T7 promoter, and giving each cDNA its own rbs. Expression of cDNAs encoding both subunits was induced in BL21(DE3) cells as indicated by SDS–PAGE (data not shown), but the cell lysate from these cells had no detectable prolyl 4-hydroxylase activity. In addition, no tetrameric protein could be detected by native PAGE (data not shown). Apparently, the assembly of active P4H tetramer was not achieved in BL21(DE3) cells. Addition of iron(II) sulfate, zinc(II) chloride, ascorbate, and/or α-ketoglutarate to the growth media did not promote assembly. Likewise, inclusion of denatured collagen or denatured $(Pro\text{-}Pro\text{-}Gly)_{10}$ in the lysis buffer did not assemble an active tetramer. Finally, attempts were made to assemble P4H tetramer in vitro. Refolding of insoluble P4Hα(I) expressed from pBK1.P4H5 with cell lysate containing PDI expressed from pBK1.PDI1 in the presence of glutathione, ascorbate, iron(II) sulfate, glycine, arginine, and/or denatured collagen was not successful.

Instead, pBK1.PDI1.P4H5 was induced in Origami B(DE3) cells at 23 °C. SDS–PAGE indicated that induction was successful (Fig. 3A). In addition, prolyl

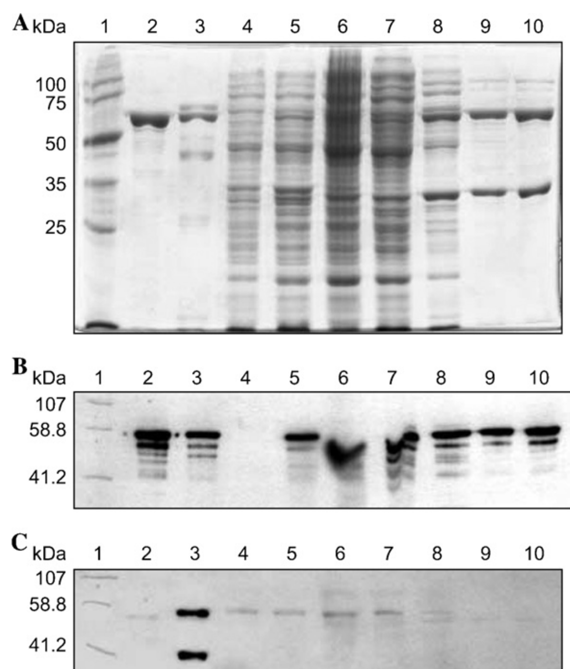

Fig. 3. Ten percent SDS–PAGE analysis of the production and purification of P4H from pBK1.PDI1.P4H5 in Origami B(DE3) cells, which results in a truncated α subunit. Lane 1, molecular mass standards; lane 2, human PDI standard; lane 3, human P4H standard; lane 4, uninduced whole cells; lane 5, induced whole cells; lane 6, crude cell lysate; lane 7, 30–65% ammonium sulfate fraction; lane 8, pooled fractions from cation-exchange chromatography; lane 9, pooled fractions from anion-exchange chromatography; and lane 10, pooled fractions from gel-filtration chromatography. (A) Analysis by Coomassie staining. (B) Analysis by immunoblotting using an anti-human PDI polyclonal antibody. (C) Analysis by immunoblotting using an anti-human α(I) subunit monoclonal antibody, which does not recognize the truncated α subunit.

4-hydroxylase activity was detected in a crude cell lysate (data not shown).

Purification of truncated P4H

SDS–PAGE of whole cell samples after induction of pBK1.PDI1.P4H5 in Origami B(DE3) cells revealed the induction of a 32-kDa protein, in addition to a protein (PDI) of 55 kDa (Fig. 3A). Interestingly, induction of the 32-kDa protein was not observed in BL21(DE3) cells (data not shown). P4H in the 30–65% ammonium sulfate fraction of this cell lysate could not be purified with poly(proline)-affinity chromatography. Instead, the enzyme was purified by cation-exchange chromatography. Although P4H did not bind to a cation-exchange resin, many other proteins did (Fig. 3A). Additional purification was achieved by anion-exchange chromatography, followed by gel-filtration chromatography. Following these two purification steps, SDS–PAGE revealed that >90% of the protein was in two bands corresponding to proteins of approximately 55 and 32 kDa (Fig. 3A). Approximately, 2 mg of this P4H oligomer was obtained per liter of growth medium.

The 32-kDa protein was digested with trypsin in situ after SDS–PAGE. The protein fragments indicated that the 32-kDa protein derived from the C-terminus of P4Hα(I). Specifically, 10 peptides had masses corresponding to trypsin fragments of the human P4Hα(I) subunit between Gly260 and Glu534, which is the most C-terminal residue. This α subunit fragment is not likely to be an *E. coli* degradation product, as it is observed in whole cell extract immediately following induction with IPTG.

No induction of full-length α subunit was apparent. In-gel trypsin digestion and MALDI-MS of the purified 55-kDa band from SDS–PAGE revealed eight peptides from PDI but none from the α subunit. Immunoblot analysis of crude cell lysate and purified protein showed a strong response with a polyclonal PDI antibody, but no detectable response with a monoclonal antibody against human P4Hα(I) (Figs. 3B and C).

A tetrameric P4H protein ($\alpha_2\beta_2$) containing PDI (55 kDa) and the C-terminal fragment of P4Hα(I) (32 kDa) would have a molecular mass of 174 kDa, whereas a dimeric P4H protein (αβ) would have a molecular mass of 87 kDa. The elution of the truncated P4H during analytical gel-filtration chromatography corresponds to a protein of 140 kDa, which is between the mass expected for dimer and tetramer. Full-length P4H eluted during analytical gel-filtration chromatography at 30 mL after the void volume, which corresponds to a protein of 380 kDa. Yet, the molecular mass of full-length P4H is 228 kDa. Thus, the structure of the wild-type tetramer is likely to be non-spherical. If the structure of the truncated P4H is also non-spherical, then the stoichiometry of its subunits cannot be determined by analytical gel-filtration chromatography.[2]

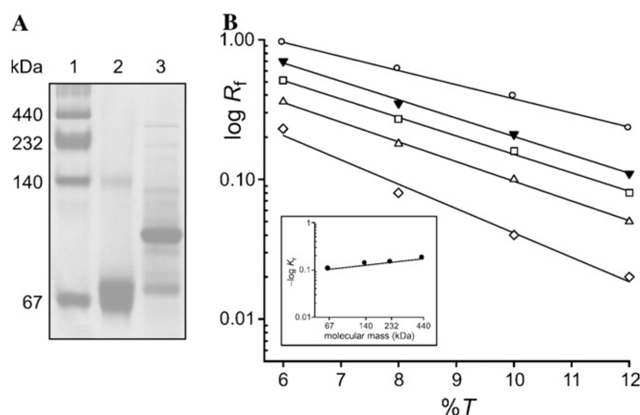

Fig. 4. Native PAGE analysis of purified truncated P4H. (A) Representative Coomassie-stained gel containing 7.5% (w/v) poly(acrylamide). Lane 1, molecular mass standards; lane 2, human PDI standard; and lane 3, truncated P4H, expressed from pBK1.PDI1.P4H5 in Origami B(DE3) cells and purified. (B) Log plot of polyacrylamide concentration (%T) versus relative mobility (R_f). BSA (circles, 67 kDa, $K_r = -0.101$), lactate dehydrogenase (squares, 140 kDa, $K_r = -0.132$), catalase (open triangles, 232 kDa, $K_r = -0.141$), and ferritin (diamonds, 440 kDa, $K_r = -0.174$) were used to construct a log plot (inset) of molecular mass versus $-K_r$. Log plot of truncated P4H (filled triangles) gives $K_r = -0.132$, which corresponds to a molecular mass of 140 kDa.

Native PAGE of the purified protein was performed with a range of poly(acrylamide) concentrations (6–12%) alongside a set of proteins of known molecular mass (Fig. 4A). A calibration plot that correlates relative mobility ($\log R_f$) to gel concentration (%T) was used to obtain a slope (K_r) for each protein standard (Fig. 4B). These data were used to construct a Ferguson plot of $-\log K_r$ versus log molecular mass. By using these standard curves, the molecular mass of truncated P4H was estimated to be 140 kDa, which is in gratifying agreement with the results of analytical gel filtration chromatography. Heretofore, we refer to this truncated P4H as P4Hα(I)235–534/β.

Catalysis by truncated P4H

The prolyl hydroxylase activity of P4Hα(I)235–534/β was determined by measuring the release of [^{14}C]CO_2 from [1-^{14}C]α-ketoglutarate during the hydroxylation of $(\text{Pro-Pro-Gly})_{10}$ (10–1000 μM) and dansyl-Gly-Phe-Pro-Gly-OEt (10–1000 μM) peptide substrates (Table 2). The truncated enzyme has weak binding constants for both peptide substrates, and the K_m for $(\text{Pro-Pro-Gly})_{10}$

[2] If the (actual mass)/(predicted mass) ratio is similar for the full-length and truncated P4H, then the actual mass of the truncated P4H is (228 kDa/380 kDa) × 140 kDa = 84 kDa, which is most consistent with the truncated P4H being an αβ dimer (87 kDa). This stoichiometry has been observed in *C. elegans* P4H [8].

(0.96 mM) is greater than for the tetrapeptide (0.23 mM). The rate of hydroxylation is an order of magnitude slower for the tetrapeptide than that for $(Pro\text{-}Pro\text{-}Gly)_{10}$, although saturating conditions could not be reached due to the limited solubility of the substrate peptides.

Production of full-length P4H

We tested two characteristics of the P4H expression vector as being responsible for the production of a truncated α subunit. First, we suspected that the translation initiation region for P4Hα(I) in the pBK1.PDI1.P4H5 vector had significant secondary structure that could hinder ribosome binding. To address this possibility, we constructed a plasmid (pBK1.PDI1.P4H6) with the same translation initiation region, taken directly from plasmid pET22b(+), for both PDI and P4Hα(I). The 32-kDa P4Hα(I) was still expressed from pBK1.PDI1.P4H6 (data not shown), suggesting that ribosome binding was not responsible for the truncation.

Second, we noted that P4Hα(I) contains nine methionine residues that could serve as alternative start codons. Translation initiation at Met235 would produce a C-terminal α subunit of 32.2 kDa, in agreement with the truncated α subunit observed by SDS–PAGE. To address this possibility, we used site-directed mutagenesis to replace Met235 with a leucine residue (which has a similar hydrophobicity and volume), yielding plasmid pBK1.PDI1.P4H7.

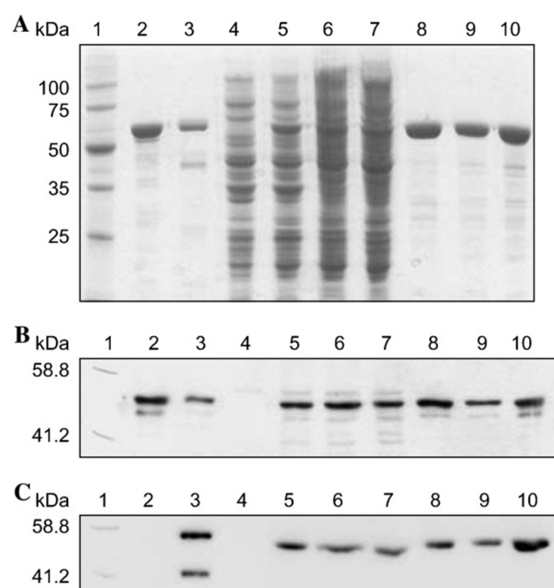

Fig. 5. Ten percent SDS–PAGE analysis of the production and purification of recombinant human P4H from pBK1.PDI1.P4H7 in Origami B(DE3) cells. Lane 1, molecular mass standards; lane 2, human PDI standard; lane 3, human P4H standard; lane 4, uninduced whole cells; lane 5, induced whole cells; lane 6, crude cell lysate; lane 7, 30–65% ammonium sulfate fraction; lane 8, pooled fractions from poly (proline)-affinity chromatography; lane 9, pooled fractions from anion-exchange chromatography; and lane 10, pooled fractions from gel-filtration chromatography. (A) Analysis by Coomassie staining. (B) Analysis by immunoblotting using an anti-human PDI polyclonal antibody. (C) Analysis by immunoblotting using an anti-human α(I) subunit monoclonal antibody.

By making this final mutation to the bicistronic expression vector, we were able to produce full-length α subunit in Origami B(DE3) cells. SDS–PAGE analysis of crude whole cell extract revealed the absence of an induced 32-kDa band (Fig. 5A). The full-length α and β subunits were of similar molecular mass (55.3 and 59.8 kDa, respectively), and hence could not be distinguished easily by Coomassie staining. Immunoblot analysis did, however, reveal the induction of both subunits in whole cell extracts (Figs. 5B and C). Apparently, Met235 provides an alternative start codon for the translation of P4Hα(I).

Production of P4H from plasmid pBK1.PDI1.P4H7 was also attempted in Rosetta-gami 2(DE3) cells. These cells combine an oxidizing cytosol with *t*RNAs that are rare in *E. coli*. Induction with 500 μM IPTG at 23 °C yielded no P4H according to SDS–PAGE analysis (data not shown).

P4H expressed from plasmid pBK1.PDI1.P4H7 bound tightly to a poly(proline)-affinity column (Fig. 5A). Purified enzyme was eluted with a 5 mg/mL solution of poly(proline), and then applied to an anion-exchange column, which also served to separate poly(proline) from the enzyme. P4H eluted as a single peak at 250 mM NaCl. A final gel-filtration step yielded protein that was >90% pure as estimated by SDS–PAGE, which represents a 148-fold purification over the ammonium sulfate fraction (Table 1). Approximately, 2 mg of P4H tetramer was obtained per liter of growth medium (41% yield).

Table 1
Purification of full-length tetrameric prolyl 4-hydroxylase[a]

Purification step	Total protein (mg)	Activity[b] (10^3 U)	Specific activity (U/mg)	Yield (%)	Purification (fold)
30–65% ammonium sulfate fraction	591	9.2	15.5	100	1
Anion-exchange chromatography	2.0	3.7	1880	40	121
Gel-filtration chromatography	1.5	3.8	2300	41	148

[a] Data are for a 1.0-L culture of Origami B(DE3) cells grown in TB medium. Data for the poly(proline)-affinity chromatography step are not shown because poly(proline) that leaches from the resin is a potent inhibitor of the enzyme.

[b] One unit of prolyl 4-hydroxylase is the amount required to hydroxylate 1 μM dansyl-Gly-Phe-Pro-Gly-OEt to form dansyl-Gly-Phe-Hyp-Gly-OEt per min.

Table 2
Kinetic parameters for prolyl hydroxylation by P4Hα (I)235–534/β and full-length P4H[a]

Parameter	P4Hα(I)235–534/β[b]		Full-length P4H[c]	
	$(PPG)_{10}$	dansyl-GFPG-OEt	$(PPG)_{10}$	dansyl-GFPG-OEt
K_m (mM)	0.96 ± 0.43	0.23 ± 0.06	0.011 ± 0.002	0.078 ± 0.032
V_{max} ($\mu M\ min^{-1}$)	4.5 ± 1.2	0.48 ± 0.05	4.7 ± 0.2	3.3 ± 0.4
k_{cat} (min^{-1})	4.5 ± 1.2	0.48 ± 0.05	13 ± 1	9.2 ± 1.2
k_{cat}/K_m ($M^{-1}\ s^{-1}$)	78 ± 46	35 ± 14	$(2.0 \pm 0.6) \times 10^4$	$(1.8 \pm 0.6) \times 10^3$

[a] Data were obtained by measuring $[^{14}C]CO_2$ released during the decarboxylation of $[1-^{14}C]\alpha$-ketoglutarate. Reactions were performed in Tris–HCl (50 mM, pH 7.8) at 30 °C for 15 min. Reaction components were as described in the text. $(Pro-Pro-Gly)_{10}$ (2–1000 μM) or dansyl-Gly-Phe-Pro-Gly-OEt (10–1000 μM) was used as substrate.
[b] Purified P4Hα(I)235–534/β (90 μg/mL) was used in the reactions. Calculations for k_{cat} are for an αβ dimer.
[c] Purified full-length P4H (82 μg/mL) was used in the reactions. Calculations for k_{cat} are for an $\alpha_2\beta_2$ tetramer.

Catalysis by full-length P4H

Kinetic analysis of the full-length P4H enzyme produced in *E. coli* was performed by measuring the release of $[^{14}C]CO_2$ from $[1-^{14}C]\alpha$-ketoglutarate during the hydroxylation of $(Pro-Pro-Gly)_{10}$ (2–200 μM) and dansyl-Gly-Phe-Pro-Gly-OEt (10–500 μM) peptide substrates at 30 °C (Table 2). $(Pro-Pro-Gly)_{10}$ binds more than 80-fold more tightly to full-length enzyme than to the truncated enzyme, whereas the tetrapeptide substrate binds 3-fold more tightly to the full-length than to the truncated enzyme. The k_{cat} for hydroxylation of both substrates is approximately 10 min^{-1}. This rate is approximately 60-fold less than that reported for the hydroxylation of $(Pro-Pro-Gly)_{10}$ at 37 °C by P4H tetramer produced in a baculovirus system [14]. The reported K_m value for this substrate is 18 μM [14], which agrees well with that measured herein for P4H produced in *E. coli.*

HPLC-based assay for enzymatic activity

An HPLC-based assay was developed to measure the in vitro enzymatic activity of P4H enzyme. This method was based on a previously published substrate [47,48] that had been monitored for hydroxylation by thin-layer chromatography. Here, the fluorescent tetrapeptide substrate (dansyl-Gly-Phe-Pro-Gly-OEt) was resolved from its hydroxylated product (dansyl-Gly-Phe-Hyp-Gly-OEt) during reversed-phase HPLC, requiring only a 15-min isocratic elution with acetonitrile:water (1:1) (Fig. 6). The identity of each peak was confirmed independently by LC/MS analysis (data not shown).

The validity of our hydroxylation assay was cNSonfirmed by comparing it to the traditional CO_2-release assay (Table 3). The K_m value for the turnover of the tetrapeptide substrate by purified full-length human P4H tetramer is 94 μM by the hydroxylation assay, which is within error of the K_m value obtained with the CO_2-release assay (87 μM). The k_{cat} value of 19 min^{-1} obtained by HPLC analysis is twofold greater than that obtained with the CO_2-release assay (9.2 min^{-1}).

Discussion

Recombinant human P4H tetramer was successfully expressed and assembled from a bicistronic expression vector in Origami B(DE3) cells. This *E. coli* expression strain contains mutations in the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which enhance the ability of the cells to form protein disulfide bonds in the *E. coli* cytosol [49,50]. Previous experiments have revealed that P4H subunits can be dissociated by reduction with dithiothreitol in vivo and in vitro [33,51]. Studies of P4H assembly in a cell-free system, using translated α subunit mRNA and endogenous PDI, showed that a properly assembled α subunit contains an intramolecular disulfide bond [52]. Site-directed mutagenesis of the five conserved α subunit cysteine residues has shown that Cys276 and Cys293 form an intramolecular disulfide bond that is essential for tetramer assembly [52] and that Cys486 and Cys511 are required for assembly of active tetramers, possibly through an intramolecular disulfide bond [9,52]. In vivo, P4H dissociated by dithiothreitol can be reassociated oxidatively through interactions with the cellular chaperone BiP [34]. In vitro, however, it has not been possible to reassociate P4H subunits [32,33]. In the reducing environment of the *E. coli* cytosol, it is likely that the intramolecular disulfide bond(s) of the α subunit cannot form, precluding it from assembling with PDI into the $\alpha_2\beta_2$ tetramer (as observed in BL21(DE3) cells). Allowing P4Hα(I) to form disulfide bond(s) in the relatively oxidizing cytosol of Origami B(DE3) cells enables tetramer assembly in the absence of any other eukaryotic proteins.

Surprisingly, induction of the initial bicistronic expression vector (pBK1.PDI1.P4H5) in Origami B(DE3) cells resulted in the production of a truncated α subunit. This subunit is synthesized from an alternative start codon (Met235) of P4Hα(I), and folds in the relatively oxidizing cytosol of the Origami B(DE3) cells but not the reducing cytosol of BL21(DE3) cells. Furthermore, under these conditions, this plasmid is unable to direct the production of full-length α subunit. Met235 is conserved among P4Hα(I) subunits from human, mouse,

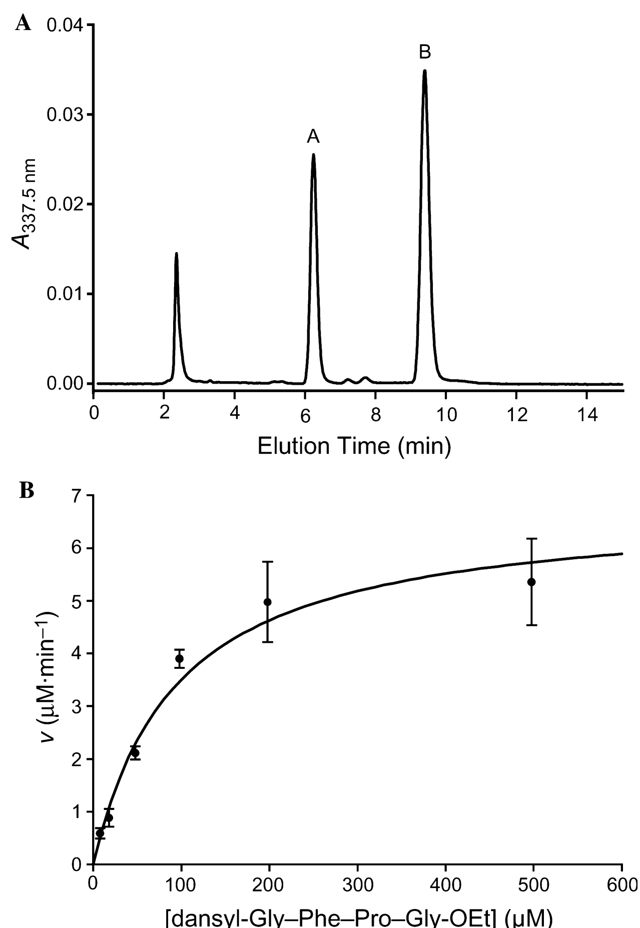

Fig. 6. HPLC-based assay for P4H activity. (A) HPLC trace of a time point in a representative assay of dansyl-Gly-Phe-Pro-Gly-OEt hydroxylation. Enzymatic reactions were injected onto an analytical reversed-phase C18 column and peptides were eluted isocratically with acetonitrile:water (1:1). Peak A is the Hyp-containing tetrapeptide product. Peak B is the Pro-containing tetrapeptide substrate. (B) Catalysis of dansyl-Gly-Phe-Pro-Gly-OEt hydroxylation by full-length human P4H. Reactions mixtures (100 µL) were 50 mM Tris–HCl buffer, pH 7.8 containing BSA (1 mg/mL), catalase (0.1 mg/mL), DTT (100 µM), ascorbate (2 mM), $FeSO_4$ (50 µM), P4H (0.36 µM), α-ketoglutarate (100 µM), and tetrapeptide substrate (10–500 µM). Reactions were performed at 30 °C for 15 min and quenched by boiling for 30 s. Aliqouts (20 µL) were analyzed as in panel A, with Peak A being quantified by integration of $A_{337.5\ nm}$. Individual points are the average (±SE) of three independent hydroxylation reactions. To obtain kinetic parameters, data were fitted to the Michaelis–Menten equation: $v = V_{max}[S]/(K_m + [S])$.

rat, and chicken [19,53–55], but is replaced by a leucine residue in the α(II) and α(III) isoforms of these organisms [17,18,54]. Whether translation at Met235 plays any natural role is unknown.

The 32-kDa α subunit that was synthesized by Origami B(DE3) cells contains only the C-terminal residues 235–534, but retains the ability to catalyze prolyl hydroxylation (Table 2). A number of site-directed mutagenesis studies have shown that the catalytic residues of P4H are positioned at the C-terminus of the α subunit. His412, Asp414, and His483 are the iron-binding ligands [8,9], while Lys493 and His501 bind α-ketoglutarate [8]. All of these residues are present in P4Hα(I)235–534/β. The binding of collagen substrates, on the other hand, involves residues that are more N-terminal than Met235. Limited proteolysis studies have revealed that the peptide-binding site is distinct from the catalytic domain and is composed of residues 140–240 [56]. Indeed, a Phe144–Ser244 polypeptide has a K_d value for $(Pro\text{-}Pro\text{-}Gly)_{10}$ that is nearly identical to the K_m value for the turnover of $(Pro\text{-}Pro\text{-}Gly)_{10}$ by the P4H tetramer [57]. These residues are predominantly missing from the truncated α subunit produced from plasmid pBK1.PDI1.P4H5. Studies with model peptides have suggested that the binding of collagen substrates to P4H involves the β-turn of a Pro-Gly sequence preceded by a polyproline type-II helix [58]. In addition to binding one proline residue in the active site for catalysis, the enzyme binds the extended strand in additional subsites that likely exist in the more N-terminal portion of the α subunit. The inability of the P4Hα(I)235–534 subunit to bind to a poly(proline)-affinity column supports this location for its substrate binding subsites.

Kinetic studies reveal that P4Hα(I)235–534/β binds substrates more weakly than does the full-length P4H tetramer (Table 2). In addition, the truncated α subunit binds $(Pro\text{-}Pro\text{-}Gly)_{10}$ even more weakly than dansyl-Gly-Phe-Pro-Gly-OEt, in opposition to the trend observed for full-length P4H (Table 2). Typically, P4H binds longer substrates more tightly than shorter substrates [31], presumably due to additional favorable interactions with subsites of the α subunit.

By removing the alternative start codon, full-length human P4H tetramer was produced in Origami B(DE3) cells and purified to >90% homogeneity. The recombinant

Table 3
Comparison of HPLC-based assay to [^{14}C]CO_2-release assay[a]

Assay method	K_m (µM)	V_{max} (µM min^{-1})	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
[^{14}C]CO_2 release[b]	78 ± 32	3.3 ± 0.4	9.2 ± 1.2	(1.8 ± 0.6) × 10^3
HPLC[c]	94 ± 19	6.8 ± 0.5	19 ± 1.5	(3.3 ± 1.3) × 10^3

[a] Reactions were performed in 50 mM Tris–HCl buffer (pH 7.8) at 30 °C for 15 min. Reaction components were as described in the text. dansyl-Gly-Phe-Pro-Gly-OEt substrate (10–1000 µM) and full-length P4H (82 µg/mL) purified from *E. coli* were used in both assays.
[b] [1-^{14}C]α-ketoglutarate (100 µM, 185,000 dpm) was added to initiate the reaction.
[c] Peptide substrate was added to initiate the reaction.

enzyme has kinetic properties that are comparable to P4H produced in other heterologous systems. The K_m value measured here for $(Pro\text{-}Pro\text{-}Gly)_{10}$ is 11 μM, which is similar to that measured for P4H purified from a baculovirus system (18 μM) [14]. The k_{cat} for hydroxylation of $(Pro\text{-}Pro\text{-}Gly)_{10}$ is lower for P4H purified in this study ($13\ min^{-1}$) than that reported for P4H produced in baculovirus ($12.5–13\ s^{-1}$) [14]. A difference in reaction temperature (30 °C for *E. coli* -produced P4H and 37 °C for baculovirus-produced P4H) could contribute to this difference.

The HPLC-based hydroxylation assay described herein has advantages over the traditional CO_2-release assay for P4H activity. CO_2(g) release necessarily provides an indirect measurement of hydroxylation, which can be problematic. In the absence of sufficient substrate, P4H mediates the uncoupled decarboxylation of α-ketoglutarate, which does not produce Hyp and inactivates the enzyme by an as yet unidentified mechanism [27,59]. Even under conditions of saturating substrate, this uncoupled reaction occurs to a small degree. For this reason, ascorbate is included in all activity assays to rescue enzyme trapped in the Fe(III) state. The HPLC-based assay provides a direct measurement of hydroxylation, thereby avoiding complications arising from the uncoupled reaction. Moreover, the HPLC-based assay avoids the use of radioactivity and the complexities of trapping a gaseous product. Finally, the CO_2-release assay method requires comparatively large amounts of enzyme, as the measurement of $[^{14}C]CO_2$ release requires that each time point derive from a separate reaction mixture. In contrast, the HPLC-based assay allows for the removal of aliquots from a single reaction mixture.

The HPLC-based assay has additional advantages over less common assays that monitor hydroxylation directly by using radiolabeled procollagen substrates [30,60]. These assays require the tedious separation of ^{14}C- or ^{3}H-labeled hydroxyproline from radiolabeled proline prior to scintillation counting. In contrast, the HPLC-based assay allows for the direct analysis of the reaction mixture in an automated manner.

To assess the accuracy of the new assay, kinetic constants obtained for hydroxylation of a tetrapeptide substrate by the HPLC-based assay were compared to those from the more traditional CO_2-release assay. The results obtained provide validation for its continued use. The K_m values obtained by the two assays are in gratifying agreement. The difference in k_{cat} values could be explained (at least in part) by incomplete trapping of $[^{14}C]CO_2$ during the CO_2-release assay, providing yet another benefit to the assay developed here.

Collagen is a common biomaterial [61,62]. Bovine collagen, which is most often used in this context, can engender allergic and immunological side effects in humans, as well as other health risks. Collagen has been produced previously in the yeast *Pichia pastoris* by the coexpression of cDNAs for P4H and procollagen [35]. Likewise, the expression system described herein could be used to produce natural human collagen in *E. coli*.

Acknowledgments

We are grateful to Drs. Kari Kivirikko and Johanna Myllyharju for valuable suggestions and their kind gift of human P4H for use as a standard in the development of our expression system. We thank Dr. Gary Case, Dr. Amy Harms, and Grzegorz Sabat of the University of Wisconsin Biotechnology Center for help with peptide synthesis and analysis. E.A.K. was supported by WARF and Steenbock predoctoral fellowships and Biotechnology Training Grant GM08349 (NIH). This work was supported by Grant AR44276 (NIH).

References

[1] R.K. Chopra, V.S. Ananthanarayanan, Conformational implications of enzymatic proline hydroxylation in collagen, Proc. Natl. Acad. Sci. USA 79 (1982) 7180–7184.

[2] N.J. Bulleid, R. Wilson, J.F. Lees, 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 (1996) 195–202.

[3] A. Snellman, M.-R. Keränen, P.O. Hägg, A. Lamberg, K. Hiltunen, K.I. Kivirikko, T. Pihlajaniemi, Type XIII collagen forms homotrimers with three triple helical collagenous domains and its association into disulfide-bonded trimers is enhanced by prolyl 4-hydroxylase, J. Biol. Chem. 275 (2000) 8936–8944.

[4] R.A. Berg, D.J. Prockop, The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen, Biochem. Biophys. Res. Commun. 52 (1973) 115–120.

[5] T.V. Burjanadze, New analysis of the phylogenetic change of collagen thermostability, Biopolymers 53 (2000) 523–528.

[6] L. Friedman, J.J. Higgin, G. Moulder, R. Barstead, R.T. Raines, J. Kimble, Prolyl 4-hydroxylase is required for viability and morphogenesis in *Caenorhabditis elegans*, Proc. Natl. Acad. Sci. USA 97 (2000) 4736–4741.

[7] A.D. Winter, A.P. Page, Prolyl 4-hydroxylase is an essential procollagen-modifying enzyme required for exoskeleton formation and the maintenance of body shape in the nematode *Caenorhabditis elegans*, Mol. Cell. Biol. 20 (2000) 4084–4093.

[8] J. Myllyharju, K.I. Kivirikko, Characterization of the iron- and 2-oxoglutarate-binding sites of human prolyl 4-hydroxylase, EMBO J. 16 (1997) 1173–1180.

[9] A. Lamberg, T. Pihlajaniemi, K.I. Kivirikko, Site-directed mutagenesis of the α subunit of human prolyl 4-hydroxylase. Identification of three histidine residues critical for catalytic activity, J. Biol. Chem. 270 (1995) 9926–9931.

[10] J. Koivu, R. Myllylä, T. Helaakoski, T. Pihlajaniemi, K. Tasanen, K.I. Kivirikko, A single polypeptide acts both as the β subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase, J. Biol. Chem. 262 (1987) 6447–6449.

[11] N.A. Guzman (Ed.), Prolyl Hydroxylase, Protein Disulfide Isomerase, and Other Structurally Related Proteins, Marcel Dekker, New York, 1998.

[12] E.A. Kersteen, R.T. Raines, Catalysis of protein folding by protein disulfide isomerase and small-molecule mimics, Antiox. Redox. Signal. 5 (2003) 413–424.

[13] H.F. Gilbert, Protein disulfide isomerase and assisted protein folding, J. Biol. Chem. 272 (1997) 29399–29402.

[14] K. Vuori, T. Pihlajaniemi, M. Marttila, K.I. Kivirikko, 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 (1992) 7467–7470.

[15] K. Vuori, T. Pihlajaniemi, R. Myllylä, K.I. Kivirikko, Site-directed mutagenesis of human protein disulphide isomerase: effect on the assembly, activity and endoplasmic reticulum retention of human prolyl 4-hydroxylase in *Spodoptera frugiperda* insect cells, EMBO J. 11 (1992) 4213–4217.

[16] K.I. Kivirikko, R. Myllylä, T. Pihlajaniemi, Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit, FASEB J. 3 (1989) 1609–1617.

[17] P. Annunen, T. Helaakoski, J. Myllyharju, J. Veijola, T. Pihlajaniemi, K.I. Kivirikko, Cloning of the human prolyl 4-hydroxylase α subunit isoform α(II) and characterization of the type II enzyme tetramer (1997).

[18] L. Kukkola, R. Hieta, K.I. Kivirikko, J. Myllyharju, Identification and characterization of a third human, rat, and mouse collagen prolyl 4-hydroxylase isoenzyme, J. Biol. Chem. 278 (2003) 47685–47693.

[19] T. Helaakoski, K. Vuori, R. Myllylä, K.I. Kivirikko, T. Pihlajaniemi, 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 (1989) 4392–4396.

[20] K.I. Kivirikko, J. Myllyharju, Prolyl 4-hydroxylases and their protein disulfide isomerase subunit, Matrix Biol. 16 (1998) 357–368.

[21] B.G. Fox, Catalysis by non-heme iron, in: M. Sinnott (Ed.), Comprehensive Biological Catalysis: A Mechanistic Reference, Academic Press, San Diego, CA, 1998.

[22] K.I. Kivirikko, T. Pihlajaniemi, Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases, Adv. Enzymol. Relat. Areas Mol. Biol. 72 (1998) 325–398.

[23] R.E. Rhoads, S. Udenfriend, Decarboxylation of α-ketoglutarate coupled to collagen proline hydroxylase, Proc. Natl. Acad. Sci. USA 60 (1968) 1473–1478.

[24] G.J. Cardinale, R.E. Rhoads, S. Udenfriend, Simultaneous incorporation of 18O into succinate and hydroxyproline catalyzed by collagen proline hydroxylase, Biochem. Biophys. Res. Commun. 43 (1971) 537–543.

[25] L. de Jong, S.P.J. Albracht, A. Kemp, Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron. The role of ascorbate in peptidyl proline hydroxylation, Biochim. Biophys. Acta 704 (1982) 326–332.

[26] R. Myllylä, E.-R. Kuutti-Savolainen, K.I. Kivirikko, The role of ascorbate in the prolyl hydroxylase reaction, Biochem. Biophys. Res. Commun. 83 (1978) 441–448.

[27] R. Myllylä, K. Majamaa, V. Günzler, H.M. Hanauske-Abel, K.I. Kivirikko, Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase, J. Biol. Chem. 259 (1984) 5403–5405.

[28] M. Costas, M.P. Mehn, M.P. Jensen, L.J. Que, Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates, Chem. Rev. 104 (2004) 939–986.

[29] C.J. Schofield, Z. Zhang, Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes, Curr. Opin. Struct. Biol. 9 (1999) 722–731.

[30] K.I. Kivirikko, R. Myllylä, Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes, Methods Enzymol. 82 (1982) 245–304.

[31] K.I. Kivirikko, Y. Kishida, S. Sakakibara, D.J. Prockop, Hydroxylation of (X-Pro-Gly)n by protocollagen proline hydroxylase. Effect of chain length, helical conformation and amino acid sequence in the substrate, Biochim. Biophys. Acta 271 (1972) 347–356.

[32] J. Koivu, R. Myllylä, Protein disulfide-isomerase retains procollagen prolyl 4-hydroxylase structure in its native conformation, Biochemistry 25 (1986) 5982–5986.

[33] J.J. Nietfeld, J. vander Kraan, A. Kemp, Dissociation and reassociation of prolyl 4-hydroxylase subunits after cross-linking of monomers, Biochim. Biophys. Acta 661 (1981) 21–27.

[34] D.C.A. John, N.J. Bulleid, Intracellular dissociation 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 (1996) 659–665.

[35] A. Vuorela, J. Myllyharju, R. Nissi, T. Pihlajaniemi, K.I. Kivirikko, 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, EMBO J. 16 (1997) 6702–6712.

[36] S.J. Higgins, B.D. Hames, Protein Expression: A Practical Approach, Oxford University Press, Oxford, UK, 1999.

[37] L. Chu, D.K. Robinson, Industrial choices for protein production by large-scale cell culture, Curr. Opin. Biotechnol. 12 (2001) 180–187.

[38] D.C. Andersen, L. Krummen, Recombinant protein expression for therapeutic applications, Curr. Opin. Biotechnol. 13 (2002) 117–123.

[39] F. Baneyx, Recombinant protein expression in *Escherichia coli*, Curr. Opin. Biotechnol. 10 (1999) 411–421.

[40] J.R. Swartz, Advances in *Escherichia coli* production of therapeutic proteins, Curr. Opin. Biotechnol. 12 (2001) 195–201.

[41] C.W. Goulding, L.J. Perry, Protein production in *Escherichia coli* for structural studies by X-ray crystallography, J. Struct. Biol. 142 (2003) 133–143.

[42] J. Sambrook, and D. Russell, Molecular Cloning: A Laboratory Manual, third ed., 2001.

[43] K.I. Kivirikko, R. Myllylä, Recent developments in posttranslational modification: intracellular processing, Methods Enzymol. 144 (1987) 96–114.

[44] L. Tuderman, E.-R. Kuutti, K.I. Kivirikko, An affinity-column procedure using poly(L-proline) for the purification of prolyl hydroxylase, Eur. J. Biochem. 52 (1975) 9–16.

[45] U. Lindberg, C.E. Schutt, E. Hellsten, A.-C. Tjäder, T. Hult, The use of poly(L-proline)-sepharose in the isolation of profilin and profilactin complexes, Biochim. Biophys. Acta 967 (1988) 391–400.

[46] S. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Biochem. 182 (1989) 319–326.

[47] M. Tandon, M. Wu, T.P. Begley, Substrate specificity of human prolyl-4-hydroxylase, Bioorg. Med. Chem. Lett. 8 (1998) 1139–1144.

[48] M. Wu, H.-S. Moon, T.P. Begley, Mechanism-based inactivation of the human prolyl-4-hydroxylase by 5-oxaproline-containing peptides: evidence for a prolyl radical intermediate, J. Am. Chem. Soc. 121 (1999) 587–588.

[49] W.A. Prinz, F. Åslund, A. Holmgren, J. Beckwith, The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm, J. Biol. Chem. 272 (1997) 15661–15667.

[50] P.H. Bessette, F. Åslund, J. Beckwith, G. Georgiou, Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm, Proc. Natl. Acad. Sci. USA 96 (1999) 13703–13708.

[51] R.A. Berg, D.J. Prockop, Affinity column purification of protocollagen proline hydroxylase from chick embryos and further characterization of the enzyme, J. Biol. Chem. 218 (1973) 1175–1182.

[52] D.C.A. John, N.J. Bulleid, Prolyl 4-hydroxylase: defective assembly of the α-subunit mutants indicates that assembled α-subunits

are intramolecularly disulfide bonded, Biochemistry 33 (1994) 14018–14025.

[53] J.A. Bassuk, W.W.-Y. Kao, P. Herzer, N.L. Kedersha, J. Seyer, J.A. DeMartino, B.L. Daugherty, G.E. Mark III, R.A. Berg, Prolyl 4-hydroxylase: molecular cloning and the primary structure of the α subunit from chicken embryo, Proc. Natl. Acad. Sci. USA 86 (1989) 7382–7386.

[54] T. Helaakoski, P. Annunen, K. Vuori, I.A. Macneil, T. Pihlajaniemi, K.I. Kivirikko, Cloning, baculovirus expression, and characterization of a second mouse prolyl 4-hydroxylase α-subunit isoform: formation of an $\alpha_2\beta_2$ tetramer with the protein disulfide-isomerase/β subunit, Proc. Natl. Acad. Sci. USA 92 (1995) 4427–4431.

[55] I. Hopkinson, S.A. Smith, A. Donne, H. Gregory, T.J. Franklin, M.E. Grant, J. Rosamond, The complete cDNA derived sequence of the rat prolyl 4-hydroxylase α subunit, Gene 149 (1994) 391–392.

[56] J. Myllyharju, K.I. Kivirikko, Identification of a novel proline-rich peptide-binding domain in prolyl 4-hydroxylase, EMBO J. 18 (1999) 306–312.

[57] R. Hieta, L. Kukkola, P. Permi, P. Pirilä, K.I. Kivirikko, I. Kipeläinen, J. Myllyharju, The peptide-substrate-binding domain of human collagen prolyl 4-hydroxylases. Backbone assignments, secondary structure, and binding of proline-rich peptides, J. Biol. Chem. 278 (2003) 34966–34974.

[58] P.L. Atreya, V.S. Ananthanarayanan, Interaction of prolyl 4-hydroxylase with synthetic peptide substrates: a conformational model for collagen proline hydroxylation, J. Biol. Chem. 266 (1991) 2852–2858.

[59] M. Wu, H.-S. Moon, A. Pirskanen, J. Myllyharju, K.I. Kivirikko, T.P. Begley, Mechanistic studies on prolyl 4-hydroxylase: the vitamin C requiring uncoupled oxidation, Bioorg. Med. Chem. Lett. 10 (2000) 1511–1514.

[60] K. Juva, D.J. Prockop, Modified procedure for the assay of H3- or C14-labeled hydroxyproline, Anal. Biochem. 15 (1966) 77–83.

[61] J.A. Ramshaw, J.A. Werkmeister, V. Glattauer, Collagen-based biomaterials, Biotechnol. Genet. Eng. Rev. 13 (1996) 335–382.

[62] J.A. Werkmeister, J.A.M. Ramshaw (Eds.), Collagen Biomaterials, Elsevier Science, Barking, Essex, England, 1992.