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

Johanna VEIJOLA*, Pia ANNUNEN*, Peppi KOIVUNEN*, Antony P. PAGE†, Taina PIHLAJANIEMI* and Kari I. KIVIRIKKO*‡

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

†Wellcome Unit of Molecular Parasitology, The Anderson College, Glasgow University, Glasgow G11 6NU, U.K.

Protein disulphide isomerase (PDI; EC 5.3.4.1) is a multifunctional polypeptide that is identical to the β subunit of prolyl 4-hydroxylases. We report here on the cloning and expression of the *Caenorhabditis elegans* PDI/ β polypeptide and its isoform. The overall amino acid sequence identity and similarity between the processed human and *C. elegans* PDI/ β polypeptides are 61 % and 85 % respectively, and those between the *C. elegans* PDI/ β polypeptide and the PDI isoform 46 % and 73 %. The isoform differs from the PDI/ β and ERp60 polypeptides in that its N-terminal thioredoxin-like domain has an unusual catalytic site sequence -CVHC-. Expression studies in insect cells demonstrated that the *C. elegans* PDI/ β polypeptide forms an active prolyl 4-hydroxylase $\alpha_2\beta_2$ tetramer with the human α subunit and an $\alpha\beta$ dimer with the *C. elegans* α subunit, whereas the *C. elegans* PDI isoform formed no prolyl 4-hydroxylase with either

α subunit. Removal of the 32-residue C-terminal extension from the *C. elegans* α subunit totally eliminated $\alpha\beta$ dimer formation. The *C. elegans* PDI/ β polypeptide formed less prolyl 4-hydroxylase with both the human and *C. elegans* α subunits than did the human PDI/ β polypeptide, being particularly ineffective with the *C. elegans* α subunit. Experiments with hybrid polypeptides in which the C-terminal regions had been exchanged between the human and *C. elegans* PDI/ β polypeptides indicated that differences in the C-terminal region are one reason, but not the only one, for the differences in prolyl 4-hydroxylase formation between the human and *C. elegans* PDI/ β polypeptides. The catalytic properties of the *C. elegans* prolyl 4-hydroxylase $\alpha\beta$ dimer were very similar to those of the vertebrate type II prolyl 4-hydroxylase tetramer, including the K_m for the hydroxylation of long polypeptide substrates.

INTRODUCTION

Protein disulphide isomerase (PDI; EC 5.3.4.1) is an abundant protein in the lumen of the endoplasmic reticulum, with multiple functions. It catalyses thiol–disulphide interchange *in vitro*, leading to net protein disulphide bond formation, reduction or isomerization, depending on the reaction conditions, and is regarded as the catalyst *in vivo* of disulphide bond formation in the biosynthesis of various secretory and cell surface proteins (reviewed in [1–3]). PDI is likely to be a modular protein comprising domains *a*, *e*, *b*, *b'*, *a'* and *c* [2,4–6]. The *a* and *a'* domains are similar to thioredoxin, a well-characterized small protein that participates in many cytoplasmic redox reactions [2,4]. Both these domains contain a -Cys-Gly-His-Cys- sequence, representing two independently acting catalytic sites for the isomerase activity [7–11]. PDI also acts as the β subunit of the animal prolyl 4-hydroxylases [5,12,13] and the microsomal triacylglycerol transfer protein [14,15], and as a chaperone-like polypeptide that binds various peptides within the lumen of the endoplasmic reticulum [1,16–23]. PDI has also been reported to have several other less-well-characterized functions [1,2,24].

Prolyl 4-hydroxylase (EC 1.14.11.2) catalyses the formation of 4-hydroxyproline in collagens and other proteins with collagen-like sequences [24] by the hydroxylation of proline residues in -Xaa-Pro-Gly- sequences (reviewed in [6,25,26]). The vertebrate type I and type II enzymes [27] are $[\alpha(I)]_2\beta_2$ and $[\alpha(II)]_2\beta_2$ tetramers respectively. Molecular cloning and baculovirus expression of the α subunit from *Caenorhabditis elegans* indicated that this polypeptide forms an active prolyl 4-hydroxylase with

the human PDI/ β polypeptide; however, surprisingly, the resulting hybrid protein was found to be an $\alpha\beta$ dimer [28]. The overall amino acid sequence identities between the processed forms of the *C. elegans* α subunit and the human and mouse $\alpha(I)$ subunits are 45 % and 43 % respectively, whereas that between the *C. elegans* α subunit and the mouse $\alpha(II)$ subunit is 41 % [27–29]. A major difference between the *C. elegans* α subunit and the vertebrate $\alpha(I)$ and $\alpha(II)$ subunits is the presence of a 32-residue C-terminal extension only in the former, the function of which is unknown [27,28]. The catalytic properties of the *C. elegans*–human hybrid protein are very similar to those of the vertebrate prolyl 4-hydroxylases [27,28].

To study further the specificity of the prolyl 4-hydroxylase α subunit–PDI/ β polypeptide interaction we cloned the PDI/ β polypeptide from *C. elegans* and expressed it in insect cells together with either the *C. elegans* α subunit or the human $\alpha(I)$ subunit. This PDI/ β polypeptide was found to form an $\alpha\beta$ dimer with the *C. elegans* α subunit and an $\alpha_2\beta_2$ tetramer with the human $\alpha(I)$ subunit, but in the case of both α subunits the human PDI/ β polypeptide was much more effective for prolyl 4-hydroxylase formation than the *C. elegans* PDI/ β polypeptide. Experiments with hybrid human–*C. elegans* polypeptides demonstrated that this is partly due to differences in the sequences within the C-terminal regions of the two polypeptides. We also characterized the catalytic properties of the *C. elegans* prolyl 4-hydroxylase $\alpha\beta$ dimer and studied the possible role of the C-terminal extension present in the *C. elegans* α subunit in the dimer assembly. The cDNA for the *C. elegans* PDI/ β polypeptide was found to cross-hybridize with genomic clones for an ad-

ditional *C. elegans* PDI isoform. This isoform was also expressed in insect cells and found to be an active disulphide isomerase, but it did not form prolyl 4-hydroxylase with the *C. elegans* or human α subunits.

EXPERIMENTAL

Isolation of cDNA clones

cDNA clones for the *C. elegans* PDI/ β polypeptide were isolated by synthesizing a 430 bp PCR fragment from a nematode cDNA library in the λ Uni-ZAP XR-vector (Stratagene no. 937006) with the primers CEB2 (5'-CCACTCAACTCAAGGAAGAAGGAT-3') and CEB4 (5'-CGCCCATCAATCGAAACTTCTTGA-3') based on the known sequences of clone CEMSA22 (see the Results section). A second PCR fragment was prepared with a nested primer CEB3 (5'-CGGAAAGCTCGATGCCACCGTCCAC-3') and the CEB4 primer, with the 430 bp PCR product as a template. The reactions were cycled 25 times as follows: denaturation for 1 min at 94 °C, annealing for 1 min at 70 °C and extension for 2 min at 72 °C. The resulting fragment from the second round of PCR was labelled with 32 P by using a random-primed DNA labelling kit (Boehringer Mannheim). Screening of the nematode cDNA library under stringent conditions [30] resulted in seven positive λ Uni-ZAP clones. The chimaeric pBluescript clones were excised in accordance with the instructions supplied by the manufacturer (Stratagene).

The cDNA clone for the *C. elegans* PDI isoform polypeptide was identified during hybridization of the gene encoding the PDI/ β polypeptide to the *C. elegans* physical map. Two distinct chromosomal hybridization patterns were obtained with this clone on the *C. elegans* yeast artificial chromosome grid, by the methods of Coulston et al. [30]. The cDNA clone for the PDI isoform was isolated by reverse transcriptase-PCR from a mixed-stage cDNA synthesized [31] with the primers PDIF (5'-GGAAGGATTTCAATGTCAGTGTCC-3') and PDIR (5'-GCTCTAGATTAAAGCTCTTCGTGATCCTG-3'), designed in accordance with the known genomic sequences of the cosmid clone C14B1 [32]. The letters underlined in these and all subsequent primers (below) indicate non-cDNA-specific nucleotides used to generate artificial restriction enzyme recognition sites at the 5' ends of the primers. The reactions were cycled 20 times as follows: denaturation for 30 s at 94 °C, annealing for 1 min at 60 °C and extension for 1 min at 72 °C. The resultant fragment was subcloned into the vector pMalc2 (New England Biolabs) and sequenced in both strands.

Reverse transcription and cDNA amplification

Attempts were made to isolate the extreme 5' end of the *C. elegans* PDI/ β cDNA by 5' rapid amplification of cDNA ends (RACE). Total RNA was transcribed with oligo(dT) as a primer [31], the cDNA was dA-tailed [31], the 3' primer CB397 (5'-CTTGGCGACTGGTCCGGT-3') was used for the first PCR reaction, and a nested 3' primer, CB312 (5'-GGAAAGCTTGAGAGTTGGGT-3') for the second, whereas the 5' primer [5'-GACTCGAGTCGACTACGA(T)₁₇-3'] contained a (dT)₁₇ stretch. The amplified products were subcloned into the vector pCRII (Invitrogen) and analysed by nucleotide sequencing.

Construction of baculovirus transfer vectors and generation of recombinant baculoviruses

The *C. elegans* PDI/ β polypeptide was expressed by constructing a pVL-CE β baculovirus transfer vector in which the missing

extreme N-terminal end of the *C. elegans* PDI/ β polypeptide was replaced by the signal peptide and the first three amino acids of the processed human PDI/ β polypeptide [5], these being followed by the *C. elegans* PDI/ β polypeptide. The clone CEB1, containing nt 1–1388 of the *C. elegans* PDI/ β cDNA at the *EcoRI*–*XhoI* site, was digested with *EcoRI* and *HindIII* or *HindIII* and *SalI*. The resulting 245 bp *EcoRI*–*HindIII* fragment, covering nt 55–299, was ligated together with a 68 bp *XbaI*–*EcoRI* fragment, prepared from four oligonucleotides encoding the signal peptide and the first three amino acids (Asp-Ala-Pro) of the human PDI/ β polypeptide [5], to the *XbaI*–*HindIII* site of the pBluescript vector (Stratagene), and was termed BS-CEBXH. The clone CEB8, containing nt 501–1773 of the *C. elegans* PDI/ β cDNA at the *XhoI*–*EcoRI* site, was digested with *PstI* and *SalI*. The resulting 449 bp *SalI*–*PstI* fragment, covering nt 1344–1477 of the coding region and nt 1478–1773 of the 3' non-coding region, and extending 19 nt in the 3' direction towards the pBluescript vector (Stratagene), was ligated together with a 1044 bp *HindIII*–*SalI* fragment, covering nt 300–1343, to the *HindIII*–*PstI* site of the pBluescript vector. The construct was termed BS-CEBHP. Finally, the 313 bp *XbaI*–*HindIII* fragment from BS-CEBXH and the 1493 bp *HindIII*–*PstI* fragment from BS-CEBHP were ligated to the *XbaI*–*PstI* site of the transfer vector pVL1393, and the construct was termed pVL-CE β .

The hybrid *C. elegans*–human PDI/ β polypeptide consisted of the signal peptide and amino acids of the processed *C. elegans* PDI/ β polypeptide as above as far as residue 435 of the processed polypeptide, followed by the last 52 amino acids of the human PDI/ β polypeptide [5]. This polypeptide was expressed by digesting the pVL-CE β with *XbaI* and *SalI* and ligating the resulting 1357 bp *XbaI*–*SalI* fragment together with a *SalI* and *BamHI*-digested 363 bp PCR fragment covering nt 1413–1779 of the human PDI/ β sequence to the *XbaI*–*BamHI* site of the transfer vector pVL1392. This PCR fragment was synthesized with the primers PK5 (5'-GGTGTGCGACTACAACGGGG-AACGCACG-3') and PK2 (5'-GATGGATCCCTTTCCAA-AACCG-3'), hybridizing to nt 1413–1430 and 1757–1775 of the human PDI/ β sequence [5] respectively. The final construct was termed pVL-CE β /H β .

The hybrid human–*C. elegans* PDI/ β polypeptide consisted of the signal peptide and the first 377 amino acids of the processed human PDI/ β polypeptide [5] followed by the last 103 amino acids of the *C. elegans* PDI/ β polypeptide. The pVL- β vector for the human PDI/ β polypeptide [33] was digested with *EcoRI* and *NcoI*, and the resulting 1227 bp *EcoRI*–*NcoI* fragment was ligated together with an *NcoI* and *SmaI*-digested 346 bp PCR fragment covering nt 1167–1505 of the *C. elegans* PDI/ β sequence to the *EcoRI*–*SmaI* site of the transfer vector pVL1392. The PCR fragment was synthesized with the primers PK6 (5'-GCCCCATGGTGCGGACACTGCAAG-3') and PK7 (5'-CTGCCCCGGGATGCGTTGAACTGTATTGG-3'), hybridizing to nt 1163–1186 and 1487–1505 of the *C. elegans* PDI/ β sequence respectively. The final construct was termed pVL-H β /CE β .

To build a pVL-dCE α transfer vector coding for a truncated form of the α subunit of *C. elegans* prolyl 4-hydroxylase, an artificial translation termination codon was added to positions 1665–1667 of the *C. elegans* α subunit cDNA, corresponding to the location of the translation stop codon in the human [29] and chicken [34] sequences. The clone pBS-CEL1, covering the full-length cDNA for the α subunit of *C. elegans* prolyl 4-hydroxylase [28], was digested with *EcoRI* and *BamHI*, and the resulting 1609 bp *EcoRI*–*BamHI* fragment was ligated together with a *BamHI* and *SmaI*-digested 58 bp PCR fragment covering nt 1610–1667 of the *C. elegans* α subunit sequence to the

EcoRI–*SmaI* site of the transfer vector pVL1392. The PCR fragment coding for residues 490–511 of the *C. elegans* α subunit was synthesized with the primers d5'C α (5'-CCAACAAATG-GATCCACGAGAAGG-3') and d3'C α (5'-TTACCCGGGCT-AGTCTGAGGATTT-3'), hybridizing to nt 1600–1623 and 1653–1666 in the *C. elegans* α subunit sequence [28] respectively.

A construct coding for the *C. elegans* PDI isoform polypeptide was obtained by digesting the plasmid PMALc2-PDI isoform with *XmnI* and *XbaI*. The resulting 1464 bp fragment was ligated to the *SmaI*–*XbaI* site of the transfer vector pVL1393, and termed pVL-CEPDI-I.

All the sequences of the baculovirus transfer vector constructs were verified by nucleotide sequencing.

The recombinant baculovirus transfer vectors were co-transfected into *Spodoptera frugiperda* Sf9 cells (Invitrogen) with a modified *Autographa californica* nuclear polyhedrosis virus DNA, BaculoGold (PharMingen), by calcium phosphate transfection, and the recombinant viruses were selected [35]. These were designated as CE β , CE β /H β , H β /CE β , CEPDI-I and dCE α . The other recombinant baculoviruses used in this work, CE α , α 59 and β , code for the *C. elegans* and human type I prolyl 4-hydroxylase α subunits and the human PDI/ β polypeptide respectively [28,33].

Nucleotide sequencing and sequence analysis

The nucleotide sequences were obtained by the Sanger dideoxy-nucleotide chain termination method [36] with T7 DNA polymerase (Pharmacia). Vector-specific or sequence-specific 17-mer primers synthesized in an Applied Biosystems DNA synthesizer (Department of Biochemistry, University of Oulu) were used. The sequence data were compiled with DNASIS and PROSIS version 7.0 sequence analysis software (Pharmacia), and the similarity comparisons were performed against the GenBank, EMBL, PIR and SwissProt databases at NCBI (National Center for Biotechnology Information) using the BLAST network service [37].

Immunizations

The rabbits were injected subcutaneously at several sites with 60–100 μ g of recombinant *C. elegans* α subunit polypeptide purified by detergent extraction and preparative SDS/PAGE, revealed with copper staining and processed for immunization [38]. Immunoglobulins were isolated and purified by affinity chromatography on a Protein A–Sepharose column (Pharmacia). Further purification was obtained by means of a negative-affinity column containing a crude mixture of proteins from insect cells infected with a baculovirus coding for the human α subunit [33].

Analysis of recombinant proteins in insect cells

The Sf9 insect cells (Invitrogen) were cultured in TNM-FH medium (Sigma) supplemented with 10% fetal bovine serum (BioClear) at 27 °C, either as monolayers or in suspension in spinner flasks (Techne). To produce recombinant proteins, insect cells seeded at a density of 10⁶/ml were infected at a multiplicity of 5 with any of the recombinant viruses nematode CE α , CE β or dCE α , human α 59 or β or hybrids CE β /H β and H β /CE β . The same multiplicity of infection was also used when two viruses were co-expressed. The cells were harvested 3 days after infection, washed with a solution of 0.15 M NaCl and 0.02 M phosphate, pH 7.4, homogenized in a buffer of 0.1 M glycine, 0.1 M NaCl, 10 μ M dithiothreitol, 0.1% Triton X-100 and 0.01 M Tris, pH 7.8, and centrifuged at 10000 *g* for 20 min at 4 °C. Aliquots of the resulting supernatants were analysed by denaturing

SDS/PAGE (8% gel) under reducing conditions or non-denaturing PAGE (7.5% gel) and assayed for enzyme activity. The remaining cell pellets were washed with the homogenizing buffer, solubilized in 1% (w/v) SDS and analysed by denaturing SDS/PAGE (8% gel) under reducing conditions.

Protein purification

Attempts were made to purify the prolyl 4-hydroxylases resulting from co-infection with the viruses nematode CE α and CE β by two separate procedures. The first was based on the use of affinity chromatography [38] on a column containing rabbit antibodies to the α subunit of *C. elegans* prolyl 4-hydroxylase coupled to Sepharose 4B (Pharmacia), and the second consisted of anion-exchange chromatography on a DEAE-cellulose column followed by gel filtration [28].

Enzyme activity assays

Prolyl 4-hydroxylase activity was assayed by a method based either on the hydroxylation-coupled decarboxylation of 2-oxo[1-¹⁴C]glutarate [39] or on the formation of hydroxy[¹⁴C]proline in a biologically prepared [¹⁴C]proline-labelled protein substrate consisting of non-hydroxylated pro- α chains of chick type I procollagen [39]. *K_m* values were determined by varying the concentration of one substrate in the presence of fixed concentrations of the second while the concentrations of the other substrates were kept constant [40]. As the *C. elegans* enzyme was not stable during purification, the Triton X-100-soluble fraction of the insect cell homogenate was used as an enzyme source for the *K_m* determinations.

PDI activity was assayed by a method based on catalysis of the oxidative refolding of reduced, denatured RNAase [41]. Before the PDI activity was measured, the polypeptides expressed were partly purified by gel filtration in an S-200 HR column (Pharmacia), equilibrated and eluted with a solution of 0.3 M NaCl and 0.1 M Tris/acetate, pH 7.8. The amounts of recombinant polypeptide obtained were compared by densitometry of the Coomassie Brilliant Blue-stained bands in SDS/PAGE using a bioimage technique (BioImage, Millipore).

Other assays

Non-denaturing PAGE (7.5% gel) analysis was performed with a running buffer of 0.192 M glycine and 25 mM Tris, pH 8.3, at 15 V for 20 h at 4 °C. The proteins for Western blot analysis were transferred to a poly(vinylidene fluoride) membrane (Immobilon P) at 80 V for 50 min in a cooled chamber (4 °C), and the analysis was performed with polyclonal antibodies K30 to the α subunit of *C. elegans* prolyl 4-hydroxylase, K60 to the human PDI/ β polypeptide or K17 to the α (I) subunit of human prolyl 4-hydroxylase [42].

Protein concentrations were determined with a Bio-Rad protein assay kit (Bio-Rad).

RESULTS

Isolation of cDNA clones

A study of the identification of genes expressed in *C. elegans* led to sequencing of the 5' end of 1517 clones in a cDNA library of 1759 clones [43]. The 5' sequence of clone CEMSA22 showed a significant similarity to the nucleotide sequence [5] of the human PDI/ β polypeptide. cDNA clones encoding the entire *C. elegans* PDI/ β polypeptide were isolated by synthesizing two PCR primers and using them to produce a 430 bp PCR fragment from a *C. elegans* cDNA library. A further PCR

H-PDI	-17	MYRRATTCGLVAALVRADAPPEEPHVLVLRKSNF
C-PDI	-14	--RLVCLFPLVLGASAAV---EEENVVLTKDNF
C-PDI-I	-20	MSLSVFFFLVLSAGVAV---VADENVVLTKDNF
H-PDI	18	AEATAHKKTLVFEYAPWCGCHCKSLAEYAKAAQKLR
C-PDI	18	DEVINQNEFFLVFEYAPWCGCHCKSLAEYAKAAQKLR
C-PDI-I	14	EEVINQNEFFLVFEYAPWCVHCKSLAEYAKAAQKLR
H-PDI	55	AEGSSEHLAKVDATEESDLAQQGVGRGYPTIKFERNG
C-PDI	55	EEGSSEHLAKVDATEESDLAQQGVGRGYPTIKFERNG
C-PDI-I	51	EEGSSEHLAKVDATEESDLAQQGVGRGYPTIKFERNG
H-PDI	92	DTASFEYTHCRFADIVVWLKKRTGPAATLPDGAAR
C-PDI	92	K---PQYNGGRDHDSTAWLKKRTGPAATLPDGAAR
C-PDI-I	88	K---PQYNGGRDHDSTAWLKKRTGPAATLPDGAAR
H-PDI	129	AEGLVSESSVAVIGCFPKDVSDFAKQELQAAEAIDDF
C-PDI	126	VKELQESADVIVIGCFPKDVSDFAKQELQAAEAIDDF
C-PDI-I	122	VEELKGRKRVVVVGYGPKDAKSDAATINNEVADSVDDA
H-PDI	166	EEGITSNSDVPFVKYQLDKGVVLEKKEDDGR---NN
C-PDI	163	EEGITSNSDVPFVKYQLDKGVVLEKKEDDGR---NN
C-PDI-I	159	EEGITSNSDVPFVKYQLDKGVVLEKKEDDGR---NN
H-PDI	199	EEGITSNSDVPFVKYQLDKGVVLEKKEDDGR---NN
C-PDI	196	EEGITSNSDVPFVKYQLDKGVVLEKKEDDGR---NN
C-PDI-I	196	EEGITSNSDVPFVKYQLDKGVVLEKKEDDGR---NN
H-PDI	236	IKSHNLLFVSKESDEPAKLEQFPKNAKQPKGKILFV
C-PDI	233	IKSHNLLFVSKESDEPAKLEQFPKNAKQPKGKILFV
C-PDI-I	233	IKSHNLLFVSKESDEPAKLEQFPKNAKQPKGKILFV
H-PDI	273	FIDSDHEDHQLRIEFLGKLEKCEPAVRLILEEPMFK
C-PDI	270	YINDDVEENARILEEFLGKLEKCEPAVRLILEEPMFK
C-PDI-I	270	YINDDVEENARILEEFLGKLEKCEPAVRLILEEPMFK
H-PDI	310	VKPEPDEETENISKFTONVLDGSKVHLMSSDPEP
C-PDI	307	VKPEPDEETENISKFTONVLDGSKVHLMSSDPEP
C-PDI-I	307	VKPEPDEETENISKFTONVLDGSKVHLMSSDPEP
H-PDI	347	WDKNPVKVLVCKNEEVARDEKKNVEVEFYAPWCGHC
C-PDI	344	WDKNPVKVLVCKNEEVARDEKKNVEVEFYAPWCGHC
C-PDI-I	340	WDKNPVKVLVCKNEEVARDEKKNVEVEFYAPWCGHC
H-PDI	384	KOLAPWDKLCGEVYKDHENIVIAKMDSTANEVAVKV
C-PDI	381	KOLAPWDKLCGEVYKDHENIVIAKMDSTANEVAVKV
C-PDI-I	377	KOLAPWDKLCGEVYKDHENIVIAKMDSTANEVAVKV
H-PDI	421	HSFPTLKFFPAGSNNKVVVDYQDRTEGCFKFLSEGC
C-PDI	418	HSFPTLKFFPAGSNNKVVVDYQDRTEGCFKFLSEGC
C-PDI-I	414	HSFPTLKFFPAGSNNKVVVDYQDRTEGCFKFLSEGC
H-PDI	458	QDGAGDDEDLDEEAEFPDMEEDDDQKAVKDEL*
C-PDI	454	QDGAGDDEDLDEEAEFPDMEEDDDQKAVKDEL*
C-PDI-I	450	QDGAGDDEDLDEEAEFPDMEEDDDQKAVKDEL*

Figure 1 Alignment of the amino acid residues of the human and *C. elegans* PDI/ β polypeptides and the *C. elegans* PDI isoform

The first row shows the human PDI/ β polypeptide; the second row the *C. elegans* PDI/ β polypeptide; and the third row the *C. elegans* PDI isoform. The amino acid sequence is shown in one-letter codes. Gaps (-) are introduced for maximal alignment of the polypeptides; * represents the stop codon of translation. Numbering of the amino acids begins with the first residue in the processed polypeptide. Negative numbers indicate signal peptides. Identity between amino acids is indicated by black boxes and similarity by grey boxes. Similar amino acids: G = A = S; A = V; V = I = L = M; I = L = M = F = Y = W; K = R = H, D = E = Q = N; S = T = E = N.

fragment was prepared with a nested 5' primer and the original 3' primer, with the previously synthesized PCR fragment as a template. The resulting 393 bp PCR fragment was used to screen the same *C. elegans* cDNA library, and seven positive clones were identified from about 500 000 recombinants. Clones CEB1 and CEB7 contained overlapping sequences that covered most of the signal peptide, the whole processed polypeptide and 295 bp of the 3' untranslated region, but none of the clones contained sequences coding for the N-terminal end of the signal peptide. Attempts to obtain the extreme 5' sequences by 5' RACE-PCR with isolated *C. elegans* total RNA as a template failed, and therefore the baculovirus expression construct for the *C. elegans* PDI/ β polypeptide (below) was built to contain sequences coding for the whole signal peptide and the first three amino acids of the processed polypeptide from human PDI/ β .

The full-length insert from the baculovirus transfer vector encoding the above *C. elegans* PDI/ β polypeptide was found to hybridize with an overlap of four yeast artificial chromosome clones (Y56D11, Y52C2, Y50C8 and Y51F11) from the left arm of chromosome X. Nevertheless an additional hybridization

signal was found, pointing to an overlap of two yeast artificial chromosome clones (Y50C6 and Y6C11) on the left arm of chromosome III. Two PCR primers were synthesized based on the known sequences of the 5' and 3' ends of the coding region of the respective gene [32], and these were used to prepare a PCR fragment covering the whole coding region of the respective gene by reverse transcriptase-PCR with *C. elegans* RNA as a template. The sequence of the PCR fragment was verified by nucleotide sequencing and comparison with those of the exons in the corresponding gene. The polypeptide encoded by this gene was termed here the PDI isoform.

cDNA-derived amino acid sequences of the *C. elegans* PDI/ β and PDI isoform polypeptides

The cDNA clones isolated here for the *C. elegans* PDI/ β polypeptide cover the whole coding region of the processed polypeptide and most of the signal peptide (Figure 1). In terms of the parameters of von Hejne [44], the most likely first amino acid of the processed polypeptide is the latter of the two consecutive alanines, and thus the size of the processed *C. elegans* PDI/ β polypeptide must be 477 residues. This polypeptide is shorter than the processed human [5] and chick [13] PDI/ β polypeptides, the sizes of which are 491 and 493 residues, respectively. The C-terminus of the *C. elegans* polypeptide has the sequence HTEL, which is a functional variant of the KDEL retention signal [45] present in the human and chick PDI/ β polypeptides [5,13].

The overall amino acid sequence identity and similarity between the processed *C. elegans* and human PDI/ β polypeptides are 61 % and 85 % respectively (Figure 1). The highest degrees of identity, 78 % and 61 %, are found within the two thioredoxin-like domains *a'* (residues 346–428, *C. elegans* numbering) and *a* (residues 9–90). The sequences around the two -CGHC- catalytic sites in these two domains are especially well conserved, a 23-residue sequence in domain *a* (residues 28–50) and a 26-residue sequence in domain *a'* (residues 366–391) being 96 % and 92 % identical between the *C. elegans* and human polypeptides (Figure 1). The lowest degree of amino acid sequence identity is found within the C-terminal region of the C-terminal domain *c*, a 27-residue sequence (residues 451–477) being only 33 % identical between the two polypeptides.

The cleavage site for the signal peptide of the *C. elegans* PDI isoform studied here is likely to be between alanine and aspartate [44], and thus the size of the signal peptide would be 20 residues and that of the processed polypeptide 465 residues (Figure 1). The C-terminus of this polypeptide has the sequence HEEL, which is also a functional variant of the KDEL retention signal [45]. The overall amino acid sequence identity and similarity between the *C. elegans* PDI isoform and the PDI/ β polypeptide are 46 % and 73 % respectively, and between the PDI isoform and the human PDI/ β polypeptide [5] 44 % and 70 %. These values are markedly higher than the overall amino acid sequence identities between the *C. elegans* PDI isoform and the human and *Drosophila* ERp60 polypeptides [46], which are 30 % and 55 %, thus indicating that the *C. elegans* polypeptide is more closely related to the PDI/ β polypeptides than to the polypeptides [2] known as ERp60. The calculated pI of the *C. elegans* PDI isoform is 4.67, which is closer to the pI values of 4.69 and 4.69 calculated for the *C. elegans* and human PDI/ β polypeptides than those of 5.61 and 5.53 for the human and *Drosophila* ERp60 polypeptides [46].

The highest degrees of amino acid sequence identity between the PDI isoform and the PDI/ β polypeptides are found within the two thioredoxin-like domains *a* and *a'*, in which the residues

Figure 2 Analysis of the expression of various normal and modified *C. elegans* polypeptides in insect cells by SDS/PAGE under reducing conditions

Lanes 1, 3, 5, 7, 9, 11 and 13 show 0.1% Triton X-100-soluble proteins from cells infected with viruses coding for the human PDI/ β polypeptide, *C. elegans* PDI/ β polypeptide, *C. elegans* PDI isoform, hybrid *C. elegans*–human PDI/ β polypeptide, hybrid human–*C. elegans* PDI/ β polypeptide, *C. elegans* prolyl 4-hydroxylase α subunit and a truncated form of this subunit respectively. After treatment with the buffer containing Triton X-100, the remaining cell pellets were solubilized with 1% SDS, and the protein samples obtained were run in lanes 2, 4, 6, 8, 10, 12, and 14 in the same order. Lanes 1–4, 5–10 and 11–14 were run on three separate gels. The gels were analysed by staining with Coomassie Brilliant Blue.

of the PDI isoform are 79% and 69% identical respectively with those present either in the human or *C. elegans* PDI/ β polypeptide sequences (Figure 1). Surprisingly, however, the sequence of the PDI isoform corresponding to the PDI catalytic site sequence in domain *a* is -CVHC- instead of the usual -CGHC- (Figure 1). The -CVHC- sequence can also be translated from the corresponding genomic DNA [32], and is thus not a PCR artifact. The lowest degree of amino acid sequence identity is again seen within the C-terminal region of domain *c*, the last 24 amino acids being only 17% identical with those present in either the human or *C. elegans* PDI/ β polypeptides (Figure 1).

Expression of the *C. elegans* PDI/ β polypeptide and PDI isoform in insect cells

Insect cells were infected with recombinant baculoviruses coding for either the PDI/ β polypeptide or the PDI isoform and harvested after 72 h, homogenized in buffer containing 0.1% Triton X-100 and centrifuged. The cell pellets were further solubilized in 1% SDS, and the proteins soluble in 0.1% Triton X-100 and 1% SDS were analysed by SDS/PAGE under reducing conditions. Both polypeptides were found in the Triton X-100 extracts (Figure 2, lanes 3 and 5), there being little, if any, of either of them in the 1% SDS extracts (Figure 2, lanes 4 and 6).

To determine the PDI activity of the polypeptides expressed, 0.1% Triton X-100 extracts of cell homogenates were partly purified by gel filtration, which removes their RNase activity. The PDI activity of the partly purified extracts was then determined with an assay based on catalysis of the renaturation of reduced, denatured RNase [41], and the amounts of polypeptide were estimated by densitometry of the respective Coomassie-stained bands in SDS/PAGE. Both the *C. elegans* PDI/ β polypeptide and the PDI isoform were found to be approximately equally active disulphide isomerases as the human PDI/ β polypeptide (results not shown).

Expression of the *C. elegans* PDI/ β polypeptide or PDI isoform together with the α subunit of *C. elegans* or human prolyl 4-hydroxylase in insect cells

To study whether the *C. elegans* PDI/ β polypeptide and PDI isoform associate with the α subunits of *C. elegans* and human prolyl 4-hydroxylases, insect cells were co-infected with two

Figure 3 Analysis of prolyl 4-hydroxylase formation from *C. elegans* or human prolyl 4-hydroxylase α and PDI/ β subunits expressed in insect cells by means of baculovirus vectors by PAGE under non-denaturing conditions

The samples were extracted with a buffer containing 0.1% Triton X-100 and subjected to PAGE with 7.5% gel. Lanes 1–4, extracts from cells co-infected with viruses coding for either the human (lanes 1 and 4) or *C. elegans* (lanes 2 and 3) α subunit and either the human (lanes 1 and 2) or *C. elegans* PDI/ β polypeptide (lanes 3 and 4). Lanes 5 and 6, extracts from cells infected with a virus coding for either the human or *C. elegans* PDI/ β polypeptide respectively, together with a virus coding for the truncated form of the α subunit. The samples were analysed by Coomassie Blue staining in (A) and by Western blotting with an anti-*C. elegans* α subunit antibody K30 in (B), an anti-human α subunit antibody K17 in (C) and an anti-human PDI/ β polypeptide antibody K38 in (D). The arrows indicate the enzymes formed from the *C. elegans* (C) or human (H) α and PDI/ β subunits, the position of the non-associated PDI/ β polypeptide also being indicated. The faint additional band seen in lanes 1 and 5 in (D) represents a PDI/ β polypeptide dimer (arrowhead).

recombinant baculoviruses, each coding for one polypeptide. The 0.1% Triton X-100 extracts of cell homogenates were then studied by PAGE performed under non-denaturing conditions (Figure 3). As reported previously [28,33] the human PDI/ β polypeptide formed an $\alpha_2\beta_2$ tetramer with the human α subunit (Figure 3, lane 1) and an $\alpha\beta$ dimer with the *C. elegans* α subunit (Figure 3, lane 2), both these prolyl 4-hydroxylases being readily seen in Coomassie Blue staining (Figure 3A) and Western blotting with an antibody to the human PDI/ β polypeptide (Figure 3D). The human enzyme tetramer could also be seen in Western blotting with an antibody to the human α subunit (Figure 3C, lane 1), and the hybrid *C. elegans*/human enzyme dimer in Western blotting with an antibody to the *C. elegans* α subunit (Figure 3B, lane 2). The *C. elegans* PDI/ β polypeptide also formed a prolyl 4-hydroxylase dimer with the *C. elegans* α subunit (Figure 3B, lane 3) and a tetramer with the human α subunit (Figure 3C, lane 4), but in both these α subunits much larger quantities of the dimer or tetramer were seen in experiments with human than with *C. elegans* PDI/ β polypeptide. The

Table 1 Prolyl 4-hydroxylase activity of Triton X-100 extracts of cells expressing human or *C. elegans* α subunits together with the human PDI/ β , *C. elegans* PDI/ β , *C. elegans*–human PDI/ β or human–*C. elegans* PDI/ β subunit

Values are given in d.p.m. per 50 μ l of extractable cell protein, means $\pm$ S.D. for at least three experiments, and as percentages of the values obtained with the wild-type human viruses or wild-type *C. elegans* α and human PDI/ β viruses. No differences were found in the amounts of the various polypeptides expressed per μ l of cell extract when studied by SDS/PAGE.

No.	Polypeptides expressed	Prolyl 4-hydroxylase activity		Significance
		(d.p.m. per 50 μ l)	(%)	
1	Human α + human PDI/ β	17 650 $\pm$ 840	100	
2	Human α + <i>C. elegans</i> PDI/ β	5 800 $\pm$ 1400	33	2 versus 1, $P < 0.001$
3	Human α + <i>C. elegans</i> –human PDI/ β	12 500 $\pm$ 1950	71	3 versus 2, $P < 0.005$
4	Human α + human– <i>C. elegans</i> PDI/ β	11 500 $\pm$ 2500	65	4 versus 1, $P < 0.005$
5	<i>C. elegans</i> α + human PDI/ β	14 600 $\pm$ 820	100	
6	<i>C. elegans</i> α + <i>C. elegans</i> PDI/ β	240 $\pm$ 100	2	6 versus 5, $P < 0.001$
7	<i>C. elegans</i> α + <i>C. elegans</i> –human PDI/ β	1 020 $\pm$ 690	7	7 versus 6, $P < 0.02$
8	<i>C. elegans</i> α + human– <i>C. elegans</i> PDI/ β	400 $\pm$ 90	3	8 versus 5, $P < 0.001$

amounts of prolyl 4-hydroxylase present in the experiments with the latter were so small that the dimer or tetramer formed could not be seen by Coomassie Blue staining (Figure 3A, lanes 3 and 4). This difference between the human and *C. elegans* PDI/ β polypeptides was consistently seen in all the experiments (more than 30).

The *C. elegans* PDI isoform was found to differ distinctly from the PDI/ β polypeptide, in that no prolyl 4-hydroxylase tetramer or dimer could be detected in any co-expression experiment involving it and the human or *C. elegans* α subunit (results not shown).

To study the possible role of the 32-amino-acid C-terminal extension present only in the *C. elegans* α subunit in prolyl 4-hydroxylase dimer formation, a baculovirus was generated that coded for a truncated form of this α subunit. The cDNA coding for the truncated polypeptide had an artificial translation stop codon in a position corresponding to the location of this codon in the human [29] and chick [34] α subunits. When the truncated α subunit was expressed in insect cells alone, almost all of it was found in the fraction of the cell homogenate soluble in 1% SDS but insoluble in 0.1% Triton (Figure 2, lanes 13 and 14), as with the wild-type *C. elegans* (Figure 2, lanes 11 and 12) and human [28, 33] α subunits. When the truncated α subunit was co-expressed with the human (Figure 3, lane 5) or *C. elegans* PDI/ β polypeptide (Figure 3, lane 6), no prolyl 4-hydroxylase tetramer or dimer was detected. The very weak band seen in Western blotting with an antibody to the human PDI/ β polypeptide (Figure 3D, lane 5) probably represents a PDI/ β dimer [28] rather than an $\alpha\beta$ dimer, as this band could not be stained with antibodies to the *C. elegans* α subunit (Figure 3B) even when tested in overloaded samples and with less dilute antibodies (results not shown), and as a similar faint PDI/ β dimer band is also seen in Figure 3(D), lane 1.

Triton X-100 extracts of cell homogenates expressing the various polypeptides were analysed for prolyl 4-hydroxylase activity with an assay based on the hydroxylation-coupled

Table 2 K_m values for co-substrates and the peptide substrate and K_i values for the poly(L-proline) inhibitor of *C. elegans*, human type I and human/mouse type II prolyl 4-hydroxylases

Abbreviation: n.i., no inhibition at 90 μ M concentration.

Co-substrate, substrate or inhibitor	Constant	K_m or K_i (μ M)		
		<i>C. elegans</i> enzyme	Human type I enzyme*	Human/mouse type II enzyme*
Fe ²⁺	K_m	4	4	4
2-Oxoglutarate	K_m	20	22	12
Ascorbate	K_m	400	330	340
(Pro-Pro-Gly) ₁₀	K_m	20	18	45
Poly(L-proline) (7000 Da)	K_i	n.i.	0.5	300†
Poly(L-proline) (44 000 Da)	K_i	n.i.	0.02	30†

* Values reported previously [27].

† Values are given as IC_{50} .

decarboxylation of 2-oxo[1-¹⁴C]glutarate [39]. As reported previously, the level of prolyl 4-hydroxylase activity in extracts from cells expressing the hybrid *C. elegans* α subunit/human PDI $\alpha\beta$ dimer enzyme was consistently very similar to that found in extracts from cells expressing the human $\alpha_2\beta_2$ tetramer (Table 1). The level in extracts from cells expressing the human α subunit/*C. elegans* PDI $\alpha_2\beta_2$ tetramer was usually approx. 25–35% of that in extracts from cells expressing the human $\alpha_2\beta_2$ tetramer, whereas the level in extracts from cells expressing the *C. elegans* α subunit/*C. elegans* PDI $\alpha\beta$ dimer was consistently very low, approx. 1–5% when reported as above (Table 1). The prolyl 4-hydroxylase activity in extracts from cells co-expressing the *C. elegans* PDI isoform with either the human or *C. elegans* α subunit, or the truncated *C. elegans* α subunit with either the human or *C. elegans* PDI/ β polypeptide, never exceeded the values of the negative controls (results not shown).

Catalytic properties of the recombinant *C. elegans* prolyl 4-hydroxylase

Attempts to purify the *C. elegans* prolyl 4-hydroxylase dimer by two procedures (see the Experimental section) failed, as this enzyme was found to be much less stable than the human enzyme tetramer or the *C. elegans* α /human PDI $\alpha\beta$ dimer. The *C. elegans* enzyme became dissociated into its subunits and degraded during both procedures. The catalytic properties of the *C. elegans* enzyme were therefore studied with the Triton X-100 extract from cell homogenates as an enzyme source.

The pH optimum of the *C. elegans* enzyme was found to be 7.1, which is slightly lower than the 7.4 reported previously for the chick enzyme [39] or the 7.5 determined here for the recombinant human enzyme (results not shown).

The K_m values for Fe²⁺, 2-oxoglutarate, ascorbate and the synthetic peptide substrate (Pro-Pro-Gly)₁₀ were very similar to those for the vertebrate type I and type II prolyl 4-hydroxylases (Table 2). The *C. elegans* enzyme was found to resemble the vertebrate type II prolyl 4-hydroxylase [27] and the previously studied *C. elegans* α /human PDI $\alpha\beta$ dimer [28], in that no inhibition was obtained with 7000 or 44000 Da poly(L-proline), even at a concentration of 90 μ M (Table 2).

The vertebrate prolyl 4-hydroxylase $\alpha_2\beta_2$ tetramers are thought

Figure 4 Analysis of prolyl 4-hydroxylase formation from human or *C. elegans* α subunits and hybrid human-*C. elegans* or *C. elegans*-human PDI/ β polypeptides in insect cells by PAGE under non-denaturing conditions

The samples were extracted with a buffer containing 0.1% Triton X-100 and subjected to PAGE with 7.5% gel. Lanes 1–4 in (A) show extracts from cells co-infected with viruses coding for the human α subunit and either the human (lane 1), *C. elegans* (lane 2), human-*C. elegans* (lane 3) or *C. elegans*-human (lane 4) PDI/ β polypeptide; lanes 5 and 6 show extracts from cells co-infected with viruses coding for the *C. elegans* α subunit and *C. elegans* (lane 5) or human (lane 6) PDI/ β polypeptide. Lanes 1–4 in (B) show extracts from cells co-infected with viruses coding for the human α subunit and the human (lane 1), *C. elegans* (lane 2), *C. elegans*-human (lane 3) or human-*C. elegans* (lane 4) PDI/ β polypeptides. Lanes 1–4 in (C) show extracts from cells co-infected with viruses coding for the *C. elegans* α subunit and the various PDI/ β polypeptides as in (B). The samples were analysed by Coomassie Blue staining in (A) and by Western blotting with an antibody against the human or *C. elegans* α subunit in (B) and (C) respectively. The arrows indicate the enzyme tetramers (T) or dimers (D) formed from the human or *C. elegans* α subunit and the various PDI/ β polypeptides, the non-associated PDI/ β polypeptide also being indicated in (A) (β).

to hydroxylate long polypeptide substrates by a mechanism in which the two catalytic sites act processively and bind to the same substrate [47]. This mechanism predicts that K_m values for the $\alpha\beta$ dimer enzymes of long polypeptide substrates should be markedly higher than those of the $\alpha_2\beta_2$ tetramer enzymes [47]. We therefore also studied the hydroxylation of [14 C]proline-labelled procollagen, a protein consisting of non-hydroxylated pro- α chains of chicken type I procollagen, with Triton X-100 extract from cell homogenates as a source of the enzyme. As the amount of prolyl 4-hydroxylase was much smaller in extracts from cells expressing the *C. elegans* prolyl 4-hydroxylase dimer than the human enzyme tetramer, the apparent maximal velocity per μ l of extract was much lower for the *C. elegans* enzyme. Nevertheless the apparent K_m values for the two enzymes were essentially identical, the data clearly excluding any marked differences between them (results not shown).

Expression of hybrid *C. elegans*-human or human-*C. elegans* PDI/ β polypeptides with the human or *C. elegans* prolyl 4-hydroxylase α subunits in insect cells

As the lowest degree of amino acid sequence identity between the *C. elegans* and human PDI/ β polypeptides was found within their C-terminal regions (see above), the possibility was studied that the differences in prolyl 4-hydroxylase formation between these two polypeptides might be largely due to differences in these sequences. Two hybrid constructs were therefore prepared.

One of them (*C. elegans*-human) codes for the above *C. elegans* PDI/ β polypeptide (where the first three amino acids are those present in the human polypeptide) up to residue 435 of the processed polypeptide (Figure 1), this being followed by the last 52 amino acids of the human polypeptide. The other (human-*C. elegans*) codes for the human PDI/ β polypeptide up to residue 377 of the processed polypeptide (Figure 1), this being followed by the last 103 amino acids of the *C. elegans* polypeptide.

When either of these hybrid polypeptides was expressed in insect cells alone, the recombinant polypeptide was found in the Triton X-100 extract of the cell homogenate (Figure 2, lanes 7 and 9). When the *C. elegans*-human PDI/ β polypeptide was co-expressed with the α subunit of human prolyl 4-hydroxylase (Figures 4A, lane 4, and 4B, lane 3), the amount of enzyme tetramer seen in PAGE performed under non-denaturing conditions was distinctly greater than that seen when the unmodified *C. elegans* PDI/ β polypeptide was expressed with the same α subunit (Figure 4A, lane 2). Also, when the human-*C. elegans* PDI/ β polypeptide was expressed with the human α subunit (Figures 4A, lane 3, and 4B, lane 4), the amount of enzyme tetramer was less than for the unmodified human PDI/ β polypeptide and the same α subunit (Figures 4A and 4B, lanes 1). Similar results were obtained with respect to prolyl 4-hydroxylase dimer formation when the hybrid polypeptides were co-expressed with the *C. elegans* α subunit (Figure 4C), except that the magnitudes of all the differences were even larger than for the human α subunit (Figures 4A and 4B).

The amounts of prolyl 4-hydroxylase activity in the Triton X-100 extracts of the cell homogenates showed a corresponding pattern (Table 1). The smallest amounts of enzyme activity in co-expression experiments with both the human and *C. elegans* α subunits were seen with the *C. elegans* PDI/ β polypeptide. The hybrid *C. elegans*-human PDI/ β polypeptide produced a consistently larger amount of prolyl 4-hydroxylase activity than the unmodified *C. elegans* polypeptide, and the hybrid human-*C. elegans* polypeptide a consistently smaller amount than the unmodified human polypeptide (Table 1).

DISCUSSION

The present and previous [28] data indicate that both the *C. elegans* and human PDI/ β polypeptides form an active prolyl 4-hydroxylase $\alpha_2\beta_2$ tetramer with the human α subunit and an active $\alpha\beta$ dimer with the *C. elegans* α subunit. The differences leading to either tetramer or dimer formation must therefore reside in the α subunits. Removal of the 32-amino-acid C-terminal extension from the *C. elegans* α subunit totally eliminated $\alpha\beta$ dimer formation, indicating that the sequences critical for dimer formation are present in this extension. Nevertheless the truncated *C. elegans* α subunit failed to form an $\alpha_2\beta_2$ tetramer even though the C-terminus of this polypeptide was in the same position as in the human [29] and chick [34] α subunits. It is thus evident that some additional structural feature present in the vertebrate α subunits but absent from the truncated *C. elegans* α subunit must be critical for tetramer assembly.

One surprising finding was that the *C. elegans* PDI/ β polypeptide consistently formed less prolyl 4-hydroxylase with both the human and *C. elegans* α subunits than did the human PDI/ β polypeptide. Furthermore the *C. elegans* PDI/ β polypeptide was much less effective in prolyl 4-hydroxylase formation with the *C. elegans* α subunit than with the human α subunit, in that the prolyl 4-hydroxylase activity generated in extracts of cell homogenates by co-expression of the *C. elegans* PDI/ β polypeptide and α subunit was usually only 1–5%, whereas it was 25–35%

on co-expression of the *C. elegans* PDI/ β and human α subunit, taking as 100% the values obtained in co-expression of the human PDI/ β polypeptide with either the human or *C. elegans* α subunits. These differences are not likely to be due to any misfolding of the *C. elegans* α subunit or PDI/ β polypeptide in insect cells, as the former effectively formed an enzyme dimer with the human PDI/ β polypeptide and the latter was a fully active disulphide isomerase. Nevertheless it is possible that effective association of the *C. elegans* α subunit and PDI/ β polypeptide requires some special condition or factor that is optimally present only in the *C. elegans* cells.

The possibility was also considered that *C. elegans* might have more than one PDI/ β polypeptide and that some of the others might form a prolyl 4-hydroxylase more effectively than the one cloned here. Indeed, analysis of the *C. elegans* database AceDB indicates the presence of several putative PDI isoforms identified during the continuing genome sequencing project [32]. The PDI isoform that was also expressed here was clearly not such a polypeptide, however, as it formed no prolyl 4-hydroxylase with either the *C. elegans* or human α subunit. Interestingly, it differs from the PDI/ β and ERp60 [1–3] polypeptides obtained from various sources in that the N-terminal thioredoxin-like domain *a* has a catalytic site sequence -CVHC- instead of the usual -CGHC-. In spite of this, the isomerase activity of the PDI isoform was found to be approximately the same as that of the human PDI/ β polypeptide.

Experiments with the *C. elegans*–human and human–*C. elegans* PDI/ β hybrid polypeptides demonstrated that sequences present within the C-terminal region are likely to be involved in prolyl 4-hydroxylase dimer and tetramer assembly. When the C-terminal region of the *C. elegans* PDI/ β polypeptide was replaced with that of the human polypeptide, a distinct increase was found in the amounts of active prolyl 4-hydroxylase $\alpha\beta$ dimer and $\alpha_2\beta_2$ tetramer, and when the human PDI/ β polypeptide received the C-terminal region from the *C. elegans* polypeptide a distinct decrease was found in the amounts of dimer and tetramer. Nevertheless the presence of the human C-terminal region in the *C. elegans* polypeptide did not increase the amount of prolyl 4-hydroxylase formation to the level seen with the unmodified human polypeptide. Thus the differences in the sequences of the C-terminal regions cannot be the only reason for the marked differences in prolyl 4-hydroxylase formation between the *C. elegans* and human PDI/ β polypeptides.

The catalytic properties of the *C. elegans* prolyl 4-hydroxylase dimer were found to be very similar to those of the vertebrate type II prolyl 4-hydroxylase tetramer [27]. Hydroxylation experiments with [¹⁴C]proline-labelled protocollagen as the substrate demonstrated that the K_m values of the *C. elegans* enzyme dimer and a vertebrate enzyme tetramer for this substrate are essentially identical. This result is not consistent with a processive model for the hydroxylation of proline residues in long polypeptide substrates [47], as the model assumes that the K_m values of an enzyme containing only one catalytic site are markedly higher than those of an enzyme containing two catalytic sites. Nevertheless our results do not exclude the possibility that the *C. elegans* enzyme dimer forms a tetramer under the conditions of the hydroxylation reaction, although this does not seem very likely.

We thank Hanna-Mari Jauho, Eeva Lehtimäki, Riitta Polojärvi and Liisa Äijälä for expert technical assistance, Ari-Pekka Kvist for help in the computational work and Alan Coulson of the MRC, Cambridge, U.K., for supplying the polytene YAC filter. This research was supported by grants from the Medical Research Council of the Academy of Finland. A.P.P. is funded by the Medical Research Council of Great Britain through a Career Development Fellowship.

REFERENCES

- Noiva, R. and Lennarz, W. J. (1992) *J. Biol. Chem.* **267**, 3553–3556
- Freedman, R. B., Hirst, T. R. and Tuite, M. F. (1994) *Trends Biochem. Sci.* **19**, 331–336
- Freedman, R. B. (1995) *Curr. Opin. Struct. Biol.* **5**, 85–91
- Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A. and Rutter, W. J. (1985) *Nature (London)* **317**, 267–270
- Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllylä, R., Huhtala, M.-L., Koivu, J. and Kivirikko, K. I. (1987) *EMBO J.* **6**, 643–649
- Kivirikko, K. I., Myllylä, R. and Pihlajaniemi, T. (1989) *FASEB J.* **3**, 1609–1617
- Hawkins, H. C. and Freedman, R. B. (1991) *Biochem. J.* **275**, 335–339
- Vuori, K., Myllylä, R., Pihlajaniemi, T. and Kivirikko, K. I. (1992) *J. Biol. Chem.* **267**, 7211–7214
- Vuori, K., Pihlajaniemi, T., Myllylä, R. and Kivirikko, K. I. (1992) *EMBO J.* **11**, 4213–4217
- Lyles, M. M. and Gilbert, H. F. (1994) *J. Biol. Chem.* **269**, 30946–30952
- Darby, N. J. and Creighton, T. E. (1995) *Biochemistry* **34**, 3576–3587
- Koivu, J., Myllylä, R., Helaakoski, T., Pihlajaniemi, T., Tasanen, K. and Kivirikko, K. I. (1987) *J. Biol. Chem.* **262**, 6447–6449
- Parkkonen, T., Kivirikko, K. I. and Pihlajaniemi, T. (1988) *Biochem. J.* **256**, 1005–1011
- Wetterau, J. R., Combs, K. A., Spinner, S. N. and Joiner, B. J. (1990) *J. Biol. Chem.* **265**, 9800–9807
- Wetterau, J. R., Combs, K. A., McLean, L. R., Spinner, S. N. and Aggerbeck, L. P. (1991) *Biochemistry* **30**, 9728–9735
- LaMantia, M.-L. and Lennarz, W. J. (1993) *Cell* **74**, 899–908
- Noiva, R., Freedman, R. B. and Lennarz, W. J. (1993) *J. Biol. Chem.* **268**, 19210–19217
- Cai, H., Wang, C.-C. and Tsou, C.-L. (1994) *J. Biol. Chem.* **269**, 24550–24552
- Otsu, M., Omamura, F., Yoshimori, T. and Kikuchi, M. (1994) *J. Biol. Chem.* **269**, 6874–6877
- Puig, A. and Gilbert, H. F. (1994) *J. Biol. Chem.* **269**, 25889–25896
- Puig, A., Lyles, M. M., Noiva, R. and Gilbert, H. F. (1994) *J. Biol. Chem.* **269**, 19128–19135
- Rupp, K., Birnbach, U., Lundström, J., Nguyen, P. and Söling, H. D. (1994) *J. Biol. Chem.* **269**, 2501–2507
- Song, J. L. and Wang, C. C. (1995) *Eur. J. Biochem.* **231**, 312–316
- Prockop, D. J. and Kivirikko, K. I. (1995) *Annu. Rev. Biochem.* **64**, 403–434
- Kivirikko, K. I., Helaakoski, T., Tasanen, K., Vuori, K., Myllylä, R., Parkkonen, T. and Pihlajaniemi, T. (1990) *Ann. N. Y. Acad. Sci.* **580**, 132–142
- Kivirikko, K. I., Myllylä, R. and Pihlajaniemi, T. (1992) in *Post-Translational Modifications of Proteins* (Harding, J. J. and Grabbe, M. J. C., eds.), pp. 1–51, CRC Press, Boca Raton, FL
- Helaakoski, T., Annunen, P., Vuori, K., MacNeil, I., Pihlajaniemi, T. and Kivirikko, K. I. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4427–4431
- Veijola, J., Koivunen, P., Annunen, P., Pihlajaniemi, T. and Kivirikko, K. I. (1994) *J. Biol. Chem.* **269**, 26746–26753
- Helaakoski, T., Vuori, K., Myllylä, R., Kivirikko, K. I. and Pihlajaniemi, T. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4392–4396
- Coulston, A., Waterston, R., Kiff, J., Soulston, J. A. and Kohare, Y. (1988) *Nature (London)* **335**, 184–188
- Frohman, M. (1990) in *PCR Protocols: a Guide to Methods and Applications* (Innis, M. A., Gelfond, D. H., Shinsky, J. J. and White, T. J., eds.), pp. 28–38, Academic Press, San Diego, CA
- Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latreille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Showkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J. and Wohldman, P. (1994) *Nature (London)* **368**, 32–38
- Vuori, K., Pihlajaniemi, T., Marttila, M. and Kivirikko, K. I. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7467–7470
- Bassuk, J. A., Kao, W. W.-Y., Herzer, P., Kedersha, N. I., Seyer, J., DeMartino, J. A., Daugherty, B. L., Mark, III, G. E. and Berg, R. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7382–7386
- Gruenewald, S. and Heitz, J. (1993) *Baculovirus Expression Vector System: Procedures and Methods Manual*, Pharmingen, San Diego, CA
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467

- 37 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410
- 38 Harlow, E. and Lane, D. (1988) *Antibodies: a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 39 Kivirikko, K. I. and Myllylä, R. (1982) *Methods Enzymol.* **82**, 245–304
- 40 Myllylä, R., Tuderman, L. and Kivirikko, K. I. (1977) *Eur. J. Biochem.* **80**, 349–357
- 41 Lyles, M. M. and Gilbert, H. F. (1991) *Biochemistry* **30**, 613–619
- 42 Veijola, J., Pihlajaniemi, T. and Kivirikko, K. I. (1996) *Biochem. J.* **315**, 613–618
- 43 MacCombie, W. R., Adams, M. D., Kelly, J. M., Fitzgerald, M. G., Utterback, T. R., Khan, M., Dubnick, M., Kerlavage, A. R., Venter, J. C. and Field, C. (1992) *Nature Genet.* **1**, 124–131
- 44 von Hejne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690
- 45 Pelham, H. R. B. (1989) *Annu. Rev. Cell. Biol.* **5**, 1–23
- 46 Koivunen, P., Helaakoski, T., Annunen, P., Veijola, J., Räisänen, S., Pihlajaniemi, T. and Kivirikko, K. I. (1996) *Biochem. J.* **316**, 599–605
- 47 de Jong, L., van der Kraan, I. and de Waal, A. (1991) *Biochim. Biophys. Acta* **103**, 103–111

Received 15 February 1996; accepted 25 March 1996