PROLYL 4-HYDROXYLASE

Genomic cloning of the human and mouse α(II) subunit, tissue distribution of type I and II isoenzymes, and cloning and characterization of a novel prolyl 4-hydroxylase from *Caenorhabditis elegans*

RITVA NISSI

Collagen Research Unit,
Biocenter Oulu and
Department of Medical Biochemistry and
Molecular Biology, University of Oulu

OULU 2002

RITVA NISSI

PROLYL 4-HYDROXYLASE

Genomic cloning of the human and mouse α(II) subunit, tissue distribution of type I and II isoenzymes, and cloning and characterization of a novel prolyl 4-hydroxylase from *Caenorhabditis elegans*

Academic Dissertation to be presented with the assent of the Faculty of Medicine, University of Oulu, for public discussion in the Auditorium of the Department of Medical Biochemistry and Molecular Biology, on July 4th, 2002, at 12 noon.

OULUN YLIOPISTO, OULU 2002

Reviewed by
Docent Mikko Lammi
Docent Sirkku Peltonen

ISBN 951-42-6756-7 (URL: http://herkules.oulu.fi/isbn9514267567/)

ALSO AVAILABLE IN PRINTED FORMAT
Acta Univ. Oul. D 686, 2002
ISBN 951-42-6755-9
ISSN 0355-3221 (URL: http://herkules.oulu.fi/issn03553221/)

OULU UNIVERSITY PRESS
OULU 2002

Nissi, Ritva, Prolyl 4-hydroxylase Genomic cloning of the human and mouse α(II) subunit, tissue distribution of type I and II isoenzymes, and cloning and characterization of a novel prolyl 4-hydroxylase from *Caenorhabditis elegans*

Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, P.O.Box 5000, FIN-90014 University of Oulu, Finland
Oulu, Finland
2002

Abstract

The collagens are a family of extracellular matrix proteins with a widespread tissue distribution. Collagen biosynthesis requires the hydroxylation of a number of proline residues by prolyl 4-hydroxylase. This posttranslational modification is essential for the synthesis of all collagens, as 4-hydroxyproline deficient collagens cannot form stable triple helices at body temperature.

The genes for the human and mouse prolyl 4-hydroxylase α(II) subunits were cloned and characterized in this study. The human and mouse genes are 34.6 and 30.3 kb in size, respectively, consisting of 16 exons and 15 introns. The intron sizes vary from 48-49 bp to over 8 kb in both genes. The 5' flanking regions contain no TATA box, but there are several motifs that may act as transcription factor binding sites. A novel mutually exclusively spliced exon 12a was identified in both genes. Both variants of the α(II) subunit were found to be expressed in a variety of tissues and both formed a fully active recombinant tetramer with the β subunit when expressed in insect cells.

Tissue distribution of the type I and type II prolyl 4-hydroxylase isoenzymes was studied in developing, mature, and malignant cells and tissues by immunofluorescence and Western blotting. The results indicate that the type I isoenzyme is the main form in many cell types. Skeletal myocytes and smooth muscle cells appeared to have the type I isoenzyme as their only prolyl 4-hydroxylase form, whereas the type II isoenzyme was clearly the main form in chondrocytes. A strong signal for the type II enzyme was detected in cultured umbilical and capillary endothelial cells, whereas the type I isoenzyme could not be detected in these cells by immunostaining or Western blotting. Similar studies on primary chondro- and osteosarcomas and benign bone tumours indicated that the type I isoenzyme is the predominant form in both types of bone sarcoma, whereas the type II isoenzyme was more abundantly expressed in benign tumours. In chondrosarcomas, the type II isoenzyme was expressed in the nonmalignant chondrocytes, whereas their malignant counterparts switched their expression pattern to that of the type I isoenzyme.

Two isoforms of the catalytic prolyl 4-hydroxylase α subunit, PHY-1 and PHY-2, have previously been characterized from *Caenorhabditis elegans*. This study reports the cloning and characterization of a third *C. elegans* α subunit isoform, PHY-3, which is much shorter than the previously characterized vertebrate and *C. elegans* α subunits. Nematodes homozygous for a *phy-3* deletion were phenotypically wild type and fertile, but the 4-hydroxyproline content of their early embryos was reduced by about 90%. The expression of PHY-3 was found to be restricted to spermatheca of late larvae and adult nematode, indicating that PHY-3 is likely to be involved in the synthesis of collagens of the early embryo egg shells.

Keywords: gene structure, prolyl 4-hydroxylase, isoenzyme, collagen biosynthesis, primary bone tumours, tissue specificity, C. elegans

Acknowledgements

This work was carried out at the Collagen Research Unit and the Department of Medical Biochemistry and Molecular Biology, University of Oulu and Biocenter Oulu.

My most sincere thanks goes to my supervisors, Docent Johanna Myllyharju and Professor Kari Kivirikko. Their invaluable contribution and always so encouraging attitude toward me have been essential for the completion of this work. I would also like to thank Professor Taina Pihlajaniemi and Biocenter Oulu for providing excellent research facilities, and similarly I would also like to acknowledge Professor Ilmo Hassinen and Professor Leena Ala-Kokko. I am very grateful to all my co-authors, Docents Helena Autio-Harmainen, Tom Böhlig, Tarja Helaakoski, Raija Sormunen, Drs Pia Marttila, Minna Nokelainen, Antony Page, Päivi Riihimaa and Ph.D. students Katriina Keskiaho and Alan Winter for their excellent cooperation. I am grateful to Docent Sirkku Peltonen and Docent Mikko Lammi for their valuable and rapid comments on the manuscript of this thesis. I also wish to acknowledge Sandra Barkoczy for revising the language of this manuscript.

I would like to acknowledge the whole staff of the Department of Medical Biochemistry and Molecular Biology, especially Pirkko Huhtala, Liisa Kukkola, Joni Mäki, Kati Rautavuoma, Laura Silvennoinen, Kati Takaluoma, Hongmin Tu, Timo Väisänen, and Jussi Vuoristo, and also those at the other Departments of the Medical Faculty, especially Heidi Eriksen. Without their support, this work would have been much harder to carry out. In addition, Joni Mäki deserves special thanks for his kind and efficient help with the computing matters. Also the members of the “computer-support-crew” are acknowledged.

I am grateful to Auli Kinnunen, Marja-Leena Kivelä, Seija Leskelä, Seppo Lähdesmäki, and Pertti Vuokila for their kind help with the numerous practical problems during these years. I would also thank Liisa Äijälä for her excellent technical assistance and friendship.

I would like to express my special thanks to my mother for her support and love throughout my life. Also my brother Lasse and sister Sirpa deserve my warmest thanks for their encouragement. Finally, I would like to thank my fiancé Vesa for his unfailing love and support.

This work was supported by the Paulo Foundation, the Finnish Cultural Foundation and Finnish Centre of Excellence Programme (2000-2005) of the Academy of Finland (44843).

Oulu, June 2002

Abbreviations

C. elegans	*Caenorhabditis elegans*
Cle-1	*C. elegans* Type XVIII collagen gene
Dpy-18	*C. elegans* PHY-1 knock-out line
GFP	Green fluorescent protein
Egl-9	*C. elegans* HPH gene
EDS	Ehlers-Danlos syndrome
ER	Endoplasmic reticulum
EST	Expressed sequence tag
HIF	Hypoxia-inducible factor
HPH	HIF prolyl 4-hydroxylase
Let-268	*C. elegans* lysyl hydroxylase gene
PDI	Protein disulfide isomerase
PHY	*C. elegans* prolyl 4-hydroxylase catalytic subunit
RACE	Rapid amplification of cDNA ends
RNAi	RNA interference
TRITC	Tetramethylrhodamine isothiocyanate
X (in -Gly-X-Y-)	any amino acid
Y (in -Gly-X-Y-)	any amino acid

List of original articles

This thesis is based on the following articles, which are referred to in the text by their Roman numerals:

I Nokelainen M, Nissi R, Kukkola L, Helaakoski, T & Myllyharju J (2001) Characterization of the human and mouse genes for the α subunit of type II prolyl 4-hydroxylase. Identification of a previously unknown alternatively spliced exon and its expression in various tissues. Eur J Biochem 268: 5300-5309.

II Nissi R, Autio-Harmainen H, Marttila P, Sormunen R & Kivirikko KI (2001) Prolyl 4-hydroxylase isoenzymes I and II have different expression patterns in several human tissues. J Histochem Cytochem 49: 1143-1153.

III Nissi R, Böhling T & Autio-Harmainen H. Immunofluorescence localization of prolyl 4-hydroxylase isoenzymes in bone tumours: Type I enzyme predominates in osteo- and chondrosarcomas, but the type II enzyme is more abundantly expressed in their benign counterparts. Submitted.

IV Riihimaa P, Nissi R, Page AP, Winter AD, Keskiaho K, Kivirikko KI & Myllyharju J (2002) Egg shell collagen formation in *Caenorhabditis elegans* involves a novel prolyl 4-hydroxylase expressed in spermatheca and embryos and possessing many unique properties. J Biol Chem 277: 18238-18243.

Contents

1 Introduction

The collagens are a family of structural proteins of the extracellular matrix consisting of three polypeptide chains with a repeated -Gly-X-Y- sequence, and are characterized by the presence of at least one triple–helical domain. They are involved in multiple biological functions, such as providing tissue strength and affecting cell attachment and tissue development.

Prolyl 4-hydroxylase is a critical enzyme in collagen biosynthesis, since it catalyzes the formation of 4-hydroxyproline residues that are essential for the stabilization of the collagen triple helix at body temperature. The molecular properties of prolyl 4-hydroxylase have been characterized in a number of species and the active enzyme in all vertebrates studied has been found to be an $\alpha_2\beta_2$ tetramer. The α subunits, and especially their C-terminal regions, are the main contributors to the two catalytic sites of the enzyme. The β subunit is identical to protein disulfide isomerase and its main function in prolyl 4-hydroxylase is to keep the α subunits in a catalytically active, nonaggregated conformation.

Two isoforms of the human and mouse α subunits have been isolated and characterized previously. The first identified α subunit is called the α(I) subunit and the more recently isolated one, the α(II) subunit. Both isoforms form a tetramer with the same β subunit, and the $[\alpha(I)_2]\beta_2$ and $[\alpha(II)_2]\beta_2$ tetramers are termed the type I and II prolyl 4-hydroxylases, respectively. The aim of the present research was to isolate and characterize the human and mouse α(II) subunit genes to obtain information on their structures and function. The expression patterns of the two types of prolyl 4-hydroxylase isoenzyme were studied by immunolocalization in developing, mature, and malignant human tissues, and the data obtained indicate that the type I prolyl 4-hydroxylase is the main isoenzyme in several cell types. It is expressed especially in developing and dedifferentiated cells, whereas the type II isoenzyme is expressed in more differentiated cells.

Caenorhabditis elegans was the first eukaryote with a completely known genomic sequence. Two prolyl 4-hydroxylase α subunit isoforms, PHY-1 and PHY-2, that are expressed in the collagen synthesizing hypodermal cells, have been characterized from *C. elegans*. The present study reports the cloning and characterization of a third *C. elegans* prolyl 4-hydroxylase isoform (PHY-3), that has a unique tissue distribution and function.

2 Review of the literature

2.1 Occurrence and functions of 4-hydroxyproline in animal proteins

Most of the 4-hydroxyproline in animal proteins is found in collagens and in several other proteins containing collagenous domains. There are also more than 15 additional proteins having collagenous domains that are not defined as collagens, including the subcomponent C1q of complement (Reid 1979), acetylcholinesterase (Rosenberry *et al.* 1982), pulmonary surfactant proteins SP-A and SP-D (Floros *et al.* 1986, Lu *et al.* 1992), mannan-binding protein (Drickamer *et al.* 1986, Colley & Baenziger 1987, Reid 1993), conglutin (Davis & Lachmann 1984), ficolins (Ichijo *et al.* 1993), type I and II macrophage scavenger receptors (Kodama *et al.* 1990, Rohrer *et al.* 1990), macrophage scavenger receptor MARCO (Elomaa *et al.* 1995), adipose -specific collagen-like factor apM1 (Maeda *et al.* 1997), collectin-43 (Jensenius *et al.* 1994), aggretin (Chung *et al.* 1999), ectodysplasin (Srivastava *et al.* 1997, Ezer *et al.* 1999), and src-homologous and collagen (SHC) protein (Pelicci *et al.* 1992, Pelicci *et al.* 1996). Like collagens, these proteins have been shown to contain 4-hydroxyproline and hydroxylysine, but unlike collagens, they may not contain 3-hydroxyproline (Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998).

Collagens are a family of closely related but distinct extracellular matrix proteins and constitute the major component of the extracellular matrix. They play a dominant role in maintaining the integrity of various tissues and also have a number of other important functions. (Mendler *et al.* 1989, Kivirikko *et al.* 1992, Kielty *et al.* 1993, Kivirikko 1993, Fichard *et al.* 1994, Pihlajaniemi & Rehn 1995, Prockop & Kivirikko 1995, O´Reilly *et al.* 1997, Ricard-Blum *et al.* 2000, Kivirikko & Pihlajaniemi 1998, Myllyharju & Kivirikko 2001). They consist of three polypeptide chains containing repeating -Gly-X-Y- triplets, in which X is often proline and Y is often 4-hydroxyproline, with the exception that there are a few –Ala-X-4Hyp- sequences in the human subcomponent C1q of the complement and also in some collagen chains. The 4-hydroxyproline content of different collagen types varies between about 60 and 130 residues per 1000 amino acids (Kivirikko & Myllylä 1980, Kivirikko *et al.* 1992).

Table 1. Tissue distribution of the known collagens. The references are indicated in the text.

Collagen type	Tissue distribution
I	Most connective tissues and developing tissues. Abundant in dermis, bone, skin, tendon, ligament, teeth, fasciae.
II	Cartilage, intervertebrate disc, inner ear, vitreous humor, cornea. Low levels in many non-chondrogenic tissues during development.
III	Abundantly expressed in elastic tissues such as skin, inner organs, and blood vessels. Found in the surface of type I collagen fibrils.
IV	All basement membranes, forms an interface between epithelial, endothelial, muscular, and neural cells and the adjacent stroma.
V	Widespread in low quantities, co-localizes with type I collagen, type V being located in the core of the fibril. Some tissues have hybrid type V/XI molecules. Schwann cells in peripheral nerves.
VI	Minor component of soft tissues, the major component of beaded microfibrils located on the surface of cells and around collagen fibers.
VII	Anchoring fibrils in skin, cornea, cervix, oral and oesophageal mucosa.
VIII	Descemet´s membrane, sclera, chorioid, optic nerve, dura mater, perichondrium of cartilage, connective tissue around hair follicles, arterioles and venules.
IX	Found in the surface of type II collagen fibrils, minor amounts of foetal, non-chondogenic tissues.
X	Hypertrophic chondrocytes.
XI	Found in the core of type II collagen fibrils, minor amounts of foetal non-chondogenic tissues.
XII	Dense connective tissue containing type I collagen *e.g.* cornea, meninges, blood vessels, endo- and perineurium.
XIII	Widespread in low quantities, often found in conjunction with cell-matrix contact sites, contains a transmembrane domain.
XIV	Minor component in many tissues, especially on skin and nerve sheaths.
XV	Many tissues in basement membrane zone, contains an endostatin domain.
XVI	Low quantities in many tissues, such as heart, kidney, intestine, ovary, testis, eye, arterial walls, smooth muscle.
XVII	Hemidesmosomes in skin, also in oral and corneal epithelial hemidesmosomes, contains a transmembrane domain.
XVIII	In a wide variety of tissues in the basement membrane zone, contains an endostatin domain.
XIX	In a wide variety of tissues in the basement membrane zone.
XX	Minor component of several connective tissues, most prevalent in chick corneal epithelium.
XXI	Higher levels in developing tissues.
XXV	Also called CLAC (collagen-like Alzheimer amyloid plaque component). Contains a transmembrane domain.

The collagen polypeptide chains are folded into a left-handed helix, and the three chains are then wrapped around each other into a right-handed triple helix. At least 22

proteins are now defined as collagens, as shown in table 1 (Fukai *et al.* 1994, Prockop & Kivirikko 1995, Pihlajaniemi & Rehn 1995, Bateman *et al.* 1996, Cremer *et al.* 1998, Blaschke *et al.* 2000, Ricard-Blum *et al.* 2000, Fitzgerald & Bateman 2001, Koch *et al.* 2001, Myllyharju & Kivirikko 2001, Chou & Li 2002, Hashimoto *et al.* 2002).
Collagen types can be classified based on their ability to form fibrils or other supramolecular structures (Prockop & Kivirikko 1995, Bateman *et al.* 1996, Myllyharju & Kivirikko 2001). The group of fibrillar or fibril-forming collagens includes types I, II, III, V, and XI. These molecules have a single uninterrupted triple-helical domain that is available for fibril formation. Their triple-helical domains consist of about 1000 residues of uninterrupted -Gly-X-Y- sequences in each of the three chains, and they form highly ordered fibrils with a 67 nm banded pattern (for further details, see Kielty *et al.* 1993, Brewton & Mayne 1994, Prockop & Kivirikko 1995, Bateman *et al.* 1996, Myllyharju & Kivirikko 2001).

Other collagen types do not form fibrils, even though type VIII and X collagens are known as the short-chain collagens, since their constituent polypeptide chains consists of a single uninterrupted collagenous domain being about half the size of those of the fibril forming collagens (Hulmes 1992, Bateman *et al.* 1996). Some collagen types have unique functions in specialized extracellular matrix structures, *e.g.* type IV in basement membranes and type VII in anchoring fibrils, which are specialized structures at the dermo-epidermal junction that link the basement membranes to the anchoring plaques (Uitto & Christiano 1992, Bruckner-Tuderman 1994, Prockop & Kivirikko 1995, Bateman *et al.* 1996, Myllyharju & Kivirikko 2001). It has been proposed that type VI collagen microfibrils may perform a bridging function between the cells and the extracellular matrix and may be of particular significance in directing tissue remodeling (Timpl & Engel 1987, Prockop & Kivirikko 1995). The largest subgroup of the non-fibrillar collagens consists of FACITs (fibril-associated collagens with interrupted triple helices), which includes types IX, XII, XIV, XVI, XIX, XX, and XXI (Bruckner *et al.* 1985, Huber *et al.* 1986, McCormick *et al.* 1987, Watt *et al.* 1992, Wu *et al.* 1992, Fitzgerald & Bateman 2001, Koch *et al.* 2001, Myllyharju & Kivirikko 2001, Chou & Li 2002).

Collagens have also been characterized from invertebrates (Engel 1997). These include minicollagens from *Cnidaria* nematocytes (Kurz *et al.* 1991, Engel *et al.* 2001), the longest known collagens from annelids (Gaill *et al.* 1991), the cuticle and basement membrane collagens from *C. elegans* (Johnstone 1994, Johnstone 2000), the collagen from mussel byssus threads (Coyne *et al.* 1997), and two basement membrane collagens from *Drosophila* (Yasothornsikul *et al.* 1997, Hynes & Zhao 2000, Kuo *et al.* 2001). Most invertebrate collagens resemble vertebrate collagens in their 4-hydroxyproline content, but there are exceptions. The 4-hydroxyproline content is only 20 residues per 1000 amino acids in *Ascaris* cuticle collagen, whereas in earthworm cuticle collagen it is 160 residues per 1000 amino acids (Adams 1978, Murray & Tanzer 1983). The earthworm cuticle collagen also differs from the others in that 4-hydroxyproline residues are found in both the X and Y positions of the repeating -Gly-X-Y- sequences (Kivirikko *et al.* 1992).

Nonhydroxylated collagen polypeptide chains can fold into a triple helix at low temperatures. The midpoint of thermal transition from helix to coil (T_m) is 24°C in a nonhydroxylated type I procollagen molecule, whereas the T_m of a fully hydroxylated

type I procollagen molecule is about 39°C (Berg & Prockop 1973, Jimenez *et al.* 1973). Therefore, almost complete 4-hydroxylation of proline residues in the Y positions of the -Gly–X-Y- repeats is required for the formation of a stable collagen molecule at the physiological body temperature. The stabilizing effect of 4-hydroxyproline residues on the collagen triple helix has been thought to arise from the generation of water-mediated hydrogen bonds within the same chain or with the adjacent chain (Fraser & McRae 1973, Brodsky & Ramshaw 1997). The hydroxy group of 4-hydroxyproline can act as a hydrogen bond acceptor or a donor, and water molecules can bind it at two sites (Bella *et al.* 1995, Bella & Berman 1996, Kramer *et al.* 1998, Kramer *et al.* 2001). However, the presence of electronegative fluoroproline residues in the Y position of synthetic peptides has also been shown to stabilize the triple helix. This indicates that the stabilization is in fact due to a stereoelectronic effect, which also applies to 4-hydroxyproline residues (DeRider *et al.* 2002, Jenkins & Raines 2002).

Elastin is the main protein component of elastic fibers. It has no collagen-like triple-helical domains, but it has repeating -Gly-X-Y- sequences containing 4-hydroxyproline residues in the Y position. The 4-hydroxyproline content of elastin varies greatly, ranging from 10 to 50 residues per 1000 amino acids (Kivirikko *et al.* 1992, Rosenbloom *et al.* 1993). The function of 4-hydroxyproline residues in elastin is currently unknown (Debelle & Tamburro 1999). A single 4-hydroxyproline residue is found in an -X-4Hyp-Gly- sequence in hydroxyproline-lysyl-bradykinin, hydroxyproline luteinizing hormone-releasing hormone (Kivirikko *et al.* 1992), and in the prion protein, PrP (Gill *et al.* 2000). Hypoxia inducible transcription factor (HIF) has recently been shown to contain 4-hydroxyproline residues in two highly unusual –Leu-X-X-Leu-Ala-4Hyp- sequences (Ivan *et al*. 2001, Jaakkola *et al.* 2001, Epstein *et al.* 2001, Bruick & McKnight 2001, for more details, see chapter 2.3.). In addition, synthetic peptides representing noncollagenous proline-rich repeats encoded by the genome of the eukaryotic algal virus *Paramecium bursaria Chlorella* virus-1 have been shown to be hydroxylated by a recombinant viral prolyl 4-hydroxylase (Eriksson *et al.* 1999).

2.2 Collagen prolyl 4-hydroxylases

2.2.1 Vertebrate collagen prolyl 4-hydroxylases

Prolyl 4-hydroxylase (EC 1.14.11.2, procollagen-proline, 2-oxoglutarate 4-dioxygenase) catalyzes the formation of 4-hydroxyproline in collagens and certain other proteins with collagen-like sequences by hydroxylating proline residues in the Y positions of the -X-Y-Gly- sequences. It is present in cells of mesenchymal and nonmesenchymal origin, and a good correlation usually exists between the amount of prolyl 4-hydroxylase activity and the rate of collagen synthesis (Cardinale & Udenfriend 1974, Kivirikko & Myllylä 1980, Annunen *et al.* 1998). The vertebrate enzyme is an $\alpha_2\beta_2$ tetramer with a molecular weight of about 240 000 Da, and consists of two types of monomers with molecular weights of about 63 000 (α subunit) and 58 000 (ß subunit). Subcellular fractionation studies have

demonstrated that prolyl 4-hydroxylase is located in the rough ER (Peterkofsky & Assad 1976) and both the active enzyme tetramer and the PDI/β polypeptide are soluble ER luminal proteins (Kivirikko 1995, Kivirikko & Myllyharju 1998, Kivirikko & Pihlajaniemi 1998).

All attempts to assemble a prolyl 4-hydroxylase tetramer from its subunits *in vitro* have been unsuccessful apparently because of the tendency of the α subunit to aggregate as a monomer. However, an active recombinant tetramer has been produced by coexpression of the α and β subunits in insect cells by means of baculovirus vectors (Vuori *et al.* 1992b), in the yeast *Pichia pastoris* (Vuorela *et al.* 1997), and in mammalian cells (John & Bulleid 1996, Wagner *et al.* 2000), and by expressing a recombinant α subunit in a cell-free protein synthesis system in the presence of canine pancreas microsomal vesicles (John *et al.* 1993).

2.2.2 Molecular properties of vertebrate collagen prolyl 4-hydroxylases

2.2.2.1 α subunit

Prolyl 4-hydroxylase had long been assumed to be of one type only, but a cDNA isoform of the vertebrate α subunit, termed the α(II) subunit, has been cloned and characterized from mouse and man (Helaakoski *et al.* 1995, Annunen *et al.* 1997). Correspondingly, the previously known vertebrate α subunit, which has been cloned from man (Helaakoski *et al.* 1989), mouse (Helaakoski *et al.* 1995), rat (Hopkinson *et al.* 1994), and chicken (Bassuk *et al.* 1989), is now called the α(I) subunit. Both types of α subunit form an $\alpha_2\beta_2$ tetramer with the same β subunit, the $[\alpha(I)_2]\beta_2$ and $[\alpha(II)_2]\beta_2$ tetramers being termed the type I and II prolyl 4-hydroxylases, respectively. Recombinant insect cell expression experiments strongly argue against the formation of mixed $[\alpha(I)\alpha(II)]\beta_2$ vertebrate tetramers (Annunen *et al.* 1997).

The processed human α(I) and α(II) subunits consist of 517 and 514 amino acids, respectively, both being synthesized in a form containing a signal peptide of 17 and 21 amino acids. The amino acid sequence of the human α(II) subunit shows a 64% identity and 81% similarity to that of the human α(I) subunit, and 93% identity and 96% similarity to that of the mouse α(II) subunit (Helaakoski *et al.* 1995, Annunen *et al.* 1997), the identity being highest within the C-terminal regions that contain the residues involved in cosubstrate binding (see Chapter 2.2.2.). Five cysteine residues are conserved between the α(I) and α(II) subunits, and site-directed mutagenesis studies suggest that intrachain disulfide bonds are formed between the second and third, and the fourth and fifth cysteines in the α(I) subunit, both of these bonds being essential for the mainteinance of the native structure of the α subunit needed for tetramer assembly (John & Bulleid 1994, Lamberg *et al.* 1995). The positions of the five cysteine residues are conserved in the human, mouse, and chicken α(I) subunits and the human and mouse α(II) subunits, but the latter two contain an additional cysteine located between the fourth and fifth cysteines. The prolyl 4-hydroxylase tetramers have no interchain disulfide

bonds. Vertebrate α(I) and α(II) subunits contain two potential attachment sites for asparagine-linked oligosaccharide units (Bassuk *et al.* 1989, Helaakoski *et al.* 1989, Hopkinson *et al.* 1994, Helaakoski *et al.* 1995, Annunen *et al.* 1997), but removal of the sugars or mutation of the glycosylation sites do not affect the enzyme tetramer formation or activity (Kedersha *et al.* 1985, Lamberg *et al.* 1995). The extent of glycosylation of the two asparagines is variable, thus causing heterogeneity in the enzyme tetramers.

The human and mouse α(II) subunit mRNAs are 2.3 and 2.4 kb in size, respectively, whereas the size of the human α(I) subunit mRNA is 3 kb due to its longer untranslated 3´-end. Two mRNAs with the sizes of 3.4 kb and 4.3 kb are generated from the mouse α(I) subunit gene, and the shorter mRNA form seems to be the predominant one (Helaakoski *et al.* 1995, Annunen *et al.* 1997). The mRNAs for the α(I) and α(II) subunits are expressed in a variety of human and mouse cells and tissues. Northern analysis revealed that they are expressed in all the tissues studied, including heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, and testis (Helaakoski *et al.* 1989, Helaakoski *et al.* 1995, Annunen *et al.* 1997). The hybridization signal for the α(II) subunit mRNA was weaker than that for the α(I) subunit mRNA in testis, liver, skeletal muscle, kidney, and brain (Helaakoski *et al.* 1995, Annunen *et al.* 1997).

The percentages of the type I and type II isoenzymes of the total prolyl 4-hydroxylase activity in several tissues and cell types have been studied using a method based on the different poly(L-proline) binding properties of the two isoenzymes (Annunen *et al.* 1997, Annunen *et al.* 1998). The type I enzyme was found to be the main active form in several chicken embryo tissues, including calvaria, sternum, tendon, and liver, representing about 85-95% of the total prolyl 4-hydroxylase activity (Annunen *et al.* 1997). These results agree with the early reports indicating that at least 80 % of the total prolyl 4-hydroxylase activity in crude extracts from whole chicken embryos is bound to a poly(L-proline) column (Tuderman *et al.* 1975). The type I enzyme accounts for 60-90%, and the type II enzyme for 10-40%, of the total prolyl 4-hydroxylase activity in many mouse cell types and tissues. In cultured chondrocytes and in cartilage, however, the type II enzyme represents about 70-80%, and in mouse bone almost 50%, of the total activity (Annunen *et al.* 1998). Type I prolyl 4-hydroxylase is the main isoenzyme in cultured adult and foetal human skin fibroblasts, mouse embryonal fibroblasts, human embryonic lung fibroblasts, human fibrasarcoma 1080 cells, and human rhabdomyosarcoma cells (Annunen *et al.* 1997, Annunen *et al.* 1998). The proportion of the type I isoenzyme is increased further in these cells in confluency when compared to the logarithmic phase of growth (Annunen *et al.* 1998). Previously, prolyl 4-hydroxylase activity has been studied in normal and transformed cell lines, and in the latter the amount of prolyl 4-hydroxylase activity has been found to be decreased (Myllylä *et al.* 1981, Pihlajaniemi *et al.* 1981). This decrease was similar for both isoenzymes (Annunen *et al.* 1998).

Immunofluorescence studies (Annunen *et al.* 1998) of undifferentiated mesenchymal cells showed strong staining for the type I isoenzyme and weak staining for the type II isoenzyme. Like fibroblasts, the smooth muscle cells stained strongly for the type I isoenzyme. Both enzymes were expressed during endochondral ossification, but the type I enzyme appears to be expressed in the earliest phases of bone morphogenesis. In capillary endothelial cells, the type II enzyme seemed to be the only enzyme form expressed, and was also the main or only form in endothelial cells of larger blood vessels (Annunen *et al.* 1998).

2.2.2.2 β subunit

The ß subunit of prolyl 4-hydroxylase is identical to protein disulfide isomerase (PDI, EC 5.3.4.1) (Pihlajaniemi *et al.* 1987, Koivu *et al.* 1987). The function of the PDI is to catalyze the disulfide bond formation. Active PDI is a nonglycosylated homodimer, which has been cloned from a number of vertebrate species and from many plants, yeasts, and micro-organisms (Freedman 1989, Noiva 1994, Freedman *et al.* 1994, Freedman *et al.* 1995, Kivirikko & Myllyharju 1998, Kivirikko & Pihlajaniemi 1998). The processed human PDI polypeptide consists of 491 amino acids and is synthesized in a form having a signal peptide of 17 additional residues (Pihlajaniemi *et al.* 1987). The human gene coding for the PDI polypeptide is located in the q25 region of chromosome 17 (Pajunen *et al.* 1988, Popescu *et al.* 1988, Pajunen *et al.* 1991), and its transcribed part is about 16.5 kb in size and contains 11 exons (Tasanen *et al.* 1988). The promoter region of the human PDI gene contains six CCAAT elements, five potential SP 1 binding sites, and other regulatory elements, some of which are functionally redundant (Tasanen *et al.* 1992, Tasanen *et al.* 1993).

PDI is a modular protein consisting of domains *a, b, b´, a´,* and a C-terminal extension *c,* in this order (Edman *et al.* 1985, Pihlajaniemi *et al.* 1987). It has two -Cys-Gly-His-Cys- sequences, which represent independently acting catalytic sites for the isomerase activity (Hawkins & Freedman 1991, Vuori *et al.* 1992a, LaMantia & Lennarz 1993, Noiva *et al.* 1993). They are located in the *a* and *a´* domains, which are the most conserved regions of the polypeptide (Kivirikko & Myllyharju 1998, Kivirikko & Pihlajaniemi 1998), and are coded by exon numbers 1 and 2, and 8 and 9, respectively (Tasanen *et al.* 1988). The β subunit of prolyl 4-hydroxylase has PDI activity even when present in the native prolyl 4-hydroxylase tetramer (Koivu *et al.* 1987), but its role in prolyl 4-hydroxylase is not related to the disulfide isomerase activity. This has been demonstrated by the formation of a fully active prolyl 4-hydroxylase tetramer by a double mutant PDI polypeptide in which both -Cys-Gly-His–Cys- catalytic sites have been inactivated to -Ser-Gly-His-Cys- (Vuori *et al.* 1992c). The *a* and *a´* domains show significant sequence identity to thioredoxin, and NMR analysis has confirmed that the recombinant *a* and *a´* domains have a thioredoxin fold (Kemmink *et al.* 1997, Dijkstra *et al.* 1999). The homologous *b* and *b´* domains do not show significant sequence similarity to *a* and *a´*, but NMR analysis has revealed that the *b* domain also has a thioredoxin fold (Kemmink *et al.* 1999). Extension *c*, which is located in the C-terminus of the PDI polypeptide, has been assumed to bind Ca^{2+} ions (Lebeche *et al.* 1994). The C-terminal end of the polypeptide contains an ER retention signal, -Lys-Asp-Glu-Leu- (Pelham 1990, Vuori *et al.* 1992c). A critical PDI region for prolyl 4-hydroxylase assembly has been located in the C-terminus of domain *a´*, which cannot be replaced by the corresponding region of the related polypeptide ERp57 (Koivunen *et. al* 1996, Koivunen *et al.* 1999). The PDI *b´* and *a´* domains fulfill the minimum sequence requirement for the assembly of an active prolyl 4-hydroxylase tetramer in insect cell coexpression experiments, but the sequential addition of *b* and *a* domains increases markedly the level of enzyme activity (Pirneskoski *et al.* 2001).

When a prolyl 4-hydroxylase tetramer is dissociated (Kivirikko *et al.* 1992) or the α(I) subunit is expressed alone in insect cells (Vuori *et al.* 1992b), the α subunits form

insoluble, inactive aggregates. This indicates, that the function of the PDI polypeptide is to keep the α subunit in a soluble, nonaggregated conformation. However, when the prolyl 4-hydroxylase α subunit was expressed with the ER chaperone BiP in insect cells, soluble complexes were seen, but no prolyl 4-hydroxylase activity was observed (Veijola *et al.* 1996b). Therefore, the function of PDI in prolyl 4-hydroxylase is probably more specific, most likely to keep the α subunits in a catalytically active, nonaggregated conformation.

PDI has multiple functions in addition to catalyzing disulfide bond formation and serving as the β subunit of prolyl 4-hydroxylase. It acts as the smaller subunit of the triglyceride transfer protein dimer, in which its main function may also be to keep the other subunit in an active, nonaggregated conformation (Wetterau *et al.* 1990, Quan *et al.* 1995, Lamberg *et al.* 1996). PDI also acts as a chaperone that binds various peptides within the lumen of the ER (Noiva *et al.* 1999) and assists in the folding of many newly synthesized proteins (LaMantia & Lennarz 1993, Cai *et al.* 1994, Otsu *et al.* 1994, Rupp *et al.* 1994, Puig & Gilbert 1994, Hayano & Kikuchi 1995, Yao *et al.* 1997, Lamandé & Bateman 1999). It has been reported to be similar to the developmentally regulated retinal protein r-cognin (Krishna Rao & Hausman 1993), and to function as a major phosphoprotein of the ER (Quemeneur *et al.* 1994) that may undergo ATP-dependent autophosphorylation (Guthapfel *et al.* 1996). Other suggested functions are as a calcium binding protein (Lebeche *et al.* 1994) and a dehydroascorbate reductase (Wells *et al.* 1990), in addition to several less-well characterized functions (Kivirikko & Pihlajaniemi 1998).

2.2.3 Invertebrate prolyl 4-hydroxylases

Two genes coding for prolyl 4-hydroxylase α subunits, *phy-1* and *phy-2*, have been characterized in the *C. elegans* genome, and two potential PDI subunit encoding genes, called *pdi-1* and *pdi-2*, are also present (Veijola *et al.* 1994, Veijola *et al.* 1996a, Friedman *et al.* 2000, Hill *et al.* 2000, Winter & Page 2000). The processed PHY-1 and PHY-2 polypeptides are 544 and 525 amino acids long, respectively (Veijola *et al.* 1994, Friedman *et al.* 2000, Winter & Page 2000), and the sizes of the processed PDI-1 and PDI-2 polypeptides are 465 and 477 amino acids, respectively (Veijola *et al.* 1996a). Interesting differences are found in the *C. elegans* PDI polypeptides when compared to the human PDI. The sequence of the PDI-1 catalytic site in domain *a* is -Cys-Val-His-Cys instead of the usual -Cys-Gly-His-Cys, and the C-terminus of PDI-2 has the sequence –His-Thr-Glu-Leu, which is a functional variant of the –Lys-Asp-Glu-Leu retention signal present in the human PDI (Pihlajaniemi *et al.* 1987, Veijola *et al.* 1996a). The PHY-1 and PHY-2 polypeptides are 44-46 % identical to the human α(I) and α(II) subunits at the amino acid level, the highest degree of conservation being found within the C-terminal regions (Veijola *et al.* 1994, Hill *et al.* 2000, Friedman *et al.* 2000, Winter & Page 2000).
The PHY-1, PHY-2, PDI-1, and PDI-2 polypeptides are expressed in the collagen synthesizing hypodermal cells at times of maximal collagen synthesis (Page 1997, Winter & Page 2000). Unlike the vertebrate α subunits, the recombinant *C. elegans* α subunit PHY-1 has been shown to form an active αβ dimer in an insect cell coexpression system

with either the human PDI or the *C. elegans* PDI-2 polypeptide (Veijola *et al.* 1994, Veijola *et al.* 1996a). The *C. elegans* PHY-1 polypeptide has a 32 amino acid C-terminal extension, and its deletion prevents assembly of the enzyme dimer (Veijola *et al.* 1994). Interestingly, recombinant *C. elegans* PHY-2 did not become assembled into an active dimer with human PDI or *C. elegans* PDI-2, but a mixed PHY-1/PHY-2(PDI-2)$_2$ tetramer was formed when the three polypeptides were coexpressed in insect cells (Myllyharju J, Kukkola L, Winter AD & Page AP, unpublished observations). The mixed tetramer was found to be the main prolyl 4-hydroxylase form in *C. elegans* also *in vivo*.

Wild type *C. elegans* has a long and sinuous body, whereas the knock-out of *phy-1* results in a dumpy (*dpy)* phenotype (Friedman *et al.* 2000, Hill *et al.* 2000, Winter & Page 2000). Nematodes homozygous for a *phy-2* deletion were phenotypically wild type, suggesting that *phy-2* is not essential, although the 4-hydroxyproline content was slightly lowered in the *phy-2* deletion line when compared to the wild type (Friedman *et al.* 2000). The *phy-1/phy-2* double mutant was either embryonic lethal or resulted in a severe dumpy phenotype (Winter & Page 2000, Friedman *et al.* 2000). The dead embryos usually elongated at the twofold stage, but were unable to maintain their shape and often exploded (Friedman *et al.* 2000). Deletion of *pdi-2* has also been shown to lead to an embryonic lethal phenotype, which became evident after the first cuticle synthesis occurred (Winter & Page 2000).

The largest number of genes coding for prolyl 4-hydroxylase α subunit-like polypeptides is found in the *Drosophila* genome, which contains 19 putative prolyl 4-hydroxylase genes, ten of them being clustered in the chromosomal area 3R (Annunen *et al.* 1999, Adams *et al.* 2000, Abrams & Andrew 2002). Six of the prolyl 4-hydroxylase related genes are in the 74D1 part of the genome, whereas three of them are dispersed throughout other regions of the genome (Abrams & Andrew 2002). One recombinant *Drosophila* α subunit has been shown to form an $\alpha_2\beta_2$ tetramer with human and *Drosophila* PDI polypeptides (Annunen *et al.* 1999).

2.2.4 Catalytic properties

Prolyl 4-hydroxylase requires ferrous ions, 2-oxoglutarate, molecular oxygen, and ascorbate, and the K_m values for the cosubstrates between various enzyme forms studied are very similar (Helaakoski *et al.* 1995, Veijola *et al.* 1996a, Annunen *et al.* 1997, Kivirikko & Myllyharju 1998, Kivirikko & Pihlajaniemi 1998). Poly(L-proline) is a highly effective competitive inhibitor of the vertebrate type I prolyl 4-hydroxylase, the K_i of a poly(L-proline) with a molecular weight of 1500 being 0.02 μM. However, it is only a weak inhibitor of the type II prolyl 4-hydroxylase, the K_i being 1 000 times higher than that for the type I enzyme. The catalytic properties of the *C. elegans* PHY-1/PDI-2 dimer (Veijola *et al.* 1994, Veijola *et al.* 1996a) and the mixed tetramer (Myllyharju J, Kukkola L, Winter AD & Page AP, unpublished obsevations) are similar to those of the vertebrate type II enzyme in that they are relatively insensitive to inhibition by poly(L-proline). The K_i values for poly(L-proline) of both vertebrate isoenzymes decrease with increasing chain length (Helaakoski *et al.* 1995, Annunen *et al.* 1997). The K_m values of the human type II isoenzyme for three peptide substrates, a collagen like synthetic peptide, an elastin

like synthetic peptide, and a biologically prepared collagen substrate are 3-6 fold higher than those obtained with the type I enzyme. These findings suggest that there are some differences in the peptide binding sites between the isoenzymes (Annunen *et al.* 1997).

Photoaffinity labelling experiments with analogs of 2-oxoglutarate (de Waal *et al.* 1987) and the peptide substrate (de Waal *et al.* 1985), and also binding studies with suicide inhibitors acting as analogs of 2-oxoglutarate (Günzler *et al.* 1987), indicate that most, if not all, of the sequences contributing to the catalytic site are located in the α subunit. There are only a few observations suggesting that some catalytically important sequences may be located in the PDI subunit. Some inhibitory monoclonal antibodies were found to react with the PDI subunit (Höyhtyä *et al.* 1984, Bai *et al.* 1986). Moreover, the closely related 2-oxoglutarate dependent dioxygenase, lysyl hydroxylase, is a homodimer without PDI, suggesting that PDI does not participate in catalyzing the hydroxylation reaction (Turpeenniemi-Hujanen *et al.* 1980, Turpeenniemi-Hujanen *et al.* 1981, Myllylä *et al.* 1988).

During hydroxylation, the 2-oxoglutarate is stoichiometrically decarboxylated, with one atom of the O_2 molecule being incorporated into the succinate and the other into the hydroxy group formed on the proline residue (Fig. 1A) (Rhoads & Udenfriend 1968, Kivirikko & Myllyharju 1998, Kivirikko & Pihlajaniemi 1998). Ascorbate is not needed in the hydroxylation reaction, and prolyl 4–hydroxylase can hydroxylate proline residues at a maximal rate for a number of reaction cycles in the absence of ascorbate (Tuderman *et al.* 1977, Myllylä *et al.* 1978), but eventually the hydroxylation ceases. Prolyl 4-hydroxylase catalyzes the decarboxylation of 2-oxoglutarate (Fig. 1B and 1C) in an uncoupled way, i.e. without hydroxylation, in the absence of the peptide substrate, and at the low rate even in the presence of a saturating concentration of the substrate (Tuderman *et al.* 1977, Counts *et al.* 1978). In the uncoupled reaction cycles, the reactive ferryl ion is converted to $Fe^{3+}{\cdot}O^{-}$, making the enzyme unavailable for subsequent reaction cycles until reduced by ascorbate (De Jong *et al.* 1982). The biological function of ascorbate in the hydroxylation reaction is thus most likely to act as an alternative oxygen acceptor in the uncoupled reaction cycles, in which it is consumed stoichiometrically (Myllylä *et al.* 1984, DeJong & Kemp 1984).

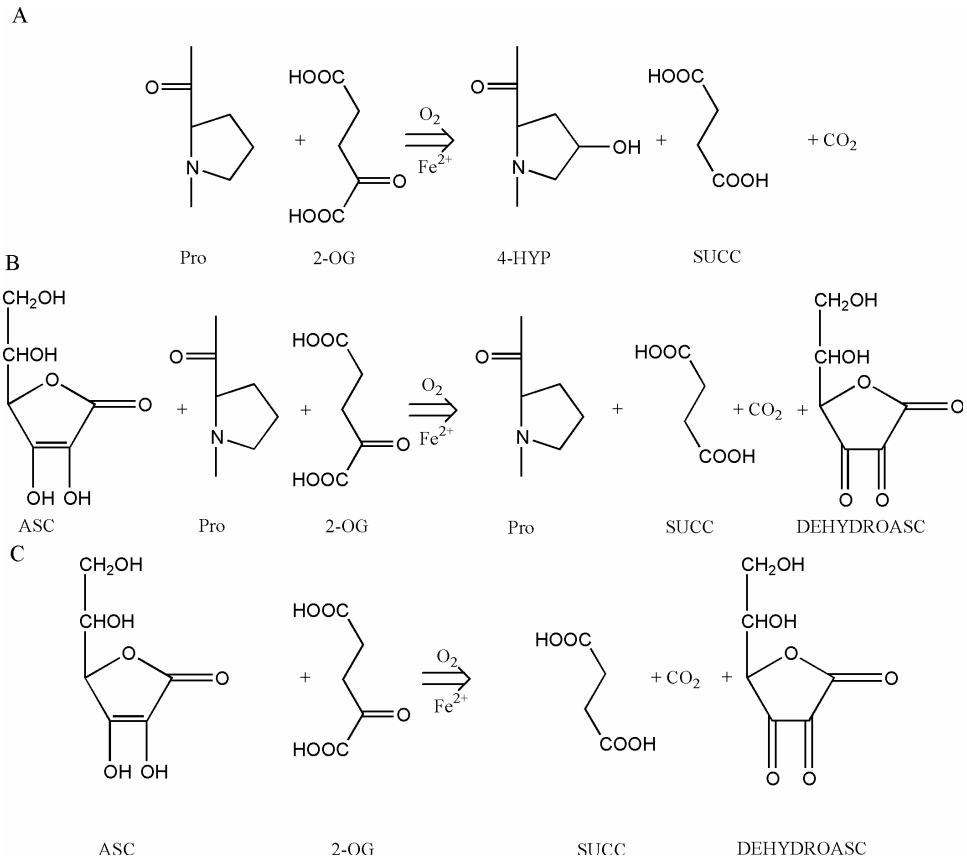

Fig. 1. Schematic representation of the reactions catalyzed by prolyl 4-hydroxylases. The 2-oxoglutarate is stoichiometrically decarboxylated during the hydroxylation reaction, which does not need ascorbate (A). Ascorbate is stoichiometrically consumed in the uncoupled reaction cycles, which may occur either in the presence (B) or absence (C) of the peptide substrate.

The Fe^{2+} is assumed to be located in a pocket coordinated with the prolyl 4–hydroxylase α subunit by three side chains (Hanauske-Abel & Günzler 1982). Site-directed mutagenesis data on the human α(I) subunit demonstrated that the His412, Asp414, and His483 residues are its iron-binding ligands (Lamberg *et al.* 1995, Myllyharju & Kivirikko 1997). Mutation of any these residues to serine or alanine inactivates the enzyme completely, with no inhibition of tetramer assembly (Lamberg *et al.* 1995, Myllyharju & Kivirikko 1997). The crystal structures of related dioxygenases isopenicillin N-synthase (IPNS), deacetoxychephalosporin synthase (DAOCS), clavaminic acid synthase (CAS), and proline 3-hydroxylase (P3H) have been resolved (Roach *et. al*, 1995b, Roach *et al.* 1997, Valegård *et al.* 1998, Lloid *et al.* 1999, Zhang *et al.* 2000, Clifton *et al.* 2001). Except for proline 3-hydroxylase, these enzymes catalyze oxidative steps in the biosynthesis of bicyclic β lactams. The metal binding amino acids in these enzymes are also two histidines and a negatively charged residue in position +2

relative to the first histidine (Roach *et. al* 1995b, Mori *et al.* 1996, Mori *et al.* 1997, Roach *et al.* 1997, Valegård *et al.* 1998, Lloid *et al.* 1999, Zhang *et al.* 2000, Clifton *et al.* 2001).

The 2-oxoglutarate binding site of prolyl 4-hydroxylase can be divided into three subsites. Subsite I is a positively charged side chain of the α subunit that ionically binds the C5 carboxyl group of 2-oxoglutarate, subsite II consists of two *cis*-positioned equatorial coordination sites of the enzyme-bound Fe^{2+} and is chelated by the C1-C2 moiety. Subsite III involves a hydrophobic binding site in the C3-C4 region of the cosubstrate (Hanauske-Abel & Günzler 1982, Majamaa *et al.* 1984, Hanauske-Abel 1991). Site-directed mutagenesis studies (Myllyharju & Kivirikko 1997) show that the residue forming subsite I in the human α(I) subunit is Lys493. Crystal structures of P3H, CAS, and DAOCS show that the corresponding residue in these enzymes is an arginine (Valegård *et al.* 1998, Zhang *et al.* 2000, Clifton *et al.* 2001). His501 is an additional critical residue in the catalytic site of type I prolyl 4-hydroxylase (Myllyharju & Kivirikko 1997). Detailed studies on various His501 mutants indicate that it probably has two roles; it directs the orientation of the C1 carboxyl group of 2-oxoglutarate to the active iron center and accelerates the breakdown of a tetrahedral ferryl intermediate to succinate, CO_2, and ferryl iron (Myllyharju & Kivirikko 1997).

The peptide substrate-binding domain of prolyl 4-hydroxylase has been shown to be distinct from the catalytic region (Myllyharju & Kivirikko 1999). Limited proteolysis experiments and site-directed mutagenesis studies on human type I prolyl 4-hydroxylase indicate that the peptide-binding domain is located approximately between the residues 140 and 240 of the α(I) subunit (Myllyharju & Kivirikko 1999).

2.2.5 Structure of the human α(I) subunit gene

The human gene for the α(I) subunit is located in the q21.3-23.1 region of chromosome 10 (Pajunen *et al.* 1989). The gene for type XIII collagen is located in the same chromosomal region (Horelli-Kuitunen *et al.* 1997). The distance between COL13A1 and the α(I) subunit genes is about 550 kb, and therefore, it is possible that several genes are located between them.

The size of the α(I) subunit gene is more than 69 kb, and the gene consists of 16 exons covering only 4% of the whole length of the gene (Helaakoski *et al.* 1994). The sizes of exons that consist entirely of protein coding sequences vary from 54 to 240 bp, and the sizes of introns vary from 750 to more than 16 000 bp (Helaakoski *et al.* 1994). Most of the introns are less than 4 kb in size, five of them being over 6 kb. The gene contains a classic polyadenylation signal 455 nucleotides downstream from the stop codon for translation, but no mRNA of this size was found in Northern blots (Helaakoski *et al.* 1994). However, the genomic clone contains an AATCAA and an AATACT sequence 1184 bp and 1218 bp after the stop codon. These sequences were shown to be utilized as polyadenylation sites (Helaakoski *et al.* 1994).

According to S1 nuclease protection analysis, the major transcription initiation site of the α(I) subunit gene is located 133 bp upstream of the translation initiation site. The 5´ flanking region of the gene contains a TATAA motif at position –29 relative to the

transcription initiation site, three variants of an AP-2 site, and one binding site for the transcription factor PuF. In addition, binding sites for the transcription factors AP-2, PEA-3, and SP-1 are found in the 5´ untranslated sequence of exon 1 (Helaakoski *et al.* 1994). The human α(I) subunit gene contains repetitive *Alu* sequences in the 5´ untranslated region and in the 3´ untranslated region in close proximity to the two putative polyadenylation sites. *Alu* sequences have also been found in introns of the human lysyl hydroxylase 1 and 3 genes (Heikkinen *et al.* 1994, Rautavuoma *et al.* 2000). An *Alu*-mediated homologous recombination in the lysyl hydroxylase 1 gene is the most commonly found mutation in Ehlers-Danlos syndrome (EDS) type VI, a disease that is attributed to lysyl hydroxylase deficiency (Heikkinen *et al.* 1994, Pousi *et al.* 1994, Heikkinen *et al.* 1997).

The RNA transcripts of the human (Helaakoski *et al.* 1989, Helaakoski *et al.* 1994) and mouse (Helaakoski *et al.* 1995) α(I) genes are alternatively spliced. Two types of mRNA are generated from the α(I) subunit genes, resulting from the mutually exclusive alternative splicing of the two consecutive, homologous 71 bp exons 9 and 10 in these genes. The alternatively spliced exons are identical in their first 5 bp, and the overall identity between them is 61% at the nucleotide level and 58% at the amino acid level. The biological significance of these two forms is unknown, but the mRNA containing exon 10 is the main form in brain, while both isoforms are abundantly expressed in placenta, gut, and skin (Helaakoski *et al.* 1994). Insect cell expression studies have shown that both recombinant α(I) subunit forms are assembled into active tetramers with no differences in their catalytic properties (Vuori *et al.* 1992b).

2.2.6 Gene regulation of prolyl 4-hydroxylase

The main function of prolyl 4-hydroxylase is to produce stable triple-helical collagen, and therefore a good correlation is found between the rate of collagen synthesis and the amount of active enzyme tetramer (Kivirikko *et al.* 1992). In most cells and tissues, the PDI polypeptide is produced in a 10-100 fold excess over the α(I) subunit, and the regulation of the active enzyme tetramer thus mainly occurs through regulation of the amount of the α subunits (Kivirikko *et al.* 1992). Several papers have reported that high lactate levels and low oxygen tension can stimulate prolyl 4-hydroxylase activity (Stassen *et al.* 1973, Kivirikko & Myllylä 1980). The mRNA levels of the prolyl 4-hydroxylase α(I) subunit have been shown to increase 2-3 fold after an 8 h exposure to hypoxia, and return to basal level after reoxygenation, indicating that the α(I) subunit gene is a target for the hypoxia-inducible transcription factor-1α (Takahashi *et al.* 2000). Under normal physiological conditions, the level of procollagen production may be affected by the availability of cofactors for prolyl 4-hydroxylase (Myllylä *et al.* 1977, Myllylä *et al.* 1978, Kivirikko *et al.* 1992).

Cells seem to have a mechanism by which they recognize unhydroxylated nonhelical collagen molecules and rapidly degrade them in the intracellular space (Breul *et al.* 1980, Berg *et al.* 1980, McLauhglin & Bulleid 1998). The molecular mechanism leading to this retention has been postulated to be mediated either by prolyl 4-hydroxylase itself (Kao *et al.* 1979, Walmsley *et al.* 1999) or by collagen binding protein HSP47 (Nagata 1996,

Nagata 1998, Tasab *et al.* 2000). Also, BiP has been shown to retain mutant collagen chains in the ER (Chessler & Byers 1992).

The half-life of recombinant prolyl 4-hydroxylase in the yeast *Pichia pastoris* has been shown to be significantly increased by simultaneous collagen expression (Vuorela *et al.* 1997). The most likely explanation for the increased half life is that in the presence of collagen the enzyme tetramer is stabilized, whereas in the absence of collagen synthesis, the tetramers are rapidly dissociated back to inactive monomeric subunits and the α subunits form insoluble aggregates (Vuorela *et al.* 1997, Vuorela *et al.* 1999). In increasing cell density, the mRNA levels of prolyl 4-hydroxylase α subunits have been shown to be decreased, whereas the amount of prolyl 4-hydroxylase protein is increased, apparently because of an increased stabilization of the enzyme tetramer (Lee *et al.* 2001). Before this finding, it was already known that the prolyl 4-hydroxylase tetramer has a half-life of 2-3 days in the presence of collagen synthesis (Chichester *et al.* 1976, Majamaa *et al.* 1979, Berg *et al.* 1980), but the half-life was only 1-2 h when L-929 fibroblasts began to divide and ceased synthesizing collagen (Hebda *et al.* 1983). Similarly, when differentiation of mouse F9 teratocarcinoma stem cells occurs, the mRNA levels of the α and β subunits and the amount of prolyl 4-hydroxylase are markedly increased simultaneously with increased synthesis of basement membrane components, especially collagen IV (Roguska & Gudas 1985, Helaakoski *et al.* 1990).

2.3 A family of prolyl 4-hydroxylases that modify HIF

In hypoxic cells, the up-regulation of many genes depends on the activation of the family of hypoxia-inducible transcription factors (HIFs), which are expressed in the cells and tissues of mammals, flies, and worms (Nambu *et al.* 1996, Bacon *et al.* 1998, Bruick *et al.* 2001, Epstein *et al.* 2001, Ivan *et al.* 2001, Jaakkola *et al.* 2001). Their target genes and the encoded proteins are involved *e.g.* in anaerobic energy production, vascularization of the embryo, and tumour angiogenesis (Sutter *et al.* 2000, Semenza 2001, Yu *et al.* 2001, Semenza 2002). At least two distinct α subunits of HIFs, HIF-1α and HIF-2α, are found in man (Epstein *et al.* 2001, Ivan *et al.* 2001, Jaakkola *et al.* 2001, Bruick & McKnight 2001). The HIFs share the same β subunit, which is commonly called ARNT (aryl hydrocarbon nuclear translocator) and is expressed independently of oxygen tension. When intracellular oxygen tension is normal, the newly synthesized HIFα subunits are modified by an oxygen-dependent prolyl 4-hydroxylase (Ivan *et al.* 2001, Jaakkola *et al.* 2001). Hydroxylation enables binding of the tumour suppressor von Hippel-Lindau protein (pVHL) (Ivan *et al.* 2001, Jaakkola *et al.* 2001), which then targets this transcription factor for proteasomal degradation (Cockman *et al.* 2000, Ohh *et al.* 2000). In normoxic cells, the HIFα subunits have an exceptionally short half life (Jewell *et al.* 2001), whereas in hypoxic cells HIFα is not hydroxylated, escapes degradation, and forms a stable heterodimer with HIFβ. The HIFαβ heterodimer then translocates into the nucleus, where it binds to hypoxia-response elements in genes that are switched on by hypoxia.

Two separate domains within HIF-1α and HIF-2α respond to the hypoxia signaling pathway. The first is an oxygen dependent degradation domain (ODDD), in which the

two proline residues subject to hydroxylation are located in conserved –Leu-X-X-Leu-Ala-Pro- core motifs. The ODDD also contains a binding site for pVHL (Huang *et al.* 1998, Sutter *et al.* 2000, Hoffman *et al.* 2001, Semenza 2002). A second hypoxia sensing region, which functions as a hypoxia-inducible transactivation domain (CAD), lies within the C-terminal part of HIF-1α and HIF-2α and contains an asparagine residue that is hydroxylated by an asparaginyl hydroxylase (Bruick & McKnight 2002, Lando *et al.* 2002).

Three human HIF prolyl 4-hydroxylases (HPH) have been identified and cloned and shown to hydroxylate the HIFα subunit by a mechanism dependent on oxygen concentration (Bruick & McKnight 2001, Epstein *et al.* 2001). The HPHs are distinct from the collagen prolyl 4-hydroxylases. Firstly, as described above, their substrate is different, *i.e.* the hydroxylated proline is present in a –Leu-X-X-Leu-Ala-Pro- sequence (Ivan *et al.* 2001, Jaakkola *et al.* 2001). Secondly, the collagen prolyl 4-hydroxylases are located in the ER, whereas the HIF prolyl 4-hydroxylases are apparently cytoplasmic monomers (Bruick & McKnight 2001, Epstein *et al.* 2001). Thirdly, the type I and II collagen prolyl 4-hydroxylases do not hydroxylate a synthetic HIF polypeptide (Jaakkola *et al.* 2001).

Hydroxylation of Pro564 and Pro402 residues is the key modification controlling the binding of pVHL to the C-terminal target site of HIF-1α subunits (Bruick & McKnight 2001, Epstein *et al.* 2001, Ivan *et al.* 2001, Jaakkola *et al.* 2001). The amino acids in close proximity to these proline residues have been shown to influence HIF-1α hydroxylation and pVHL binding. The mutation of Leu562 to Ala, Ala563 to Gly, or Tyr565 to Ala inhibits modification by HPH-1 and HPH-2. Interestingly, when Pro567 was mutated to Gly, the HIF-1α polypeptide served as an equal or even better substrate for HPH enzymes (Bruick & McKnight 2001, Ivan *et al.* 2001, Jaakkola *et al.* 2001, Yu *et al.* 2001). The HPH-2 and HPH-3 enzymes hydroxylate both Pro564 and Pro402, whereas HPH-1 acts only on Pro564 (Epstein *et al.* 2001).

The *C. elegans* HPH was first identified on the basis of an abnormal, though viable, egg laying (*egl-9*) phenotype (Epstein *et al.* 2001). Complete loss of *C. elegans* HIF regulation in *egl-9* mutants suggests that there may be only one HPH in the *C. elegans* genome. The e*gl-9* mutants constitutively expressed HIF-1α, but formation of a complex between pVHL and HIF-1α was rescued only by *in vitro* translated EGL-9 (Epstein *et al.* 2001). A HIF prolyl 4-hydroxylase (HPH) is also encoded in the genome of *D. melanogaster*, and blocking it by the RNAi technique upregulates hypoxia-inducible genes, such as lactate dehydrogenase (Bruick & McKnight 2001, Epstein *et al.* 2001).

The human HIFα prolyl 4-hydroxylases require the same cofactors and respond to the same inhibitors as collagen prolyl 4-hydroxylases. Complete inhibition of HPH activity was observed with N-oxalylglycine, which is a competitive inhibitor with respect to 2-oxoglutarate (Jaakkola *et al.* 2001). Iron chelators and cobaltous ions are also capable of inhibiting HPH (Bruick & McKnight 2001, Epstein *et al.* 2001, Jaakkola *et al.* 2001). In human type I prolyl 4-hydroxylase, the iron-binding ligands are His412, Asp414, and His483 (Lamberg *et al.* 1995, Myllyharju & Kivirikko 1997). Site directed mutagenesis studies showed that conversion of the HPH-1 His135, Asp137, and His196 residues to Ala totally eliminated its activity (Bruick & McKnight 2001). In collagen prolyl 4-hydroxylases the residue forming the subsite I for 2-oxoglutarate binding is Lys493 in the human α(I) subunit in position +10 with respect to the second iron binding His

(Myllyharju & Kivirikko 1997). As shown in Fig. 2., the corresponding basic amino acid in HPHs is an arginine at position +9 (Bruick & McKnight 2001, Epstein *et al.* 2001).

Fig. 2. Comparison of sequences of the catalytic regions of human (HuHPH), D*rosophila* (DmHPH), and *C. elegans* (EGL-9) HIF prolyl 4-hydroxylases. The Fe^{2+} binding residues are indicated by an asterix (*) and the 2-oxoglutarate binding residue is indicated by a diamond (◊), respectively.

2.4 Prolyl 3-hydroxylase

Prolyl 3-hydroxylase catalyzes hydroxylation of proline residues in the -Pro-4Hyp-Gly- sequences in vertebrates (Kivirikko *et al.* 1992). 3-Hydroxyproline has also been identified in the sequences –Val-3Hyp-Asp-, -Val-3Hyp-Glu-, and –Tyr-3Hyp-Tyr- of the secreted cathepsin L-like proteinases of the trematode *Fasciola hepatica* (Wijffels *et al.* 1994). As these sequences are completely different than these around 3-hydroxyproline in collagens, it seems likely that this trematode prolyl 3-hydroxylase is quite different from those collagen prolyl 3-hydroxylase (Wijffels *et al.* 1994). The structure of prolyl 3-hydroxylase is unknown, and it has not yet been cloned, but its molecular weight is about 160 000 Da in gel filtration (Tryggvason *et al.* 1979, Kivirikko *et al.* 1992).

2.5 The nematode *Caenorhabditis elegans*

The *Caenorhabditis elegans* lives in the soil in many parts of the world, where it survives by feeding microbes such as bacteria. It is of no economic importance to man. This nematode was the first eukaryote with a completely sequenced genome (Hodgkin *et al.* 1995, Hodgkin & Herman 1998, *C. elegans* sequencing consortium 1998, Hodgkin 1999). The adult nematode is about 1 mm in length and consists of a tube, the exterior cuticle, which contains two smaller tubes, the pharynx and gut, and the reproductive

system (Fig. 3.). Most of the volume of the animal is taken up by the reproductive system. *C. elegans* has two sexes: a self-fertilizing hermaphrodite and a male, which arises by nondisjunction and permits sexual reproduction as well. The shape of the worm is maintained by internal hydrostatic pressure. A hermaphrodite that has not mated lays about 300 eggs during its reproductive life span. Of the 959 somatic cells of the hermaphrodite, some 300 are neurons, which mediate responses to taste, smell, temperature, and touch. There is also evidence that the nematode is capable of rudimentary learning (Wood 1988).

Juvenile worms hatch and develop through four larval stages, which are punctuated by molts. The nematode cuticle is first synthesized in the egg prior to hatching and subsequently at the end of each larval stage prior to molting. When the food supply is limited during larval development, *C. elegans* can take an alternative developmental pathway at the L2/L3 molt and produce a dauer larva, a specialized L3 stage that does not feed, is resistant to desiccation, and can survive up to 3 months without further development. If food becomes available, the dauer larva molts to become an L4 larva, which resumes normal development. The animals are usually grown on petri dishes seeded with bacteria as a food supply. They reproduce with a life cycle of about 3 days under optimal conditions, and have an average life span of 2-3 weeks (Sulston & Hodgkin 1988, Wood 1988).

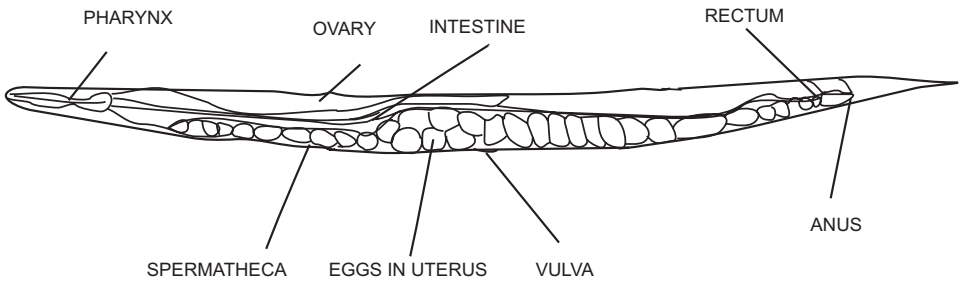

Fig. 3. Structure of an adult *C. elegans* hermaphrodite.

The 97 megabase genome of *C. elegans* contains 19 000 genes (Blumenthal & Steward 1997, *C. elegans* sequencing consortium 1998). The genes are carried on five autosomes and a sex chromosome (X). Hermaphrodites are diploid for all chromosomes, whereas males are diploid for the autosomes but have only one X chromosome (Sulston & Hodgkin 1983, Hodgkin 1999). Nematodes are transparent throughout their life cycle, and thus their development can be followed at the cellular level by light microscopy. *C. elegans* is also useful for biological studies due to its simple structure, ease of cultivation, short life cycle, suitability for genetic analysis, and small genome size (Wood *et al.* 1988).

2.5.1 C. elegans collagens

The cuticle collagen genes form one of the largest and possibly oldest nematode specific gene families (Johnstone 1994). The cuticula of *C. elegans* is mainly composed of small, collagen-like proteins encoded by a multi-gene family of approximately 160 members (Kramer 1997, Johnstone 2000). Fibrillar collagens, elastin, and fibronectin are absent from the *C. elegans* cuticle, suggesting that these genes were elaborated within the vertebrate lineage (Hutter *et al.* 2000). Despite the large number of collagen genes, only the basement membrane collagens type IV and XVIII are conserved between human and *C. elegans* (Kramer 1997, Hutter *et al.* 2000, Ackley *et al.* 2001).

C. elegans cuticle collagen molecules are synthesized by hypodermal epithelial cells. Since more than 80% of the cuticle proteins are collagens, *C. elegans* expresses these molecules throughout its life cycle, but the amounts and types of collagen expressed vary greatly according to the developmental stage (Cox *et al.* 1981, Cox *et al.* 1989, Kramer 1994a, Kramer 1994b). In general, the expression of cuticle collagens increases during the four larval molts, when the cuticle is synthesized (Hirsh *et al.* 1985, Hirsh *et al.* 1987, Cox *et al.* 1989). In addition, several collagen encoding genes are expressed only in dauer larvae (Kramer *et al.* 1985).

C. elegans has been used as a model organism to study the consequences of collagen mutations. The phenotypes caused by these mutations fall predominantly into three categories: blisters, blistering of the cuticular material; dumpy, shortening of the body length of the animal; and roller, a helical twisting of the body (Johnstone 2000). In addition, mutation of the NC1 domain of the *C. elegans* type XVIII collagen homolog, *cle-1*, affects cell migration and axon guidance (Ackley *et al.* 2001). The type IV collagen underlies the body wall muscles and surrounds the pharynx, intestine, nerves, and gonad. The cells expressing type IV collagen in *C. elegans* are generally mesodermal, but type IV collagen also assembles on tissues that do not express it (Graham *et al.* 1997). Two type IV collagen genes have been identified in *C. elegans*, and their protein products always co-localize. Mutations in either of the two genes can arrest embryonic development. The most severe phenotype arrests at the twofold stage of embryonic development, and the mildest phenotype is nearly wild type at 15-20 °C (Guo *et al.* 1991, Sibley *et al.* 1993, Gupta *et al.* 1997). In the absence of the *C. elegans* lysyl hydroxylase (Let-268), type IV collagen is not found in the basement membrane, but is retained within the type IV collagen producing cells. The homozygous *let-268* mutant nematodes fail to complete embryogenesis and arrest at the twofold stage shortly after elongation commences (Norman & Moerman 2000). It has been shown that body wall muscle function is required for elongation of the *C. elegans* embryo beyond the twofold stage, and obviously the disruption of the body wall muscles is the cause of the embryonic arrest at this stage (Waterston 1989, Barstead & Waterston 1991, Williams & Waterston 1994).

2.5.2 Composition of the C. elegans egg shell

The egg is generally the most resistant stage in the nematode life cycle, and it consists of three layers, an outer vitelline layer, middle chitinous layer, and an inner lipid layer (Bird & Bird 1991). The inner two layers are both formed by secretion from the egg cytoplasm after fertilization, whereas the vitelline layer is derived from the oolemma. An important function of the chitinous layer is to provide structural strength to the egg shell (Bird & Bird 1991).

The egg shell itself is a structure distinct from the embryo, and there is a space between the inner lipid layer of the egg shell and the plasma membrane of the embryo (Rappleye *et al.* 1999). The inner lipid layer of the egg shell also contains collagenous molecules, and it is hydrolyzed before hatching, making the egg more permeable. At 20°C, *C. elegans* embryogenesis takes approximately 800 minutes from first division to hatching. Most cell divisions are completed during the first 350 minutes, after which the embryo begins to elongate from an ovoid ball into a tube (Sulston *et al.* 1983).

3 Outlines of the present study

The amino acid sequences of the collagen prolyl 4-hydroxylase catalytic α subunits are highly conserved among vertebrates and nonvertebrates. In this work, the structures of the human and mouse α(II) subunit genes were determined in order to lay a basis for future studies on prolyl 4-hydroxylase gene expression and regulation. Another purpose of the present research was to study by immunofluorescence techniques the expression of the two enzyme isoforms in different human tissues in order to obtain information about their potential tissue or collagen specificity.

The availability of the complete genomic sequence of *Caenorhabditis elegans* has enabled the identification of prolyl 4-hydroxylase isoenzymes in this organism. Two prolyl 4-hydroxylase α subunit isoforms have already been characterized from this nematode. The present thesis describes the identification of a third *C. elegans* prolyl 4-hydroxylase isoform and the study of its tissue distribution and molecular properties. As a model organism, *C. elegans* provides many possibilities for future studies of the prolyl 4-hydroxylase isoenzyme family.

The specific aims were:

1. To isolate and characterize the genes for the α(II) subunits of human and mouse type II prolyl 4-hydroxylases.

2. To study the expression of human type I and II prolyl 4-hydroxylases in foetal, mature, and malignant tissues.

3. To clone a third *C. elegans* prolyl 4-hydroxylase α subunit isoform and to study its function.

4 Materials and methods

Detailed descriptions of the materials and methods are presented in the original papers I-IV.

4.1 Isolation and characterization of genomic and cDNA clones (I-IV)

A human lung fibroblast genomic library (Stratagene), a human chromosome 5 specific library (ATCC), and P1 clones (Genome Systems Inc.) were screened for the human prolyl 4-hydroxylase α(II) subunit gene. The probe used for the screening of the libraries was a ^{32}P-labelled nick-translated α(II) subunit cDNA (Annunen *et al.* 1997). The libraries were plated and screened as described elsewhere (Sambrook *et al.* 1989). P1 Human Library screening services (Genome Systems Inc.) were used to find 5´ genomic clones for the gene. The P1 DNA was isolated according to the manufacturer's recommended protocol. A cosmid library from the 129 SvJ mouse strain (Stratagene) was used to screen for the mouse α(II) subunit gene. The cosmid DNA was bound to duplicate filters (Sambrook *et al.* 1989) and probed with a ^{32}P-labelled mouse α(II) cDNA (Helaakoski *et al.* 1995). BAC mouse library screening services (Genome Systems Inc.) were used to obtain 5´ genomic clones for the gene, and the BAC DNA was isolated using the QIAGEN Large Construct kit.

A *C. elegans* database search indicated the presence of an open reading frame showing a 23-30% amino acid sequence identity to the conserved C-terminal regions of human and *C. elegans* prolyl 4-hydroxylase α subunits. The full-length cDNA was cloned in two parts. Two cDNA fragments, one corresponding to the predicted open reading frame T20B3.7 that contains exons 2-5, and the other spanning the sequence from the start codon in the first exon to a *BamHI* site in the fourth exon, were amplified from the *C. elegans* mixed-stage cDNA library (Stratagene). The cDNAs were cloned into the pUC18 vector using the SureClone kit (Pharmacia). For the cloning of the full-length *phy-3* cDNA, the 3´ fragment (exons 2-5) was digested with *BamHI* and the 5´ fragment with *BamHI* and *EcoRI*, and the digested fragments were coligated into pBluescript (Stratagene).

4.2 Nucleotide sequencing and sequence analysis (I, III, IV)

The positive clones from the various human and mouse genomic libraries were isolated and analyzed by restriction enzyme mapping and Southern blotting. Suitable genomic fragments were subcloned into pUC18 (Pharmacia) or pBluescript (Stratagene) vectors for further analysis, and P1, BAC, and cosmid clones were sequenced directly.
DNA sequencing was performed on an automated sequencer (ABI Prism 377, Applied Biosystems) using BigDye or dRhodamine Terminator Cycle Sequencing Ready Reaction kits (PE Biosystems). The primers used for sequencing were synthesized in an Applied Biosystems DNA synthesizer at the Department of Biochemistry, University of Oulu or ordered from Lifetech (GIBCO).

DNASIS™ and PROSIS™ version 6.00 sequence analysis software (Amersham Biosciences) were used to analyze the sequence data. Sequence homology comparisons were carried out in the GenBank, EMBL, and Swiss-Prot databases using the BLAST service at the National Center of Biotechnology Information and the ClustalW service at the European Bioinformatics Institute. Consensus transcription factor binding sites were searched from the Transcription Factor Data Base using the Transcription Element Search Software version 3.2, MatInspector version 2.2, and Gene-Search (BCM search launcher and GenBank databases) programs.

4.3 Determination of the transcription start site and intron sizes of the human and mouse prolyl 4-hydroxylase α(II) subunit genes (I)

The transcription initiation sites were mapped by S1 nuclease protection analysis using the Multi-NPA RNA/DNA/Oligo Probe Protection Assay Kit (Ambion). Two probes corresponding to nucleotides –680 to +38 and –223 to +6 relative to the previously identified 5´ ends of the human (Annunen *et al.* 1997) and mouse (Helaakoski *et al.* 1995) α(II) subunit cDNAs, respectively, were generated by PCR and restriction enzyme digestion. The probes were 5´ end-labelled using a KinaseMax 5´ End-Labelling Kit (Ambion) and hybridized with total RNA from human placenta (Clontech) and mRNA from mouse placenta. The protected fragments were analyzed on 6% polyacrylamide sequencing gels, and their sizes were determined by comparison with adjacent dideoxynucleotide sequencing reactions (Sanger *et al.* 1977) and pUC19/*Sau3A* markers (Ambion).

The sizes of introns were determined by nucleotide sequencing, PCR, restriction enzyme mapping, and Southern blotting.

4.4 Semiquantative PCR assays (I, IV)

In order to study the alternative splicing of exons 12a and 12b in the human α(II) gene, PCR analysis of adult and foetal multiple tissue cDNA (MTC™) panels (Clontech) was performed using exon specific primers. The forward primer HGAII981f was from exon 7

and the reverse primers HGAIIAlt2r and HGAIIAlt1r from the 5′ ends of the exons 12a and 12b, respectively. The PCR products were electrophoresed on 1% agarose gels and the sequences of the PCR products were verified.

Reverse transcription followed by PCR (RT-PCR) was performed to detect the relative abundance of the *phy-3* transcripts at different developmental stages of *C. elegans* (L1, L2, L3, and L4 larval stages plus adult stage). The cDNA samples used as templates were from synchronous nematode cultures sampled at 2 h intervals from shortly after hatching to the young adult stage (Johnstone & Barry 1996). The abundance of the *phy-3* transcript was measured in relation to the constitutively expressed control gene *ama-1*, which codes for the large subunit of RNA polymerase II (Johnstone & Barry 1996). PCR products corresponding to *phy-3* and *ama-1* were simultaneously amplified, electrophoresed on 2% agarose gels, Southern blotted, and hybridized under stringent conditions with ^{32}P labelled probes. The hybridized bands were excised and counted by scintillation.

4.5 Expression of recombinant proteins in insect cells and cell culturing (I, II, III, IV)

For recombinant expression of the two alternatively spliced isoforms of the human type II prolyl 4-hydroxylase, the baculovirus expression vector pVL1392-α(II)hu (Annunen *et al.* 1997) was modified so that it contained the exon 12a sequence instead of the exon 12b sequence. The recombinant vector was cotransfected into *Spodoptera frugiperida* Sf9 insect cells (Invitrogen) with a modified *Autographa californica* nuclear polyhedrosis virus DNA using the BaculoGold transfection kit (PharMingen) (Crossen & Gruenwald 1998). Sf9 cells were cultured in TNM-FH medium (Sigma) supplemented with 10% foetal bovine serum (Gibco) at 27°C as monolayers. To produce enzyme tetramers, insect cells were coinfected with the α(II)Ex12a or the original α(II)hu (termed here α(II)Ex12b and the PDI (Vuori *et al.* 1992b) viruses in a 1:1 ratio. The cells were harvested 72 after infection, washed with PBS, homogenized in a Triton X-100 containing buffer, and centrifuged at 10 000 g for 20 min. The resulting supernatants were analyzed on 8% nondenaturing PAGE gels.

The full-length cDNA of *phy-3* was cloned into the pVL1392 vector (PharMingen), and the recombinant baculovirus was generated as above. Insect cells were infected with viruses coding for PHY-3, PHY-1 (Veijola *et al.* 1994), or the human α(I) subunit (Vuori *et al.* 1992b) alone or together with viruses coding for *C. elegans* PDI-1, PDI-2 (Veijola *et al.* 1994, Veijola *et al.* 1996a), or human PDI (Vuori *et al.* 1992b). Insect cells were harvested and homogenized as described above, and aliquots of the resulting supernatants and insoluble fractions were analyzed on 12% SDS-PAGE under reducing conditions. Western transfer analysis was performed with a polyclonal antibody 2681.2 generated against a synthetic PHY-3 peptide.

Prolyl 4-hydroxylase activities of the Triton X-100 soluble fractions of the insect cell homogenates were assayed by a method based on the hydroxylation coupled decarboxylation of 2-oxo-[1-^{14}C] glutarate (Kivirikko & Myllylä 1982). An immortalized hepatocyte cell line HepG2 and a human umbilical vein endothelial cell line EA were grown in glass plates under standard conditions. Rat osteosarcoma cells (ATCC ROS

17/2.8) were cultured in Alpha Mem medium with ribonucletides and deoxyribonucletides (Gibco) under standard conditions.

4.6 Generation of α(I) and α(II) subunit antibodies and confirmation of their specificity (II, III, IV)

Two new antibodies, one polyclonal and one monoclonal, were generated for the present work. A polyclonal rabbit antibody against the human α(I) subunit (R17) was raised by injecting rabbits subcutaneously at several sites with 60-100 μg of the purified and denaturated recombinant α(I) polypeptide. The sera was analyzed with enzyme-linked immunoassay (EIA) using the native type I and II prolyl 4-hydroxylases as antigens according to the instructions of the manufacturer of the EIA-kit (Vectastain ABC, Vector Laboratories). To purify the α(I) antibody, 20 mg of recombinant type I prolyl 4-hydroxylase was coupled to epoxy-activated Sepharose A 6B according to manufacturer's protocol (Amersham Pharmacia). A type II prolyl 4-hydroxylase column was used as a negative affinity purification step for the α(I) antibody. The purified antibody was concentrated to 0.5 mg/ml using Microsep microconcentrators (Filtron Technology). A monoclonal antibody against the human α(II) subunit (M14) was generated by immunizing mice with a purified and denatured recombinant α(II) polypeptide. The α(II) antibody was used either in a form of nonpurified medium or was purified with Sepharose A. No difference was observed between the results obtained with the purified and nonpurified α(II) antibodies. The specificity of the antibodies was confirmed by Western blotting and antigen blocking. Protein concentrations were estimated with the Bio-Rad protein assay kit (Bio-Rad).

4.7 Immunofluorescence staining and immunoelectron microscopy (II, III, IV)

For immunofluorescence, tissue cryosections and cells cultured on glass coverslips were fixed in precooled methanol. Nonspecific staining was blocked by incubating the slides in 1% BSA-PBS, after which the samples were subjected to the primary antibody incubations for 1 h or overnight. After exhaustive washing in PBS, the TRITC (tetramethylrhodamine isothiocyanate)-conjugated secondary antibodies were applied and the samples were incubated in the dark in a moist chamber, washed again, mounted with Glycergel (DAKO), and studied with a Leitz Aristoplan microscope equipped with filters optimized for TRITC epifluorescence. Representative fields were photographed on Kodak Ektachrome Elite II 400 ASA diapositive film. Frozen sections were also stained with haematoxylin and eosin for more accurate histological evaluation.

For the nematode whole mount immunofluorescence staining, the wild type and *phy-3*$^{-/-}$ nematodes were washed thoroughly with ice-cold PBS, pipetted onto poly-L-lysine coated slides, permeabilized by freeze-cracking, and fixed with methanol and acetone (Rogalski *et al.* 1993). The slides were incubated with the polyclonal PHY-3 antibody

2681, washed with PBS, incubated with an Alexafluor 546-conjugated goat anti-rabbit antibody (Molecular Probes), washed again, and examined under an epifluorescence microscope.

For immunoelectron microscopy, the tissue samples were fixed in 4% paraformaldehyde or immersed in sucrose and frozen in liquid nitrogen, and ultrathin cryosections were cut with a Leica Ultracut UCT microtome. The sections were incubated in 5% BSA and 0.1% gelatin in PBS. For the single immunolabeling experiment, sections were then incubated with the M14 α(II) antibody, washed, and exposed to rabbit anti-mouse IgG (Zymed Laboratories) followed by a protein A-gold complex. For the double immunolabeling experiment, the sections were first exposed to the primary cross-reactive K4 α(II) antibody (Annunen *et al.* 1998) followed by rabbit anti-mouse IgG and the protein A-gold complex (size 10 nm). The sections were then incubated with a second monoclonal antibody to type IV collagen (DAKO) followed by rabbit anti-mouse IgG and a protein A-gold complex (size 5 nm). The sections were examined in a Philips CM100 transmission electron microscope.

Table 2. The polyclonal and monoclonal antibodies used in this work.

Antibody	Manufacturer
Rabbit non immune isotype IgG	DAKO
Osteocalcin	BioDesign
β-subunit	DAKO
CD 34	Novocastra laboratories ltd
Gb200	NeoMarkers
α-Sma	NeoMarkers
Col I	BioDesign
Col II	BioDesign
Col IV	DAKO
SAR-TRITC (swine anti-rabbit)	DAKO
RAM-TRITC (rabbit anti-mouse)	DAKO
PHY-3	
α(I)	
α(II)	

4.8 Generation of the *phy-3* reporter gene construct and germ line transformation (IV)

The promoterless nucleus-localized vector pPD96.04 was used for the construction of a *phy-3*/*lacZ* reporter gene fusion vector. A DNA fragment from -1480 to +6 relative to the ATG translation start codon was amplified from the *phy-3* genomic clone and ligated into pPD96.04, allowing in-frame translational fusion of the *phy-3* promoter sequence and GFP/*lacZ* reporter genes.

Transgenic nematode strains were generated as previously described by microinjection into the gonad of adult nematodes (Mello *et al.* 1991). Two independent lines of

transformed animals were maintained and inspected for reporter gene expression as described previously (Fire *et al.* 1990), except that slides were stained at 22°C with a solution containing 0.03% X-gal or at 37°C with a solution containing 0.3% X-gal (sensitive staining). The GFP expression levels of this construct were too low to be detected *in vivo* fluorescence studies. The stained nematodes were viewed under Nomarski optics using a Zeiss Axoscop 2 microscope.

4.9 Isolation and analysis of *phy-3*$^{-/-}$ deletion mutants (IV)

Deletion mutants of *phy-3* were made by Dr. Gary Moulder at the Oklahoma Medical Research Foundation as part of the *C. elegans* gene knockout consortium. Genomic DNA was first isolated from several populations of mutagenized nematodes to identify a population with animals carrying deletions in *phy-3*. These populations were subdivided until homozygotes *phy-3*$^{-/-}$ nematodes were obtained. Each deletion mutant was outcrossed to wild type N2 nematodes to clean up the genetic background for any additional mutations. Outcrossed homozygous *phy-3*$^{-/-}$ nematodes were examined by microscopy. End points of the deletion were determined by sequencing the PCR products spanning the deleted region.

In order to determine the 4-hydroxyproline content in egg cells, the amino acid analysis of nematodes and early embryos was performed. Synchronous cultures of wild type and *phy-3*$^{-/-}$ *C. elegans* strains were prepared by letting alkaline hypochlorite purified embryos hatch in the absence of food (Sulston & Hodgkin 1988). The resulting synchronously arrested L1 larvae were then inoculated onto agar plates containing a lawn of *E. coli* OP50 (Sulston & Hodgkin 1988), and the cultures were incubated at 20°C until the nematodes developed to the late L4 stage. The nematodes were then washed from the plates with M9 buffer, bleached, and incubated at 37°C for 10 min with shaking. The resulting worm debris was passed though a narrow gauge needle 40-50 times, washed with distilled water, and centrifuged at 3500 rpm for 5 min. A fraction of the obtained egg pellet was analyzed by light microscopy to confirm that the embryos were at an early developmental stage. The egg pellet was resuspended in 400 µl of ST buffer, heated at 100°C for 2 min, and incubated at 22°C overnight. The sample was then centrifuged down, and the resulting supernatant was hydrolyzed and subjected to amino acid analysis in an Applied Biosystems 421 amino acid analyzer. The remaining pellet was further resuspended in 400 µl of ST buffer containing 5% β-mercaptoethanol and the procedure described above was repeated. The removal of sodium dodecyl sulphate from the protein solution was performed by addition of potassium salts (Suzuki & Terada 1988).

5 Results

5.1 Characterization of the human and mouse prolyl 4-hydroxylase α(II) subunit genes (I)

Five positive human and three positive mouse prolyl 4-hydroxylase α(II) subunit clones were obtained from the genomic libraries. These clones did not contain the 5´ ends of the genes, therefore, P1 human and BAC mouse library screening services were used to obtain the 5´ ends. The human and mouse α(II) subunit genes were approximately 34.6 and 30.3 kb in size, respectively, both containing 16 exons and 15 introns. The first exons containing the transcription initiation sites consisted entirely of 5´ untranslated sequences in both genes, and the translation initiation sites were located in exon two. The exons represented about 7% of the sequences in both genes and they were well conserved in terms of sizes when compared to the exon sizes of the human α(I) subunit gene (Helaakoski *et al.* 1994). The boundaries of the exons and introns starting from exon 2 in the human and mouse genes followed the GT/AG consensus rule. The sizes of exons 3-14, which consisted entirely of amino acid coding sequences, varied from 54 to 240 bp. The last exons in the human and mouse genes contained 71 bp of coding region and the entire 3´ untranslated regions with classic polyadenylation signals 371 bp and 370 bp downstream of the translation stop codons. The size range of the introns was similar in both genes, being from 48-49 bp to over 8 kb, with the first intron being the largest one in both genes.

The transcription initiation sites of the human and mouse α(II) subunit genes were determined by S1 nuclease protection analysis. One major protected fragment was seen in the analyses of the 5´ ends of both genes. The sizes of the first exons were different, being 546 bp in the human and 293 bp in the mouse gene. The 5´ flanking regions in both genes contained several potential transcription factor binding sites and GC rich areas. Similar elements have been shown to be important for efficient expression of the PDI polypeptide gene (Tasanen *et al.* 1992, Tasanen *et al.* 1993). BLAST homology search in the GenBank database identified a human EST sequence that was highly similar to that of the

prolyl 4-hydroxylase α(II) subunit, differing only in the amino acids coded by exon 12. Sequencing of introns 11 and 12 of the human and mouse α(II) subunit genes showed that both genes contained this alternatively spliced exon, named here as exon 12a, which was located upstream from the previously known exon, named here exon 12b. The size of the novel exon 12a was 60 bp, whereas the previously known exon 12b was 66 bp in size.

A recombinant baculovirus coding for a human α(II) subunit containing amino acids coded by exon 12a was generated and used to coinfect insect cells with a virus coding for the human PDI polypeptide (Vuori *et al.* 1992b). The recombinant α(II)ex12a subunit formed a fully active enzyme tetramer with the recombinant human PDI polypeptide as efficiently as the α(II)ex12b subunit, and both enzyme forms had identical K_m values for a synthetic peptide substrate and the cosubstrates.

Expression of human prolyl 4-hydroxylase α(II) subunit sequences corresponding to the two alternatively spliced exons 12a and 12b in various tissues was studied by PCR amplification of Multiple Tissue cDNA panels and cartilage and adult fibroblast cDNAs. Sequences corresponding to both types of exon were found to be expressed in all tissues studied, except that sequences corresponding to exon 12a were not detected in adult leukocytes.

5.2 Tissue specificity of prolyl 4-hydroxylase type I and II isoenzymes in developing and mature human tissues (II)

The previously generated monoclonal antibody K4 to the prolyl 4-hydroxylase α(II) subunit had been found to give a weak signal in the basement membrane zone below the epidermis (Annunen *et al.* 1998). However, the present immunoelectron microscopy studies revealed that the antibody also recognized a basement membrane component outside the cells. The basement membrane signal seemed quite specific and clusters of label were observed in close contact with the endothelial cell membrane. To study the specificity of this signal further, a new monoclonal antibody M14 to the human α(II) subunit was generated. The new M14 antibody, like the K4 antibody, gave a signal in the ER of capillary endothelial cells in immunoelectron microscopy, but no signal in the basement membrane outside the cells was observed. A new polyclonal antibody R17 against the human α(I) subunit was also generated and shown to give a slightly stronger signal than the previously used L7P antibody (Annunen *et al.* 1998). Expression of the type I and II prolyl 4-hydroxylases in various tissues was studied by immunofluorescence staining using the M14 and R17 antibodies.

No signal was observed when cultured human umbilical vein endothelial cells were stained with the R17 antibody for the type I isoenzyme. However, the K4 antibody and the new M14 antibody gave a strong signal for the type II isoenzyme. These results were confirmed by Western blotting with identical results. In kidney specimens, the type II isoenzyme was expressed in the endothelial cells of blood vessels, and the staining was essentially similar in the foetal and adult kidney. Also cells of the developing glomeruli stained positively in foetal kidney. Tubular structures of collecting duct caliber stained positively for the α(II) subunit in both foetal and adult kidneys. A clear signal for the type I prolyl 4-hydroxylase was present in the undifferentiated mesenchymal cells of the

developing interstitium, some immature tubules, and the capsule of the foetal kidney. In the adult kidney, the type I enzyme was expressed in interstitial fibroblasts around tubular structures, fibroblastic cells of the capsule, smooth muscle cells of arterial walls, and mesangial cells of the glomeruli.

In foetal liver samples, the expression level of the type II prolyl 4-hydroxylase was generally low, the strongest signal being seen in the periportal hepatocytes of the developing liver. In the adult liver, occasional sinusoidal structures were stained with the α(II) antibody, and staining was also seen in the epithelium of the bile ducts. As in other tissues, the capillaries were stained with the α(II) antibody. Type I prolyl 4-hydroxylase was found to be expressed in fibroblastic cells and smooth muscle cells in the portal tracts of the foetal liver, but no detectable signal was seen in the parenchyme. Staining with the α(I) antibody was also seen in some hepatocytes in the adult liver. Analysis of specimens with hepatocellular carcinoma with or without chirrosis showed that malignant hepatocytes stained positively for both isoenzymes, although the signal for the type I enzyme was much stronger. Fibroblasts in the cirrhotic areas and smooth muscle cells of larger vessels gave a strong signal with the α(I) antibody. Cultured hepatoblastoma cells became strongly stained with the antibody to the α(I) subunit, while only a very faint signal was obtained with the antibody to the α(II) subunit.

In adult human striated muscle, the antibody to the α(II) subunit stained the capillaries but not the myocytes. The myocytes, however, were strongly stained with the α(I) antibody as was the musculature of the large arteries. Also, the arrector muscle of the hair in the skin gave a strong signal with the α(I) antibody. In developing and mature placenta, the decidual cells were either negative or became only faintly stained with the α(II) antibody, whereas the decidual cells and the smooth muscle cells of the spiral arteries stained positively with the α(I) antibody.

During intramembranous ossification, the bone forms via the condensation and differentiation of mesenchymal cells in the osteoid secreted by osteoblasts, and therefore the developing bone might stain almost exclusively with the α(I) antibody. However, the osteoblasts were strongly stained with both the α(I) and the α(II) antibody. The signal for the α(I) antibody was seen earlier in ossification, in the undifferentiated mesenchymal cells, whereas the signal for the α(II) antibody became evident only later during ossification.

5.3 Expression of type I and II prolyl 4-hydroxylases and type I and II collagens in primary bone tumours (III)

Expression of type I and II prolyl 4-hydroxylases in bone tumours was studied by immunofluorescence staining. The type II isoenzyme was rarely seen in malignant primary bone tumour cells, with the main isoenzyme in these cells being the type I prolyl 4-hydroxylase. In normal chondrocytes, the type II isoenzyme was the main form (Annunen *et al.* 1998), but according to these immunofluorescence studies it was strongly down-regulated in malignant chondrous cells. A surprising finding was an intracellular signal for the type I and II collagens in osteosarcomas and chondrosarcomas, while an extracellular labeling for type II collagen in chondrosarcomas was practically absent.

Western blot analysis of rat osteosarcoma cells showed a reduced mobility for type II collagen, indicating that it was incompletely processed and still contained the propeptides.

The expression of prolyl 4-hydroxylase isoenzymes was variable in benign bone tumours, although there was a tendency for slightly higher expression levels for the type II isoenzyme in chondrous tumours. Giant cell tumours differed from the other benign tumours in that only low levels of expression were seen for both isoenzymes. Benign bone tumours expressed variable amounts of type I and II collagens and the type II collagen was more often secreted into the extracellular space, when compared to their malignant counterparts. Giant cell tumours expressed mainly type I collagen, whereas the expression levels for type II collagen were low. A positive signal for osteocalcin, a specific marker for osteocytes, was detected in both benign and malignant chondrous tumours (Vincent *et al.* 1991, Desbois *et al.* 1994).

5.4 Identification and characterization of a novel prolyl 4-hydroxylase isoenzyme in *Caenorhabditis elegans* (IV)

A gene encoding a new prolyl 4-hydroxylase isoenzyme, named PHY-3, was identified by sequence homology search in the *C. elegans* database and cloned. The *phy-3* cDNA encoded a 318 amino acid polypeptide, which was much shorter than the previously characterized vertebrate prolyl 4-hydroxylase α subunits and the *C. elegans* PHY-1 and PHY-2 subunits with a size range from 514 to 544 residues (Veijola *et al.* 1994, Helaakoski *et al.* 1995, Annunen *et al.* 1997, Friedman *et al.* 2000, Hill *et al.* 2000, Winter & Page 2000). The genomic sequence of *phy-3* was found on cosmid T20B3, which maps to chromosome V. Comparison of the *phy-3* cDNA and the genomic clone T20B3 revealed that the coding sequence was organized into five exons, the first exon being separated from the second one by an exceptionally long intron (4151 nucleotides), as typical intron lengths in *C. elegans* are only 44-52 nucleotides (Blumenthal & Steward 1997).

To study the spatial expression pattern of *phy-3* in live nematodes, a reporter gene plasmid was constructed by ligating the putative promoter region of *phy-3* and six first nucleotides of the coding sequence in-frame to a reporter gene encoding β-galactosidase (*lacZ*). This construct was then microinjected into the *C. elegans* germ line. Expression of *phy-3::lacZ* was detected in the spermatheca, where the fertilization of the oocytes occurs. In contrast to *phy-1* and *phy-2*, which were expressed in hypodermal cells at all life stages from embryo to adult, no staining was observed in these cells even after overstaining. Expression of the PHY-3 polypeptide in spermatheca was confirmed by immunofluorescence staining with a polyclonal antibody against PHY-3.

The temporal expression pattern of *phy-3* was studied by RT-PCR using mRNA samples prepared from highly synchronous postembryonic cultures of *C. elegans* extracted at 2-h intervals from hatched L1 larvae to young adults (Johnstone & Barry 1996). The *phy-3* transcripts were detected in the late L3 and L4 larval stages and in adult nematodes. However, an additional PCR analysis of a cDNA pool prepared from embryos gave a strong signal, indicating that *phy-3* was also expressed in embryos.

Animals homozygous for a *phy-3* deletion were phenotypically wild type. No defects were found in their morphology, fertility features, or behavior. As PHY-3 was expressed in the embryos and the spermatheca of the L4 larvae and adult nematodes, we especially looked for defects in the early embryos. Amino acid analyses showed that the 4-hydroxyproline content of the *phy-3* knock-out early embryos was decreased by about 90% when compared to the wild type strain. This drastically decreased 4-hydroxyproline content had no detectable effect on the viability of the early embryos in the nonstressed laboratory conditions. No decrease in the 4-hydroxyproline content of adult nematodes was detected.

Recombinant PHY-3 polypeptide was produced in insect cells, and the cell lysate was analyzed by SDS-PAGE followed by Coomassie staining and Western blotting. The majority of the PHY-3 polypeptide formed insoluble aggregates as reported previously also for the *C. elegans* PHY-1 and PHY-2 polypeptides (Veijola *et al*. 1994, Myllyharju J, Kukkola L, Winter AD & Page AP, unpublished observations) and the human prolyl 4-hydroxylase α subunits (Vuori *et al*. 1992b, Annunen *et al.* 1997). The recombinant PHY-3 polypeptide exhibited prolyl 4-hydroxylase activity only when it was coexpressed with *C. elegans* PDI-1, but not with human PDI or *C. elegans* PDI-2 (Veijola *et al*. 1996a). The activity level in the lysates from insect cells coexpressing PHY-3 and PDI-1 was about 22-27 % of that in extracts from cells expressing the *C. elegans* PHY-1/human PDI dimer or the human type I prolyl 4-hydroxylase. However, the amount of PHY-3 in the soluble fraction was very small and thus the specific activity of PHY-3 may not be significantly lower than that of the other prolyl 4-hydroxylases expressed. The PHY-3 polypeptide contains five potential N-glycosylation sites. Several forms of the PHY-3 polypeptide corresponding to differentially glycosylated forms were seen in Western blots, whereas only one major band and two minor bands probably representing degradation products were present after N-glycosidase F treatment.

6 Discussion

6.1 Structures of the human and mouse prolyl 4-hydroxylase α(II) subunit genes

The transcription initiation sites of the human and mouse α(II) subunit genes were determined by S1 nuclease protection analysis, which showed that the sizes of the first exons were 546 bp and 293 bp in the human and mouse genes, respectively. However, the majority of the size difference in exons 1 was due to the presence of an exon-intron boundary in the human gene 390 bp downstream from that in the mouse gene, and in fact the transcription start site in the mouse gene was 145 bp upstream of that in the human gene.

Unlike the human α(I) subunit gene (Helaakoski *et al.* 1994), the human and mouse α(II) subunit genes contained no TATA box in their putative promoter regions. The GC content in the 5´ flanking regions was 72% and 69% in the human and mouse α(II) subunit genes, respectively, and several potential transcription factor binding sites were identified in these regions. The transcription factor Sox9 has been shown to be required for the differentiation of chondrocytes from mesenchymal cells (de Crombrugghe *et al.* 2000), and mutations in Sox9 are responsible for a congenital bone and cartilage malformative syndrome known as campomelic dysplacia (Zenteno-Ruiz *et al.* 2001). The type II isoenzyme was the main prolyl 4-hydroxylase form in chondrocytes (Annunen *et al.* 1998) and interestingly, the first introns in the human and mouse α(II) subunit genes were found to contain 20 and 16 putative Sox9 binding sites, whereas none was found in the promoter region or in the first intron of the human α(I) gene. Sox9 has been shown to be involved in cartilage specific transcription of the *Col2a1*, *Col9a2*, *Col11a2*, and *aggregan* genes (de Crombrugghe *et al.* 2000), but the functional significance of the regulatory elements identified in the α(II) subunit genes remained to be determined. The promoter regions of the human and mouse α(II) subunit genes showed very little nucleotide sequence similarity to that of the human α(I) subunit gene (Helaakoski *et al.*

1994), suggesting that major differences exist in the regulation of expression of the α(I) and α(II) subunit genes.

The translation initiation site was located in exon 2 in the human α(I) (Helaakoski *et al.* 1994), human α(II), and mouse α(II) subunit genes. The human and mouse α(II) subunit genes had very similar exon-intron structures when compared to the human α(I) subunit gene (Helaakoski *et al.* 1994), with the exception of alternatively spliced exons. Starting from exon 2, the exon sizes and exon-intron boundaries were conserved in all three genes, with the exception that exon 7 was 3 bp longer in the human α(I) subunit gene.

The human α(I) subunit gene has been shown to contain two homologous 71 bp exons 9 and 10, which are subjected to mutually exclusive alternative splicing (Helaakoski *et al.* 1989). Sequencing of introns 8 and 9 in the human and mouse α(II) subunit genes showed that these genes contain no alternatively spliced exon at the corresponding site. However, BLAST homology search against the human α(II) subunit amino acid sequence identified an EST clone differing in the amino acids coded by exon 12 from the previously identified human α(II) subunit cDNA (Annunen *et al.* 1997). Sequencing of introns 11 and 12 in the human and mouse genes showed that the novel alternatively spliced exon, called 12a, existed in both genes and was located upstream of the previously known exon 12, named here exon 12b. The identity between the 60 bp exon 12a and the 66 bp exon 12b were 45% at the amino acid level, while exons 12a in the human and mouse gene were 100% identical and exons 12b were 77% identical at the amino acid level. Thus the similarity between exons 12a and 12b was lower than that between exons 9 and 10 in the human α(I) subunit gene, which were 58% identical at the amino acid level (Helaakoski *et al.* 1989). Similarly to the alternatively spliced α(I) subunit mRNAs (Helaakoski *et al.* 1989), no major differences were seen in the expression patterns of the alternatively spliced α(II) subunit mRNA forms in several human tissues studied, except that exon 12a sequences were not detected in mature leukocytes. The biological significance of the two alternatively spliced forms of the α(II) subunit mRNA remained to be explained, but similarly to the alternatively spliced α(I) subunit (Helaakoski *et al.* 1989), both recombinant α(II) subunit isoforms efficiently assembled into active prolyl 4-hydroxylase tetramers with identical K_m values for the substrate and cosubstrates.

The last exons in the human and mouse α(II) subunit genes contained 71 bp of coding sequences, and unlike the human α(I) subunit gene (Helaakoski *et al.* 1994), both α(II) subunit genes contained a classic polyadenylation signal. The predicted sizes of the human and mouse α(II) subunit mRNAs from the identified transcription initiation sites to the polyadenylation sites were about 2.6 and 2.4 kb respectively, when poly(A) tails were excluded, whereas the size of the human α(I) subunit mRNA was 3 kb (Helaakoski *et al.* 1989). The human and mouse α(II) subunit genes were approximately 35 and 30 kb in size, respectively, thus being considerably smaller than the more than 69 kb human α(I) subunit gene (Helaakoski *et al.* 1994). The size difference between the α(I) and α(II) subunit genes was due to considerably longer introns in the α(I) gene, ranging from 750 bp to over 16 kb in size (Helaakoski *et al.* 1994), whereas the intron sizes in the human and mouse α(II) subunit genes varied from 48-49 bp to over 8 kb. The first intron was the largest one in all three genes.

The human prolyl 4-hydroxylase α(I) subunit gene has been mapped to chromosome 10 (Pajunen *et al.* 1989), whereas the human α(II) subunit gene was located on chromosome 5 according to the GenBank database information. The human PDI gene was located on chromosome 17 (Pajunen *et al.* 1988, Popescu *et al.* 1988, Pajunen *et al.* 1991), and the mouse α(II) subunit and PDI genes were both located on chromosome 11 based on the GenBank information. According to preliminary GenBank data, the mouse α(I) subunit gene has an identical exon-intron structure starting from exon 2 as the human α(I) subunit gene, and the size of the mouse α(I) subunit gene was approximately 35 kb, which was about half that of the human α(I) subunit gene.

Prolyl 4-hydroxylase appeared to be an evolutionarily old enzyme as it was found in species from *Chlorella* viruses to a number of invertebrates and vertebrates. Phylogenetic analysis of prolyl 4-hydroxylase sequences from human, mouse, rat, chicken, *C. elegans*, *Drosophila*, and *Paramecium bursaria Chlorella* virus suggested that the α(I) and α(II) subunit genes may have arisen by gene duplication (Friedman *et al.* 2000). Homologous recombination of *Alu* sequences has been shown to be involved in many gene rearrangements (Rudiger *et al.* 1995, Strout *et al.* 1998), and the human α(I) subunit gene has been reported to contain one *Alu* sequence in the 5´ untranslated region and one in close proximity to the two putative polyadenylation signals (Helaakoski *et al.* 1994). BLAST homology searches indicated that the mouse α(II) subunit gene had no *Alu* repeats, whereas the human α(II) subunit gene had four full-length *Alu* repeats in introns 1, 4, and 12b, and three *Alu* fragments in introns 1 and 6.

6.2 Expression of the type I and II prolyl 4-hydroxylases in various human tissues

The previously generated human α(II) subunit antibody (Annunen *et al.* 1998) was shown here to crossreact with an unknown basement membrane component. Such crossreaction was not observed with the new α(II) subunit antibody prepared in this work. The present data indicated that the two prolyl 4-hydroxylase isoforms had major differences in their expression patterns in many tissues. The type I isoenzyme was the main form in most cells and tissues studied, whereas the type II isoenzyme was the main form in chondrocytes and the main or only form in capillary endothelial cells. In general, the type I enzyme was expressed especially in less differentiated cells in developing and malignant tissues and seemed to be the main form in immature mesenchymal cells of foetal tissues, while during differentiation of these cells the expression pattern switches and the expression level of the type II enzyme increases.

The type I isoenzyme was the predominant form in osteo- and chondrosarcomas. In normal chondrocytes the type II isoenzyme was the main form (Annunen *et al.* 1998), but its expression level were low in chondrosarcomas, thus supporting the predominance of type I isoenzyme in less differentiated cells and tissues. Type I and II collagens were detected in osteosarcomas and chondrosarcomas predominantly intracellularly, and in the latter tumours the extracellular signal for type II collagen was very faint. Western blot analysis of rat osteosarcoma cells supported the finding of intracellular retention of type II procollagen by showing a reduced mobility when compared to pepsinized type II

collagen. Intracellular retention of type I and II collagens in both malignant and benign primary bone tumours inspite of prolyl 4-hydroxylase expression may indicate that the amount of cofactors such as ascorbate or oxygen was insufficient in the microenvironments of the tumours. It has been shown previously that in the absence of calcium, chick embryonic calvaria possesses cells which can form genuine cartilage during the process of intramembranous ossification. This resulted the detection of type II collagen instead of type I collagen in these cells (Jacenko & Tuan 1986, Jacenko *et al.* 1995). The present results showed that the woven bone cells in fibrous dysplacia produced type II collagen, forming genuine cartilage.

Neoplastic chondrocytes had distinctly different phenotypes when compared to nonmalignant chondrocytes: even well differentiated chondrocytic cells can express type I collagen (Aigner *et al.* 2000). Hypertrophic chondrocytes can be induced to transdifferentiate into bone-forming cells if their matrix is cut off (Roach *et al.* 1995a), and the tumorigenic cells in chondrosarcomas were also able to transdifferentiate into bone-forming cells (Aigner *et al.* 2000). According to the present results, the neoplastic chondrocytes in chondrosarcomas and chondromas expressed osteocalcin, a highly osteoblast-specific protein (Towler *et al.* 1994, Ducy & Karsenty 1995). It was thus evident that transformed chondrous tumours may express proteins that are specific to mature osteocytes.

6.3 The *C. elegans* prolyl 4-hydroxylases and their functions

Collagens are major components of *C. elegans*, comprising approximately 1% of the weight of an adult nematode (Johnstone 1994). This nematode lives at cooler temperatures than the warm-blooded vertebrates, and its collagens contain about 12% hydroxyproline. Most *C. elegans* collagens are located in the cuticle and the basement membrane (Kramer 1994a, Kramer 1994b).

A third *C. elegans* prolyl 4-hydroxylase α subunit isoform, PHY-3, was cloned and characterized in this study. The processed PHY-3 consisted of 295 amino acids, being thus much shorter than the previously characterized vertebrate and *C. elegans* α subunits. An even shorter prolyl 4-hydroxylase monomer with a size of 210 residues has been characterized from the *Paramecium bursaria Chlorella* virus-1 (Eriksson *et al.* 1999) and a 261 amino acid prolyl 4-hydroxylase subunit has been cloned from *Arabidobsis thaliana* (Hieta & Myllyharju 2002). The two histidines and one aspartate that bind Fe^{2+} in human prolyl 4-hydroxylases were conserved, as well as the lysine that binds 2-oxoglutarate (Myllyharju & Kivirikko 1997). The C-terminal region was the most conserved part between species, the PHY-3 amino acids 150-295 being 23-30% identical to the corresponding residues of PHY-1 and PHY-2 and human α(I) and α(II) (Helaakoski *et al.* 1989, Annunen *et al.* 1997, Friedman *et al.* 2000, Winter & Page 2000). During the course of this work, we identified two additional α subunit isoforms from the *C. elegans* genome, named here as *phy-4* and *phy-5*. The iron-binding amino acids, two histidines and one aspartate, were conserved within the *C. elegans* prolyl 4-hydroxylase isoenzyme family according to sequence alignment studies. Also the lysine that binds the C5 carboxyl group of 2-oxoglutarate was conserved between the *C. elegans* PHY isoforms

(Myllyharju & Kivirikko 1997). Coexpression of recombinant PHY-3 in insect cells with *C. elegans* PDI-1 produced prolyl 4-hydroxylase activity, whereas coexpression of PHY-3 with PDI-2 produced no activity. We were not able to determine whether PHY-3 formed a tetramer or a dimer with PDI-1, or whether PHY-3 is a monomer like the *A. thaliana* (Hieta & Myllyharju 2002) and viral prolyl 4-hydroxylases (Eriksson *et al.* 1999).

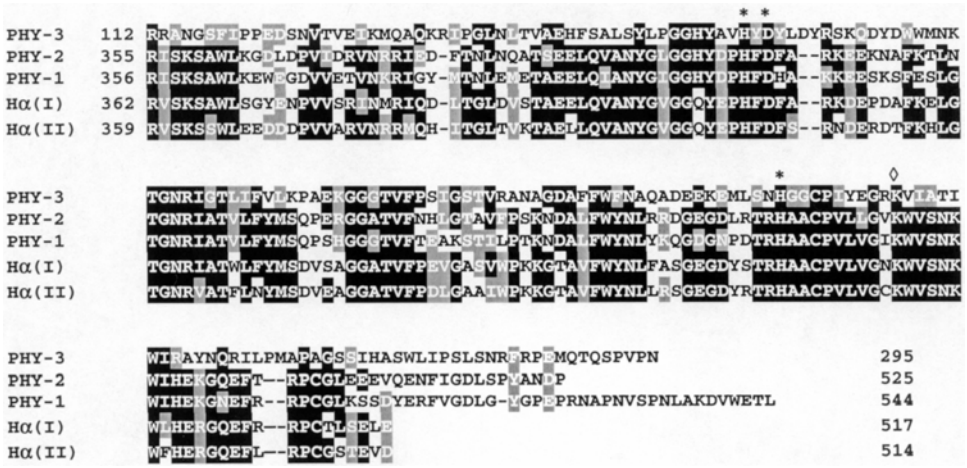

Fig. 4. Comparison of the amino acid sequence between the unprocessed *C. elegans* prolyl 4-hydroxylase α subunits (PHY-1-PHY-3) and the human α(I) and α(II) subunits. White letters on a black background indicate identity and white letters in a gray background indicate similarity. The iron-binding sites are indicated by an asterix (*) and the lysine binding the C5 carboxyl group of 2-oxoglutarate is marked by a diamond (◊).

The *phy-1* and *phy-2* genes were expressed in the hypodermal cells, which are the cells that secrete the cuticle components, and their expression patterns were temporally and spatially associated with those of the cuticular collagens (Kramer 1994a, Kramer 1994b, Friedman *et al.* 2000, Hill *et al.* 2000, Winter & Page 2000). However, expression of *phy-3* was only detected in embryos, late larval stages, and adult nematodes, and in the latter two its expression was restricted to spermatheca. No hypodermal signal was observed for PHY-3 indicating that it was not involved in the processing of cuticular collagens. The results obtained from the time-course PCR analysis of the expression of the *phy-3* mRNA also supported this finding. Such a highly restricted expression pattern has not been reported for any other prolyl 4-hydroxylase in any other species.

The majority of *phy-1*$^{-/-}$ nematodes were viable but had a dumpy phenotype (Hill *et al.* 2000, Friedman *et al.* 2000, Winter & Page 2000), suggesting that the primary function of PHY-1 is to ensure normal body morphology. However, a fraction of the *phy-1*$^{-/-}$ nematodes arrested as embryos (Hill *et al.* 2000) or had a severely disrupted body morphology (Hill *et al.* 2000, Winter & Page 2000). Inactivation of the *phy-2* gene displayed no observable phenotype (Friedman *et al.* 2000, Winter & Page 2000), but the *phy-1*$^{-/-}$/*phy-2*$^{-/-}$ double mutants arrested during embryogenesis (Friedman *et al.* 2000, Winter & Page 2000).

Winter & Page (2000) used RNAi technology to inactivate the *phy-2* function in *phy-1*$^{-/-}$ mutants. Organogenesis and elongation were normal in the mutant embryos, but the fully elongated embryos began to retract to a less elongated form indicating that the cuticule could not maintain the morphology of the embryos. This phenotype was identical to that caused by a cuticular collagen mutation, *sqt-3* (van der Keyl *et al.* 1994). Friedman *et al.* (2000) used the *phy-1*$^{-/-}$ and *phy-2*$^{-/-}$ genomic knock-out lines for their experiments. They observed that the *phy-1*$^{-/-}$*/phy-2*$^{-/-}$ embryos elongated usually to the threefold stage, but were unable to maintain their shape and often exploded. This phenotype was similar to that of the *emb-9* embryos having a mutation in the gene coding for type IV collagen, which is an essential component of basement membranes. They suggested that *phy-1* and *phy-2* may play an important role in basement membrane synthesis, even though a positive signal for *phy-1* expression was seen neither with a *LacZ* fusion vector nor by immunofluorescence (Friedman *et al.* 2000, Winter & Page 2000), and only faint *phy-2* expression was detected in the body wall muscle cells, where the synthesis of type IV collagen occurred (Graham *et al.* 1997, Winter & Page 2000).

The decreased 4-hydroxyproline content in the *phy-3*$^{-/-}$ early embryos had no detectable effect on the fertility of the nematodes at least in the nonstressed laboratory conditions, and the *phy-3*$^{-/-}$nematodes seemed to be perfectly normal. However, the 4-hydroxyproline content was not decreased to zero in the early embryos, indicating that some other prolyl 4-hydroxylase isoform is likely to be active in the developing embryo. The viable *phy-3*$^{-/-}$ phenotype supported this suggestion. The *phy-1/phy-3* double knock-out mutant gave no additional phenotype when compared to the *phy-1* knock-out, and the *phy-2*$^{-/-}$*/phy-3*$^{-/-}$ mutants gave no observable phenotype (Winter A & Page A, unpublished results).

7 Conclusions

Disruption of collagen biosynthesis leads to human syndromes with phenotypes ranging from mild skin abnormalities and joint hyperextensibility to pre- or perinatal lethality (Myllyharju & Kivirikko 2001). These syndromes are most frequently caused by mutations in genes that encode collagen polypeptides, leading to abnormal, reduced, or absent collagen triple helix formation. An exception is Ehlers-Danlos sydrome (EDS) type VI, which is usually due to impaired lysyl hydroxylase activity, and EDS type VIIC, which is caused by deficiency in procollagen N proteinase activity (Myllyharju & Kivirikko 2001). Prolyl 4-hydroxylase has not been implicated in any inherited human disorder, but it is obvious that a proper functioning of this enzyme is essential in humans. The data on the prolyl 4-hydroxylase α(II) subunit genes obtained in this thesis offer a number of opportunities. With the known exon-intron structure of the human gene, it is possible to start screening for mutations in patient samples from candidate diseases. The 5´ genomic clones can be used in studies on the gene regulation. With the known structure of the mouse α(II) subunit gene, the function of this gene can be studied by generation of transgenic and knock-out mouse lines. Additionally, *C. elegans* can be used as a model organism for gathering information for the consequences of prolyl 4-hydroxylase gene mutations.

The collagen family consists of more than 20 proteins formally defined as collagens (Myllyharju & Kivirikko 2001). The existence of several forms of prolyl 4-hydroxylase raises the possibility that these isoenzymes may show different specificities with respect to the collagen types to be hydroxylated. However, as all collagens contain similar –Gly-X-Y- triplets it seems possible that the main differences between the isoenzymes may be found in their expression patterns rather than specificities with respect to various collagen types. It has been reported that there may be no collagen specificity for lysyl hydroxylase isoforms (Wang et al. 2000).

Before identification of the type II prolyl 4-hydroxylase, mutations in the α(I) gene were thought to be lethal. However, it now seems possible that mutations in the gene for one isoenzyme may not be lethal if expression of the other isoenzyme can compensate for the missing activity. Currently no candidate diseases have been mapped at the genomic areas of the α(I) or the α(II) subunit genes. A knock-out mouse line was generated in order to study the involvement of the *OTCN2* gene in lipid metabolism (Zhu et al. 2000).

However, the deleted gene loci comprised of a 450 kb genomic area, which, based on computational analysis, coded for nine genes including the mouse α(II) subunit gene. Surprisingly, the null mice were viable, but exhibited a variety of abnormalities including severe hypertriglyceridemia, hepatic and cardiac enlargement, growth retardation, and premature mortality. However, no further experiments were made, such as studies on the structure of the cartilage or studies to indicate whether such mice develop osteoporosis or joint disorders more easily (Zhu et al. 2000).

In this work antibodies against the prolyl 4-hydroxylase α(I) and α(II) subunits were generated, providing an essential tool for the analysis of the tissue distribution of the type I and II enzymes. According to the immunofluorescence studies, the type II isoenzyme is the main form, or the only form, in endothelial cells, whereas skeletal myocytes and smooth muscle cells appear to have the type I isoenzyme as their main form. Some cell types that are likely to synthesize collagen expressed only very small amounts of either known human prolyl 4-hydroxylase isoenzyme. At least one additional gene coding for a prolyl 4-hydroxylase α subunit like polypeptide is present in the human genome (Kukkola et al. 2000), and in the light of the data obtained with double knock-out worms, it would seem possible that different isoenzymes show genetic redundancy thus preventing underhydroxylation, which would have deleterious effects on connective tissues and skeletal structures.

8 References

Abrams EW & Andrew DJ (2002) Prolyl 4-hydroxylase α-related proteins in *Drosophila melanogaster:* tissue-specific embryonic expression of the 99F8-9 cluster. Mechan Develop 112: 165-171.

Ackley BD, Crew JR, Elamaa H, Pihlajaniemi T, Kuo CJ & Kramer JM (2001) The NC1/endostatin domain of *Caenorhabditis elegans* type XVIII collagen affects cell migration and axon guidance. J Cell Biol 152: 1219-1232.

Adams E (1978) Invertebrate collagens. Science 202: 591-598.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amantides PG, Scherer SE, Li PW, *et al*. (2000) The genome sequence of *Drosophila melanogaster*. Science 287: 2185-2195.

Aigner T, Loos S, Inwards C, Perris R, Perissinotto D, Unni KK & Kirchner T (1999) Chondroblastoma is an osteoid-forming, but not cartilage-forming neoplasm. J Pathol 189: 463-469.

Aigner T, Loos S, Muller S, Sandell LJ, Unni KK & Kirchner T (2000) Cell differentiation and matrix gene expression in mesenchymal chondrosarcomas. Am J Pathol 156: 1327-1335.

Annunen P, Autio-Harmainen H & Kivirikko KI (1998) The novel type II prolyl 4-hydroxylase is the main enzyme form in chondrocytes and capillary endothelial cells, whereas the type I enzyme predominates in most cells. J Biol Chem 273: 5989-5992.

Annunen P, Helaakoski T, Myllyharju J, Veijola J, Pihlajaniemi T & Kivirikko KI (1997) Cloning of the human prolyl 4-hydroxylase α subunit isoform α(II) and characterization of the type II enzyme tetramer. The α(I) and α(II) subunits do not form a mixed $[\alpha(I)\alpha(II)]\beta_2$ tetramer. J Biol Chem 272: 17342-17348.

Annunen P, Koivunen P & Kivirikko KI (1999) Cloning of the α subunit of prolyl 4-hydroxylase from *Drosophila* and expression and characterization of the corresponding enzyme tetramer with some unique properties. J Biol Chem 274: 6790-6796.

Bacon NC, Wappner P, O'Rourke JF, Bartlett SM, Shilo B, Pugh CW & Ratcliffe PJ (1998) Regulation of the *Drosophila* bHLH-PAS protein Sima by hypoxia: functional evidence for homology with mammalian HIF-1α. Biochem Biophys Res Commun 249: 811-816.

Bai Y, Muragaki Y, Obata K, Iwata K & Ooshima A (1986) Immunological properties of monoclonal antibodies to human and rat prolyl 4-hydroxylase. Biochem J 99: 1563-1570.

Barstead RJ & Waterston RH (1991) Vinculin is essential for muscle function in the nematode. J Cell Biol 114: 715-724.

Bassuk JA, Kao WW, Herzer P, Kedersha NL, Seyer J, DeMartino JA, Daugherty BL, Mark GE, III & Berg RA (1989) Prolyl 4-hydroxylase: molecular cloning and the primary structure of the α subunit from chicken embryo. Proc Natl Acad Sci USA 86: 7382-7386.

Bateman JF, Lamandé SR & Ramshaw JAM (1996) Collagen superfamily. In: Comper, WD (ed) Extracellular Matrix. Harwood academic publishers, Amsterdam, The Netherlands, p 22-67.

Bella J & Berman HM (1996) Crystallographic evidence for C α-H...O=C hydrogen bonds in a collagen triple helix. J Mol Biol 264: 734-742.

Bella J, Brodsky B & Berman HM (1995) Hydration structure of a collagen peptide. Structure 3: 893-906.

Berg RA & Prockop DJ (1973) 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: 115-120.

Berg RA, Schwartz ML & Crystal RG (1980) Regulation of the production of secretory proteins: intracellular degradation of newly synthesized "defective" collagen. Proc Natl Acad Sci USA 77: 4746-4750.

Bird AF & Bird J (1991) The structure of nematodes 2nd edition, Academic press, p 7-74.

Blaschke UK, Eikenberry EF, Hulmes DJ, Galla HJ & Bruckner P (2000) Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. J Biol Chem 275: 10370-10378.

Blumenthal T & Steward K (1997) RNA processing and gene structure. In: Riddle DL, Blumenthal T, Meyer BJ & Priess JR (eds) *C. elegans* vol II. Cold Spring Harbor, NY, p 117-145.

Breul SD, Bradley KH, Hance AJ, Schafer MP, Berg RA & Crystal RG (1980) Control of collagen production by human diploid lung fibroblasts. J Biol Chem 255: 5250-5260.

Brewton RG & Mayne R (1994) Heterotypic type II, IX, and XI fibrils: Comparison of vitreous and cartilage forms. In: Yurchenko, P D, Birk, D E, & Mecham, R P (eds) Extracellular Matrix Assembly and Sctructure. Academic Press, Orlando, Florida, p 129-170.

Brodsky B & Ramshaw JA (1997) The collagen triple-helix structure. Matrix Biol 15: 545-554.

Bruckner-Tuderman L (1994) Collagen VII and bullous disorders of the skin. Dermatology 189 Suppl 2: 16-20.

Bruckner P, Vaughan L & Winterhalter KH (1985) Type IX collagen from sternal cartilage of chicken embryo contains covalently bound glycosaminoglycans. Proc Natl Acad Sci USA 82: 2608-2612.

Bruick RK & McKnight SL (2001) A conserved family of prolyl 4-hydroxylases that modify HIF. Science 294: 1337-1340.

Bruick RK & McKnight SL (2002) Oxygen sensing gets a second wind. Science 295: 807-808.

C. elegans sequencing consortium (1998) Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science 282: 2012-2018.

Cai H, Wang CC & Tsou CL (1994) Chaperone-like activity of protein disulfide isomerase in the refolding of a protein with no disulfide bonds. J Biol Chem 269: 24550-24552.

Cardinale GJ & Udenfriend S (1974) Prolyl 4-hydroxylase. Adv Enzymol Relat Areas Mol Biol 41: 245-300.

Chessler SD & Byers PH (1992) Defective folding and stable association with protein disulfide isomerase/prolyl hydroxylase of type I procollagen with a deletion in the pro-α2(I) chain that preserves the Gly-X-Y repeat pattern. J Biol Chem 267: 7751-7757.

Chichester CO, 3d, Fuller GC & Cardinale GJ (1976) *In vivo* labeling and turnover of prolyl hydroxylase and a related immunoreactive protein. Biochem Biophys Res Commun 73: 1056-1062.

Chou M-Y & Li H-C (2002) Genomic organization and characterization of the human type XXI collagen (COL21A1) gene. Genomics 79: 395-401.

Chung CH, Au LC & Huang TF (1999) Molecular cloning and sequence analysis of aggretin, a collagen-like platelet aggregation inducer. Biochem Biophys Res Commun 263: 723-727.

Clifton IJ, Hsueh LC, Baldwin JE, Harlos K & Schofield CJ (2001) Structure of proline 3-hydroxylase. Evolution of the family of 2-oxoglutarate dependent oxygenases. Eur J Biochem 268: 6625-6636.

Cockman ME, Masson N, Mole DR, Jaakkola P, Chang GW, Clifford SC, Maher ER, Pugh CW, Ratcliffe PJ & Maxwell PH (2000) Hypoxia inducible factor-1α binding and ubiquitylation by the von Hippel-Lindau tumour suppressor protein. J Biol Chem 275: 25733-25741.

Colley KJ & Baenziger JU (1987) Identification of the post-translational modifications of the core-specific lectin. The core-specific lectin contains hydroxyproline, hydroxylysine and glucogalaktosylhydroxylysine residues. J Biol Chem 262: 10290-10295.

Counts DF, Cardinale GJ & Udenfriend S (1978) Prolyl hydroxylase half reaction: peptidyl prolyl-independent decarboxylation of alpha-ketoglutarate. Proc Natl Acad Sci USA 75: 2145-2149.

Cox GN, Fields C, Kramer JM, Rosenzweig B & Hirsh D (1989) Sequence comparisons of developmentally regulated collagen genes of *Caenorhabditis elegans*. Gene 76: 331-344.

Cox GN, Staprans S & Egdar RS (1981) The cuticle of *Caenorhabditis elegans* II. Stage-specific changes in ultratructure and protein composition during postembryonic development. Dev Biol 86: 456-470.

Coyne KJ, Qin XX & Waite JH (1997) Extensible collagen in mussel byssus: a natural block copolymer. Science 19: 1785-1786.

Cremer MA, Rosloniec EF & Kang AH (1998) The cartilage collagens: a review of their structure, organization, and role in the pathogenesis of experimental arthritis in animals and in human rheumatic disease. J Mol Med 76: 275-288.

Crossen R & Gruenwald S (1998) Baculovirus expression vector system: Instruction manual, PharMingen San Diego, CA.

Davis AE, III & Lachmann PJ (1984) Bovine conglutinin is a collagen-like protein. Biochemistry 23: 2139-2144.

Debelle L & Tamburro AM (1999) Elastin: molecular description and function. J Biochem & Cell Biol 31: 261-272.

de Crombrugghe B, Lefebvre V, Behringer RR, Bi W, Murakami S & Huang W (2000) Transcriptional mechanisms of chondrocyte differentiation. Matrix Biol 19: 389-394.

de Jong L, Albracht SP & Kemp A (1982) 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: 326-332.

de Jong L & Kemp A (1984) Stoicheiometry and kinetics of the prolyl 4-hydroxylase partial reaction. Biochim Biophys Acta 787: 105-111.

DeRider ML, Wilkens SJ, Waddell MJ, Bretscher LE, Weinhold F, Raines RT & Markley JL (2002) Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. J Am Chem Soc 124: 2497-2505.

de Waal A, de Jong L, Hartog AF & Kemp A (1985) Photoaffinity labeling of peptide binding sites of prolyl 4-hydroxylase with N-(4-azido-2-nitrophenyl)glycyl-$(Pro\text{-}Pro\text{-}Gly)_5$. Biochemistry 24: 6493-6499.

de Waal A, Hartog AF & De Jong L (1987) Photoaffinity labelling of the 2-oxoglutarate binding site of prolyl 4-hydroxylase with 5-azidopyridine-2-carboxylic acid. Biochim Biophys Acta 912: 151-155.

Desbois C, Hogue DA & Karsenty G (1994) The mouse osteocalcin gene cluster contains three genes with two separate spatial and temporal patterns of expression. J Biol Chem 269: 1183-1190.

Dijkstra K, Karvonen P, Pirneskoski A, Koivunen P, Kivirikko KI, Darby NJ, van Straaten M, Scheek RM & Kemmink J (1999) Assignment of ^{1}H, ^{13}C and ^{15}N resonances of the *a´* domain of protein disulfide isomerase. J Biomol NMR 2: 195-196.

Drickamer K, Dordal MS & Reynolds L (1986) Mannose-binding proteins isolated from rat liver contain carbohydrate- recognition domains linked to collagenous tails. Complete primary structures and homology with pulmonary surfactant apoprotein. J Biol Chem 261: 6878-6887.

Ducy P & Karsenty G (1995) Two distinct osteoblast-specific cis-acting elements control expression of a mouse osteocalcin gene. Mol Cell Biol 15: 1858-1869.

Edman JC, Ellis L, Blacher RW, Roth RA & Rutter WJ (1985) Sequence of protein disulfide isomerase and implications of its relationship to thioredoxin. Nature 317: 267-270.

Engel J (1997) Versatile collagens in invertebrates. Science 277: 1785-1786.

Engel U, Perz O, Fauser C, Engel J, David CN & Holstein TW (2001) A switch in disulfide linkage during minicollagen assembly in *Hydra* nematocytes. EMBO J 20: 3063-3073.

Elomaa O, Kangas M, Sahlberg C, Tuukkanen J, Sormunen R, Liakka A, Thesleff I, Kraal G & Tryggvason K (1995) Cloning of a novel bacterial-binding receptor structurally related to scavenger receptors and expressed in a subset of macrophages. Cell 80: 603-609.

Epstein AC, Gleadle JM, McNeill LA, Hewitson KS, O'Rourke J, Mole DR, Mukherji M, Metzen E, Wilson MI, Dhanda A, Tian YM, Masson N, Hamilton DL, Jaakkola P, Barstead R, Hodgkin J, Maxwell PH, Pugh CW, Schofield CJ & Ratcliffe PJ (2001) *C. elegans* EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. Cell 107: 43-54.

Eriksson M, Myllyharju J, Tu H, Hellman M & Kivirikko KI (1999) Evidence for 4-hydroxyproline in viral proteins. Characterization of a viral prolyl 4-hydroxylase and its peptide substrates. J Biol Chem 274: 22131-22134.

Ezer S, Bayés M, Elomaa O, Schlessinger D & Kere J (1999) Ectodysplasin is a collagenous trimeric type II membrane protein with tumour necrosis factor-like domain and co-localizes with cytoskeletal structures at lateral and apical surfaces of cells. Hum Mol Genet 8: 2079-2086.

Fichard A, Kleman J-P & Ruggiero F (1994) Another look at collagen V and XI molecules. Matrix Biol 14: 515-531.

Fire A, Harrison SW & Dixon D (1990) A modular set of lacZ fusion vectors for studying gene expression in *Caenorhabditis elegans*. Gene 93: 189-198.

Fitzgerald J & Bateman JF (2001) A new FACIT of the collagen family: COL21A1. FEBS Lett 505: 275-280.

Floros J, Steinbrink R, Jacobs K, Phelps D, Kriz R, Recny M, Sultzman L, Jones S, Taeusch HW & Frank HA (1986) Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactant-associated protein. J Biol Chem 261: 9029-9033.

Fraser RBD & McRae TP (1973) Conformation in fibrous proteins. Academic Press, New York.

Freedman RB (1989) Protein disulfide isomerase: multiple roles in the modification of nascent secretory proteins. Cell 57: 1069-1072.

Freedman RB (1994) Protein folding. Folding helpers and unhelpful folders. Curr Biol 4: 933-935.

Freedman RB, Hawkins HC & McLaughlin SH (1995) Protein disulfide-isomerase. Methods Enzymol 251: 397-406.

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

Fukai N, Apte S & Olsen BR (1994) Nonfibrillar collagens. Meth Enzymol 245: 3-28.

Gaill F, Wiedermann H, Mann K, Kühn K, Timpl R & Engel J (1991) Molecular characterization of cuticle and interstitial collagens from worms collected at deep sea hydrothermal vents. J Mol Biol 5: 209-223

Gill AC, Ritchie MA, Hunt LG, Steane SE, Davies KG, Bocking SP, Rhie AG, Bennett AD & Hope J (2000) Post-translational hydroxylation at the N-terminus of the prion protein reveals presence of PPII structure *in vivo*. EMBO J 19: 5324-5331.

Graham PL, Johnson JJ, Wang S, Sibley MH, Gupta MC & Kramer JM (1997) Type IV collagen is detectable in most, but not all, basement membranes of *Caenorhabditis elegans* and assembles on tissues that do not express it. J Cell Biol 137: 1171-1183.

Guo XD, Johnson JJ & Kramer JM (1991) Embryonic lethality caused by mutations in basement membrane collagen of *C. elegans*. Nature 349: 707-709.

Gupta MC, Graham PL & Kramer JM (1997) Characterization of α1(IV) collagen mutations in *Caenorhabditis elegans* and the effects of α1(IV) and α2(IV) mutations on type IV collagen distribution. J Cell Biol 137: 1185-1196.

Guthapfel R, Gueguen P & Quemeneur E (1996) ATP binding and hydrolysis by the multifunctional protein disulfide isomerase. J Biol Chem 271: 2663-2666.

Günzler V, Hanauske-Abel HM, Myllylä R, Mohr J & Kivirikko KI (1987) Time-dependent inactivation of chick-embryo prolyl 4-hydroxylase by coumalic acid. Evidence for a syncatalytic mechanism. Biochem J 242: 163-169.

Hanauske-Abel (1991) Prolyl 4-hydroxylase, a target enzyme for drug development. Design of suppressive agents and the in vitro effects of inhibitors and proinhibitors. J Hepatol 13 Suppl 3: S8-15.

Hanauske-Abel HM & Günzler V (1982) A stereochemical concept for the catalytic mechanism of prolyl 4-hydroxylase: applicability to classification and design of inhibitors. J Theor Biol 94: 421-455.

Hashimoto T, Wakabayashi T, Watanabe A, Kowa H, Hosoda R, Nakamura A, Kanazawa I, Arai T, Takio K, Mann DAM & Iwatsubo T (2002) CLAC: a novel Alzheimer amyloid plaque component derived from a transmembrane precursor, CLAC-P/collagen type XXV. EMBO J 21: 1524-1534.

Hawkins HC & Freedman RB (1991) The reactivities and ionization properties of the active-site dithiol groups of mammalian protein disulfide-isomerase. Biochem J 275 (Pt 2): 335-339.

Hayano T & Kikuchi M (1995) Cloning and sequencing of the cDNA encoding human P5. Gene 164: 377-378.

Hebda PA, Ebert J, Chou KL, Shields M & Kao WW (1983) The association between prolyl 4-hydroxylase metabolism and cell growth in cultured L-929 fibroblasts. Biochim Biophys Acta 758: 128-134.

Heikkinen J, Hautala T, Kivirikko KI & Myllylä R (1994) Structure and expression of the human lysyl hydroxylase gene (PLOD): introns 9 and 16 contain *Alu* sequences at the sites of recombination in Ehlers-Danlos syndrome VI patients. Genomics 24: 464-471.

Heikkinen J, Toppinen T, Yeowell H, Krieg T, Steinmann B, Kivirikko KI & Myllylä R (1997) Duplication of seven exons in the lysyl hydroxylase gene is associated with longer forms of repetitive sequence within the gene and is a common cause for the type VI variant of Ehlers-Danlos syndrome. Am J Hum Gen 60: 48-56.

Helaakoski T, Annunen P, Vuori K, MacNeil IA, Pihlajaniemi T & Kivirikko KI (1995) Cloning, baculovirus expression, and characterization of a second mouse prolyl 4-hydroxylase α subunit isoform: formation of an α_2 β_2 tetramer with the protein disulfide-isomerase/β subunit. Proc Natl Acad Sci USA 92: 4427-4431.

Helaakoski T, Pajunen L, Kivirikko KI & Pihlajaniemi T (1990) Increases in mRNA concentrations of the α and β subunits of prolyl 4-hydroxylase accompany increased gene expression of type IV collagen during differentiation of mouse F9 cells. J Biol Chem 265: 11413-11416.

Helaakoski T, Veijola J, Vuori K, Rehn M, Chow LT, Taillon-Miller P, Kivirikko KI & Pihlajaniemi T (1994) Structure and expression of the human gene for the α subunit of prolyl 4-hydroxylase. The two alternatively spliced types of mRNA correspond to two homologous exons the sequences of which are expressed in a variety of tissues. J Biol Chem 269: 27847-27854.

Helaakoski T, Vuori K, Myllylä R, Kivirikko KI & Pihlajaniemi T (1989) Molecular cloning of the α subunit of human prolyl 4-hydroxylase: the complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. Proc Natl Acad Sci USA 86: 4392-4396.

Hieta R & Myllyharju J Cloning and characterization of a low-molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*. Effective hydroxylation of proline-rich, collagen-like and HIFα -like peptides. J Biol Chem, in press.

Hill KL, Harfe BD, Dobbins CA & L'Hernault SW (2000) *dpy-18* encodes an α subunit of prolyl 4-hydroxylase in *Caenorhabditis elegans*. Genetics 155: 1139-1148.

Hirsh D, Cox GN & Kramer JM (1987) Gene expression during formation of the cuticle of *Caenorhabditis elegans*. Curr Probl Dermatol 17: 19-31.

Hirsh D, Cox GN, Kramer JM, Stinchcomb D & Jefferson R (1985) Structure and expression of the collagen genes of *C. elegans*. Ann N Y Acad Sci 460: 163-171.

Hodgkin J (1999) Sex, cell death and the genome of *C. elegans*. Cell 98: 277-280.

Hodgkin J & Herman RK (1998) Changing styles in *C. elegans* genetics. Trends Genet 14: 352-357.

Hodgkin J, Plasterk RH & Waterston RH (1995) The nematode *Caenorhabditis elegans* and its genome. Science 270: 410-414.

Hoffman MA, Ohh M, Yang H, Klco JM, Ivan M & Kaelin WG Jr (2001) von Hippel-Lindau protein mutants linked to type 2C VHL disease preserve ability to downregulate HIF. Hum Mol Gen 10: 1019-1027.

Hopkinson I, Smith SA, Donne A, Gregory H, Franklin TJ, Grant ME & Rosamond J (1994) The complete cDNA derived sequence of the rat prolyl 4-hydroxylase α subunit. Gene 149: 391-392.

Horelli-Kuitunen N, Kvist A-P, Helaakoski T, Kivirikko KI, Pihlajaniemi T & Palotie A (1997) The order and transcriptional orientation of the human COL13A1 and P4HA genes on chromosome 10 long arm determined by high-resolution FISH. Genomics: 46: 299-302.

Huang LE, Gu J, Schau M & Bunn HF (1998) Regulation of hypoxia-inducible factor 1α is mediated by an O_2- dependent degradation domain via the ubiquitin-proteasome pathway. Proc Natl Acad Sci USA 95: 7987-7992.

Huber S, Van der Rest M, Bruckner P, Rodriguez E, Winterhalter KH & Vaughan L (1986) Identification of the type IX collagen polypeptide chains. The α2(IX) polypeptide carries the chondroitin sulfate chain(s). J Biol Chem 261: 5965-5968.

Hulmes DJS (1992) The collagen superfamily - diverse structures and assemblies. Essays Biochem 27: 49-67.

Hutter H, Vogel BE, Plenefisch JD, Norris CR, Proenca RB, Spieth J, Guo C, Mastwal S, Zhu X, Scheel J & Hedgecock EM (2000) Conservation and novelty in the evolution of cell adhesion and extracellular matrix genes. Science 287: 989-994.

Hynes RO & Zhao Q (2000) The evolution of cell adhesion. J Cell Biol 150: 89-95.

Höyhtyä M, Myllylä R, Piuva J, Kivirikko KI & Tryggvason K (1984) Monoclonal antibodies to human prolyl 4-hydroxylase. Eur J Biochem 141: 472-482.

Ichijo H, Hellman U, Wernstedt C, Gonez LJ, Claesson-Welsh L, Heldin CH & Miyazono K (1993) Molecular cloning and characterization of ficolin, a multimeric protein with fibrinogen- and collagen-like domains. J Biol Chem 268: 14505-14513.

Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, Salic A, Asara JM, Lane WS & Kaelin WG, Jr. (2001) HIF-1α targeted for VHL-mediated destruction by proline hydroxylation: implications for O_2 sensing. Science 292: 464-468.

Jaakkola P, Mole DR, Tian YM, Wilson MI, Gielbert J, Gaskell SJ, Kriegsheim A, Hebestreit HF, Mukherji M, Schofield CJ, Maxwell PH, Pugh CW & Ratcliffe PJ (2001) Targeting of HIF-1α to the von Hippel-Lindau ubiquitylation complex by O_2-regulated prolyl hydroxylation. Science 292: 468-472.

Jacenko O & Tuan RS (1986) Calcium deficiency induces expression of cartilage-like phenotype in chick embryonic calvaria. Dev Biol 115: 215-232.

Jacenko O, San Antonio JD & Tuan RS (1995) Chondrogenic potential of chick embryonic calvaria: II. Matrix calcium may repress cartilage differentiation. Dev Dyn 202: 27-41.

Jenkins CL & Raines RT (2002) Insights of the conformational stability of collagen. Nat Prod Rep 19: 49-59.

Jensenius JC, Laursen SB, Zheng Y & Holmskov U (1994) Conglutinin and CL-43, two collagenous C-type lectins (collectins) in bovine serum. Biochem Soc Trans 22: 95-100.

Jewell UR, Kvietikova I, Scheid A, Bauer C, Wenger RH & Gassmann M (2001) Induction of HIF-1α in response to hypoxia is instantaneous. FASEB J 15: 1312-1314.

Jimenez SA, Harch M & Rosenbloom J (1973) Hydroxyproline stabilizes the triple-helix of chick tendon collagen. Biochem Biophys Res Commun 52: 106-114.

John DC & Bulleid NJ (1994) Prolyl 4-hydroxylase: defective assembly of α subunit mutants indicates that assembled α subunits are intramolecularly disulfide bonded. Biochemistry 33: 14018-14025.

John DC & Bulleid NJ (1996) Intracellular dissociation and reassembly of prolyl 4-hydroxylase: the α subunits associated with the immunoglobulin-heavy-chain binding protein (BiP) allowing reassembly with the β subunit. Biochem J 317 (Pt 3): 659-665.

John DC, Grant ME & Bulleid NJ (1993) Cell-free synthesis and assembly of prolyl 4-hydroxylase: the role of the β subunit (PDI) in preventing misfolding and aggregation of the α subunit. EMBO J 12: 1587-1595.

Johnstone IL (1994) The cuticle of the nematode *Caenorhabditis elegans*: a complex collagen structure. BioEssays 16: 171-178.

Johnstone IL (2000) Cuticle collagen genes. Expression in *Caenorhabditis elegans*. Trends Genet 16: 21-27.

Johnstone IL & Barry JD (1996) Temporal reiteration of a precise gene expression pattern during nematode development. EMBO J 15: 3633-3639.

Kadler KE, Holmes DF, Trotter JA & Chapman JA (1996) Collagen fibril formation. Biochem J 316: 1-11.

Kao WW, Prockop DJ & Berg RA (1979) Kinetics for the secretion of nonhelical procollagen by freshly isolated tendon cells. J Biol Chem 254: 2234-2243.

Kedersha NL, Tkacz JS & Berg RA (1985) Biosynthesis of prolyl hydroxylase: evidence for two separate dolichol- media pathways of glycosylation. Biochemistry 24: 5960-5967.

Kemmink J, Darby NJ, Dijkstra K, Nilges M & Creighton TE (1997) The folding catalyst protein disulfide isomerase is constructed of active and inactive thioredoxin molecules. Curr Biol 4: 239-245.

Kemmink J, Dijkstra K, Mariani M, Scheek RM, Penka E, Nilges M & Darby NJ (1999) The structure in solution of the *b* domain of protein disulfide isomerase. J Biomol NMR 4: 357-368.

Kielty CM, Hopkinson I & Grant ME (1993) Collagen: The collagen family: Structure, assembly and organization in the extracellular matrix. In: Royce, P M & Steinmann, B (eds) Connective Tissue and Its Heritable Disorders. Molecular, Genetics, and Medical Aspects. Wiley-Liss, Inc., New York, p 103-147.

Kivirikko KI (1993) Collagens and their abnormalities in a wide spectrum of diseases. Ann Med 25: 113-126.

Kivirikko KI (1995) Posttranslational processing of collagens. In: Bittar EE, Bittar N & Greenwich CT (eds) Principles of Medical Biology, Vol 3. Cellular organelles the Extracellular Matrix. JAI Press, Greenwich, Conn, p 233-254.

Kivirikko KI & Myllyharju J (1998) Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. Matrix Biol 16: 357-368.

Kivirikko KI & Myllylä R (1980) The enzymology of post-translational modifications of proteins. In: Freedman RB & Hawkins HC (eds). Academic, London, p 53-104.

Kivirikko KI & Myllylä R (1982) Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. Methods Enzymol 82: 245-304.

Kivirikko KI, Myllylä R & Pihlajaniemi T (1992) Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In: Harding, J J & Crabbe, M J C (eds) Post-Translational Modifications of Proteins. CRC Press, Inc., Boca Raton, Florida, p1-51.

Kivirikko KI & Pihlajaniemi T (1998) Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. Adv Enzymol Relat Areas Mol Biol 72: 325-398.

Kivirikko KI & Prockop DJ (1967) Hydroxylation of proline in synthetic polypeptides with purified protocollagen hydroxylase. J Biol Chem 242: 4007-4012.

Koch M, Foley JE, Hahn R, Zhou P, Burgeson RE, Gerecke DR & Gordon MK (2001) α1(XX) collagen, a new member of the collagen superfamily, fibril-associated collagens with interrupted triple helices. J Biol Chem 276: 23120-23126.

Kodama T, Freeman M, Rohrer L, Zabrecky J, Matsudaira P & Krieger M (1990) Type I macrophage scavenger receptor contains α-helical and collagen-like coiled coils. Nature 343: 531-535.

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

Koivunen P, Helaakoski T, Annunen P, Veijola J, Räisänen S, Pihlajaniemi T & Kivirikko KI (1996) ERp60 does not substitute for protein disulfide isomerase as the β subunit of prolyl 4-hydroxylase. Biochem J 316 (Pt 2): 599-605.

Koivunen P, Pirneskoski A, Karvonen P, Ljung J, Helaakoski T, Notbohm H & Kivirikko KI (1999) The acidic C-terminal domain of protein disulfide isomerase is not critical for the enzyme subunit function or for the chaperone or disulfide isomerase activities of the polypeptide. EMBO J 18: 65-74.

Kramer JM (1994a) Structures and functions of collagens in *Caenorhabditis elegans*. FASEB J 8: 329-336.

Kramer JM (1994b) Genetic analysis of extracellular matrix in *C. elegans*. Annu Rev Genet 28: 95-116.

Kramer JM (1997) Extracellular matrix in *C. elegans*, Vol II. In: Riddle DL *et al.* (eds), Cold spring harbor, NY p 471-500.

Kramer JM, Cox GN & Hirsh D (1985) Expression of the *Caenorhabditis elegans* collagen genes *col-1* and *col-2* is developmentally regulated. J Biol Chem 260: 1945-1951.

Kramer JM, French RP, Park EC & Johnson JJ (1990) The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. Mol Cell Biol 10: 2081-2089.

Kramer RZ, Bella J, Brodsky B & Berman H (2001) The crystal and molecular structure of a collagen like peptide with biologically relevant sequence. J Mol Biol 3: 131-143.

Kramer RZ, Vitagliano L, Bella J, Berisio R, Mazzarella L, Brodsky B, Zagari A & Berman HM (1998) X-ray crystallographic determination of a collagen-like peptide with the repeating sequence (Pro-Pro-Gly). J Mol Biol 280: 623-638.

Krishna Rao AS & Hausman RE (1993) cDNA for R-cognin: homology with a multifunctional protein. Proc Natl Acad Sci USA 90: 2950-2954.

Kukkola L, Hieta R, Kivirikko K & Myllyharju J (2000) Cloning a novel α(III) subunit of human prolyl 4-hydroxylase. FECTS meeting, Patras, Greece, 1: E9.

Kuo CJ, LaMontagne KR, Garcia-Gardenã G, Ackley BD, Kalman D, Park S, Christofferson R, Kamihara J, Yuan-Hua D, Kin-Ming L, Gillies S, Folkman J, Mulligan RC & Javaherian K (2001) Oligomerization-dependent regulation of motility and morphogenesis by the collagen XVIII NC1/endostatin domain. J Cell Biol 6: 1233-1246.

Kurz EM, Holstein TW, Petri BM, Engel J & David CN (1991) Mini-collagens in *Hydra* nematocytes. J Cell Biol 115: 1159-1169.

Lamandé SR & Bateman JF (1999) Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones. Semin Cell Dev Biol 10: 455-464.

LaMantia ML & Lennarz WJ (1993) The essential function of yeast protein disulfide isomerase does not reside in its isomerase activity. Cell 74: 899-908.

Lamberg A, Jauhiainen M, Metso J, Enholm C, Shoulders C, Scott J, Pihlajaniemi T & Kivirikko KI (1996) The role of protein disulfide isomerase in the microsomal triacylglycerol transfer protein does not reside in its isomerase activity. Biochem J 315 (Pt 2): 533-536.

Lamberg A, Pihlajaniemi T & Kivirikko KI (1995) 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: 9926-9931.

Lando D, Peet DJ, Whelan DA, Gorman JJ & Whitelaw ML (2002) Asparagine hydroxylation of the HIF transactivation domain: A hypoxic switch. Science 295: 858-861.

Lebeche D, Lucero HA & Kaminer B (1994) Calcium binding properties of rabbit liver protein disulfide isomerase. Biochem Biophys Res Commun 202: 556-561.

Lee EH, Kao WW & Schwarz RI (2001) Cell density regulates prolyl 4-hydroxylase activity independent of mRNA levels. Matrix Biol 19: 779-782.

Lloid MD, Lee HJ, Harlos K, Zhang ZH, Baldwin JE, Schofield CJ, Charnock JM, Garner CD, Hara T, van Scheltinga ACT, Valegård K, Viklund JAC, Hajdu J, Andersson I, Danielsson A & Bhikhabhai RJ (1999) Studies on the active site of deactoxycephalosporin C synthase. J Mol Biol 287: 943-960.

Lu J, Willis AC & Reid KB (1992) Purification, characterization and cDNA cloning of human lung surfactant protein D. Biochem J 284 (Pt 3): 795-802.

Maeda K, Okubo K, Shimomura I, Funahashi T, Matsuzawa Y & Matsubara K (1996) cDNA cloning and expression of a novel adipose specific collagen-like factor apM1 (AdiPose Most abundant Gene transcript 1). Biochem Biophys Res Commun 221: 286-289.

Majamaa K, Hanauske-Abel HM, Günzler V & Kivirikko KI (1984) The 2-oxoglutarate binding site of prolyl 4-hydroxylase. Identification of distinct subsites and evidence for 2-oxoglutarate decarboxylation in a ligand reaction at the enzyme- bound ferrous ion. Eur J Biochem 138: 239-245.

Majamaa K, Kuutti-Savolainen ER, Tuderman L & Kivirikko KI (1979) Turnover of prolyl hydroxylase tetramers and the monomer-size protein in chick-embryo cartilaginous bone and lung *in vivo*. Biochem J 178: 313-322.

McCormick D, Van der Rest M, Goodship J, Lozano G, Ninomiya Y & Olsen BR (1987) Structure of the glycosaminoglycan domain in the type IX collagen-proteoglycan. Proc Natl Acad Sci USA 84: 4044-4048.

McLaughlin SH & Bulleid NJ (1998) Molecular recognition in procollagen chain assembly. Matrix Biol 16: 369-377.

Mello CC, Kramer JM, Stinchcomb D & Ambros V (1991) Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J 10: 3959-3970.

Mendler M, Eich-Bender SG, Vaughan L, Winterhalt KH & Bruckner P (1989) Cartilage contains mixed fibrils of collagen types II, IX, and XI. J. Cell Biol 108: 191-197.

Mori H, Shibasaki T, Uozaki Y, Ochiai K & Ozaki A (1996) Detection of novel proline 3-hydroxylase activities in *Streptomyces* and *Bacillus spp.* by regio- and stereospecific hydroxylation of L-proline. Appl Environ Microbiol 62: 1903-1907.

Mori H, Shibasaki T, Yano K & Ozaki A (1997) Purification and cloning of a proline 3-hydroxylase, a novel enzyme which hydroxylates free L-proline to *cis*-3-hydroxy-L-proline. J Bacteriol 179: 5677-5683.

Murray LW & Tanzer ML (1983) Characterization of a large fragment from annelid cuticle collagen and its relationship to the intact molecule. Coll Relat Res 3: 445-458.

Myllyharju J & Kivirikko KI (1997) Characterization of the iron- and 2-oxoglutarate-binding sites of human prolyl 4-hydroxylase. EMBO J 16: 1173-1180.

Myllyharju J & Kivirikko KI (1999) Identification of a novel proline-rich peptide-binding domain in prolyl 4-hydroxylase. EMBO J 18: 306-312.

Myllyharju J & Kivirikko KI (2001) Collagens and collagen-related diseases. Ann Med 33: 7-21.

Myllylä R, Alitalo K, Vaheri A & Kivirikko KI (1981) Regulation of collagen post-translational modification in transformed human and chick-embryo cells. Biochem J 196: 683-692.

Myllylä R, Kuutti-Savolainen ER & Kivirikko KI (1978) The role of ascorbate in the prolyl hydroxylase reaction. Biochem Biophys Res Commun 83: 441-448.

Myllylä R, Majamaa K, Günzler V, Hanauske-Abel HM & Kivirikko KI (1984) Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. J Biol Chem 259: 5403-5405.

Myllylä R, Pajunen L & Kivirikko KI (1988) Polyclonal and monoclonal antibodies to human lysyl hydroxylase and studies on the molecular heterogeneity of the enzyme. Biochem J 253: 489-496.

Myllylä R, Tuderman L & Kivirikko KI (1977) Mechanism of the prolyl hydroxylase reaction 2. Kinetic analysis of the reaction sequence. Eur J Biochem 80: 349-357.

Nagata K (1996) Hsp47: a collagen specific molecular chaperone. Trends Biochem Sci 21: 22-26.

Nagata K (1998) Expression and function of heat shock protein 47: a collagen-specific molecular chaperone in the endoplasmic reticulum. Matrix Biol 16: 379-386.

Nambu JR, Chen W, Hu S & Crews ST (1996) The *Drosophila melanogaster* similar bHLH-PAS gene encodes a protein related to human hypoxia-inducible factor 1α and *Drosophila* single-minded. Gene 172: 249-254.

Noiva R (1994) Enzymatic catalysis of disulfide formation. Protein Expr Purif 5: 1-13.

Noiva R (1999) Protein disulfide isomerase: the multifunctional redox chaperone of the endoplasmic reticulum. Semin Cell Dev Biol 10: 481-493.

Noiva R, Freedman RB & Lennarz WJ (1993) Peptide binding to protein disulfide isomerase occurs at a site distinct from the active sites. J Biol Chem 268: 19210-19217.

Norman KR & Moerman DG (2000) The let-268 locus of *Caenorhabditis elegans* encodes a procollagen lysyl hydroxylase that is essential for type IV collagen secretion. Dev Biol 227: 690-705.

Ohh M, Park CW, Ivan M, Hoffman MA, Kim TY, Huang LE, Pavletich N, Chau V & Kaelin WG (2000) Ubiquitination of hypoxia-inducible factor requires direct binding to the β-domain of the von Hippel-Lindau protein. Nat Cell Biol 2: 423-427.

O´Reilly MS, Boehm T, Shing Y, Fukai N, Vasios G, Lane WS, Flynn E, Birkheak JR, Olsen BR & Folkman J (1997) Endostatin: An endogenous inhibitor of angiogenesis and tumour growth. Cell 88: 277-285.

Otsu M, Omura F, Yoshimori T & Kikuchi M (1994) Protein disulfide isomerase associates with misfolded human lysozyme *in vivo*. J Biol Chem 269: 6874-6877.

Page AP (1997) Cyclophilin and protein disulfide isomerase genes are co-transcribed in a functionally related manner in *Caenorhabditis elegans*. DNA Cell Biol 11: 1335-1343.

Pajunen L, Jones TA, Goddard A, Sheer D, Solomon E, Pihlajaniemi T & Kivirikko KI (1991) Regional assignment of the human gene coding for a multifunctional polypeptide (P4HB) acting as the β-subunit of prolyl 4-hydroxylase and the enzyme protein disulfide isomerase to 17q25. Cytogenet Cell Genet 56: 165-168.

Pajunen L, Jones TA, Helaakoski T, Pihlajaniemi T, Solomon E, Sheer D & Kivirikko KI (1989) Assignment of the gene coding for the α •subunit of prolyl 4-hydroxylase to human chromosome region 10q21.3-23.1. Am J Hum Genet 45: 829-834.

Pajunen L, Myllylä R, Helaakoski T, Pihlajaniemi T, Tasanen K, Hoyhtya M, Tryggvason K, Solomon E & Kivirikko KI (1988) Assignment of the gene coding for both the β •subunit of prolyl 4-hydroxylase and the enzyme disulfide isomerase to human chromosome region 17p11-qter. Cytogenet Cell Genet 47: 37-41.

Pelham HR (1990) The retention signal for soluble proteins of the endoplasmic reticulum. Trends Biochem Sci 15: 483-486.

Pelicci G, Dente L, De Giuseppe A, Verducci-Galletti B, Giuli S, Mele S, Vetriani C, Giorgio M, Pandolfi PP, Cesareni G & Pelicci PG (1996) A family of Shc related proteins with conserved PTB, CH1 and SH2 regions. Oncogene 13: 633-641.

Pelicci G, Lanfrancone L, Grignani F, McGlade J, Cavallo F, Forni G, Nicoletti I, Grignani F, Pawson T & Pelicci PG (1992) A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. Cell 70: 93-104.

Peterkofsky B & Assad R (1976) Submicrosomal localization of prolyl 4-hydroxylase from chick embryo limb bone. J Biol Chem 251: 4770-4777.

Pihlajaniemi T, Helaakoski T, Tasanen K, Myllylä R, Huhtala ML, Koivu J & Kivirikko KI (1987) Molecular cloning of the β •subunit of human prolyl 4-hydroxylase. This subunit and protein disulfide isomerase are products of the same gene. EMBO J 6: 643-649.

Pihlajaniemi T, Myllylä R, Alitalo K, Vaheri A & Kivirikko KI (1981) Posttranslational modifications in the biosynthesis of type IV collagen by a human tumour cell line. Biochemistry 20: 7409-7415.

Pihlajaniemi T & Rehn M (1995) Two new collagen subgroups: Membrane associated collagens and types XV and XVIII. Proc Nucleic Acid Res Mol Biol 50: 255-262.

Pirneskoski A, Ruddock LW, Klappa P, Freedman RB, Kivirikko KI & Koivunen P (2001) Domains *b´* and *a´* of protein disulfide isomerase fulfill the minimum requirement for function as a subunit of prolyl 4-hydroxylase. J Biol Chem 276: 11287-11293.

Popescu NC, Cheng SY & Pastan I (1988) Chromosomal localization of the gene for a human thyroid hormone- binding protein. Am J Hum Genet 42: 560-564.

Pousi B, Hautala T, Heikkinen J, Pajunen L, Kivirikko KI & Myllyla R (1994) *Alu-Alu* recombination results in a duplication of seven exons in the lysyl hydroxylase gene in a patient with the type VI variant of Ehlers- Danlos syndrome. Am J Hum Genet 55: 899-906.

Prockop DJ & Kivirikko KI (1995) Collagens: molecular biology, diseases, and potentials for therapy. Annu Rev Biochem 64: 403-434.

Puig A & Gilbert HF (1994) Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme. J Biol Chem 269: 7764-7771.

Quan H, Fan G & Wang CC (1995) Independence of the chaperone activity of protein disulfide isomerase from its thioredoxin-like active site. J Biol Chem 270: 17078-17080.

Quemeneur E, Guthapfel R & Gueguen P (1994) A major phosphoprotein of the endoplasmic reticulum is protein disulfide isomerase. J Biol Chem 269: 5485-5488.

Rautavuoma K, Passoja K, Helaakoski T & Kivirikko KI (2000) Complete exon-intron organization of the gene for human lysyl hydroxylase (LH3). Matrix Biol 1: 73-79.

Rappleye CA, Paredez AR, Smith CW, McDonald KL & Aroian RV (1999) The coronin-like protein POD-1 is required for anterior-posterior axis formation and cellular architecture in the nematode *Caenorhabditis elegans*. Gen Dev 13: 2838-2851.

Reid KB (1979) Complete amino acid sequences of the three collagen-like regions present in subcomponent C1q of the first component of human complement. Biochem J 179: 367-371.

Reid KBM (1993) Structure/function relationships in the collectins (mammalian lectins containing collagen-like regions). Biochem Soc Transact 21: 464-468.

Rhoads RE & Udenfriend S (1968) Decarboxylation of α-ketoglutarate coupled to collagen proline hydroxylase. Proc Natl Acad Sci USA 60: 1473-1478.

Ricard-Blum S, Dublet B & van der Rest M (2000) Unconventional collagens. Types VI, VII, VIII, IX, X, XIV, XVI and XIX. Oxford: Oxford University Press.

Roach HI, Erenpreisa J & Aigner T (1995a) Osteogenic differentiation of hypertrophic chondrocytes involves asymmetric cell divisions and apoptosis. J Cell Biol 131: 483-494.

Roach PL, Clifton IJ, Fulop V, Harlos K, Barton GJ, Hajdu J, Andersson I, Schofield CJ & Baldwin JE (1995b) Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature 375: 700-704.

Roach PL, Clifton IJ, Hensgens CM, Shibata N, Schofield CJ, Hajdu J & Baldwin JE (1997) Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. Nature 387: 827-830.

Rogalski TM, Williams BD, Mullen GP & Moerman DG (1993) Products of the *unc-52* gene in *Caenorhabditis elegans* are homologous to the core protein of the mammalian basement membrane heparan sulfate proteoglycan. Genes Dev 7: 1471-84.

Roguska MA & Gudas LJ (1985) An increase in prolyl 4-hydroxylase activity occurs during the retinoic acid-induced differentiation of mouse teratocarcinoma stem cell lines F9 and P19. J Biol Chem 260: 13893-13896.

Rohrer L, Freeman M, Kodama T, Penman M & Krieger M (1990) Coiled-coil fibrous domains mediate ligand binding by macrophage scavenger receptor type II. Nature 343: 570-572.

Rosenberry TL, Barnett P & Mays C (1982) Acetylcholinesterase. Methods Enzymol 82 Pt A: p 325-339.

Rosenbloom J, Abrams WR & Mecham R (1993) Extracellular matrix 4: The elastic fiber. FASEB J 13: 1208-1218.

Rudiger NS, Gregersen N & Kielland-Brandt MC (1995) One short well conserved region of *Alu*-sequences is involved in human gene rearrangements and has homology with prokaryotic chi. Nucleic Acids Res 23: 256-260.

Rupp K, Birnbach U, Lundstrom J, Van PN & Soling HD (1994) Effects of CaBP2, the rat analog of ERp72, and of CaBP1 on the refolding of denatured reduced proteins. Comparison with protein disulfide isomerase. J Biol Chem 269: 2501-2507.

Sambrook J, Fritch EF & Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger F, Nicklen S & Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467.

Semenza GL (2001) Hypoxia-inducible factor 1: oxygen homeostasis and disease patophysiology. Trends Mol Med 8: 345-350.

Semenza GL (2002) Involvement in hypoxia-inducible factor 1 in human cancer. Intern Med 2: 79-89.

Sibley MH, Johnson JJ, Mello CC & Kramer JM (1993) Genetic identification, sequence, and alternative splicing of the *Caenorhabditis elegans* α2(IV) collagen gene. J Cell Biol 123: 255-264.

Srivastava AK, Pispa J, Hartung AJ, Du Y, Ezer S, Jenks T *et al*. (1997) The tabby phenotype is caused by mutation in a mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. Proc Natl Acad Sci USA 94: 13069-13074.

Stassen FL, Cardinale GJ & Udenfriend S (1973) Activation of prolyl 4-hydroxylase in L-929 fibroblasts by ascorbic acid. Proc Natl Acad Sci USA 70: 1090-1093.

Strout MP, Marcucci G, Bloomfield CD & Caligiuri MA (1998) The partial tandem duplication of ALL1 (MLL) is consistently generated by *Alu*-mediated homologous recombination in acute myeloid leukemia. Proc Natl Acad Sci USA 95: 2390-2395.

Sulston J & Hodgkin J (1988) Methods. In: Wood WB (ed) The nematode *Caenorhabditis elegans.* Cold Spring Harbor Laboratory Press, Plainview, NY, p 587-606.

Sulston JE, Schierenberg E, White JG & Thomson JN (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev Biol 100: 64-119.

Sutter CH, Laughner E & Semenza GL (2000) Hypoxia-inducible factor 1α protein expression is controlled by oxygen-regulated ubiquitination that is disrupted by deletions and missense mutations. Proc Natl Acad Sci USA 97: 4748-4753.

Suzuki H & Terada T (1988) Removal of dodecyl sulfate from protein solution. Anal Biochem 172: 259-263.

Takahashi Y, Takahashi S, Shiga Y, Yoshimi T & Miura T (2000) Hypoxic induction of prolyl 4-hydroxylase α(I) in cultured cells. J Biol Chem 275: 14139-14146.

Tasab M, Batten MR & Bulleid NJ (2000) Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen. EMBO J 19: 2204-2211.

Tasanen K, Oikarinen J, Kivirikko KI & Pihlajaniemi T (1992) Promoter of the gene for the multifunctional protein disulfide isomerase polypeptide. Functional significance of the six CCAAT boxes and other promoter elements. J Biol Chem 267: 11513-11519.

Tasanen K, Oikarinen J, Kivirikko KI & Pihlajaniemi T (1993) Interaction of transcription factor Sp1 with the promoter of the gene for the multifunctional protein disulfide isomerase polypeptide. Biochem J 292: 41-45.

Tasanen K, Parkkonen T, Chow LT, Kivirikko KI & Pihlajaniemi T (1988) Characterization of the human gene for a polypeptide that acts both as the β subunit of prolyl 4-hydroxylase and as protein disulfide isomerase. J Biol Chem 263: 16218-16224.

Timpl R & Engel J (1987) Type VI collagen. In: Mayne, R & Burgeson, RE (eds) Structure and Function of Collagen Types. Academic Press, New York, p 105-143.

Towler DA, Bennett CD & Rodan GA (1994) Activity of the rat osteocalcin basal promoter in osteoblastic cells is dependent upon homeodomain and CP1 binding motifs. Mol Endocrinol 8: 614-624.

Tryggvason K, Majamaa K & Kivirikko KI (1979) Prolyl 3-hydroxylase and 4-hydroxylase activities in certain rat and chick-embryo tissues and age-related changes in their activities in the rat. Biochem J 178: 127-131.

Tuderman L, Kuutti ER & Kivirikko KI (1975) An affinity-column procedure using poly(L-proline) for the purification of prolyl hydroxylase. Purification of the enzyme from chick embryos. Eur J Biochem 52: 9-16.

Tuderman L, Myllylä R & Kivirikko KI (1977) Mechanism of the prolyl hydroxylase reaction. Role of co- substrates. Eur J Biochem 80: 341-348.

Turpeenniemi-Hujanen TM, Puistola U & Kivirikko KI (1980) Isolation of lysyl hydroxylase, an enzyme of collagen synthesis, from chick embryos as a homogenous protein. Biochem J 189: 247-253.

Turpeenniemi-Hujanen TM, Puistola U & Kivirikko KI (1981) Human lysyl hydroxylase: purification to homogeneity, partial characterization and comparison of catalytic properties with those of a mutant enzyme from Ehlers-Danlos syndrome type VI fibroblasts. Coll Relat Res 1: 355-366.

Uitto J & Christiano AM (1992) Molecular genetics of the cutaneuos basement membrane zone. Perspectives on epidermolysis bullosa and other blistering skin diseases. J Clin Invest 90: 687-692.

Valegård K, van Scheltinga AC, Lloyd MD, Hara T, Ramaswamy S, Perrakis A, Thompson A, Lee HJ, Baldwin JE, Schofield CJ, Hajdu J & Andersson I (1998) Structure of a cephalosporin synthase. Nature 394: 805-809.

van der Keyl H, Kim H, Espey R, Oke CV & Edwards MK (1994) *Caenorhabditis elegans sqt-3* mutants have mutations in the *col-1* collagen gene. Dev Dyn 201: 86-94.

Veijola J, Annunen P, Koivunen P, Page AP, Pihlajaniemi T & Kivirikko KI (1996a) Baculovirus expression of two protein disulfide isomerase isoforms from *Caenorhabditis elegans* and characterization of prolyl 4-hydroxylases containing one of these polypeptides as their β subunit. Biochem J 317 (Pt 3): 721-729.

Veijola J, Koivunen P, Annunen P, Pihlajaniemi T & Kivirikko KI (1994) Cloning, baculovirus expression, and characterization of the α subunit of prolyl 4-hydroxylase from the nematode *Caenorhabditis elegans*. This α subunit forms an active αβ dimer with the human protein disulfide isomerase/β subunit. J Biol Chem 269: 26746-26753.

Veijola J, Pihlajaniemi T & Kivirikko KI (1996b) Co-expression of the β subunit of human prolyl 4-hydroxylase with BiP polypeptide in insect cells leads to the formation of soluble and insoluble complexes. Soluble β-subunit-BiP complexes have no prolyl 4-hydroxylase activity. Biochem J 315: 613-618.

Vincent F, Corral M, Defer N & Adolphe M (1991) Effects of oxygen free radicals on articular chondrocytes in culture: c- myc and c-Ha-ras messenger RNAs and proliferation kinetics. Exp Cell Res 192: 333-339.

Vuorela A, Myllyharju J, Nissi R, Pihlajaniemi T & Kivirikko KI (1997) 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: 6702-6712.

Vuorela A, Myllyharju J, Pihlajaniemi T & Kivirikko KI (1999) Coexpression with collagen markedly increases the half-life of the recombinant human prolyl 4-hydroxylase tetramer in the yeast *Pichia pastoris*. Matrix Biol 18: 519-522.

Vuori K, Myllylä R, Pihlajaniemi T & Kivirikko KI (1992a) Expression and site-directed mutagenesis of human protein disulfide isomerase in *Escherichia coli*. This multifunctional polypeptide has two independently acting catalytic sites for the isomerase activity. J Biol Chem 267: 7211-7214.

Vuori K, Pihlajaniemi T, Marttila M & Kivirikko KI (1992b) Characterization of the human prolyl 4-hydroxylase tetramer and its multifunctional protein disulfide-isomerase subunit synthesized in a baculovirus expression system. Proc Natl Acad Sci USA 89: 7467-7470.

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

Wagner K, Pöschl E, Turnay J, Baik J, Pihlajaniemi T, Frischholz S & van der Mark K (2000) Coexpression of α and β subunits of prolyl 4-hydroxylase stabilizes the triple helix of recombinant human type X collagen. Biochem J 352: 907-911.

Walmsley AR, Batten MR, Lad U & Bulleid NJ (1999) Intracellular retention of procollagen within the endoplasmic reticulum is mediated by prolyl 4-hydroxylase. J Biol Chem 274: 14884-14892.

Wang C, Valtavaara M & Myllylä R (2000) Lack of collagen type specificity for lysyl hydroxylase isoforms. DNA Cell Biol 19: 71-77.

Waterston RH (1989) Molecular genetic approaches to the study of motility in *Caenorhabditis elegans*. Cell Motil Cytoskeleton 14: 136-145.

Watt SL, Lunstrum GP, McDonough AM, Keene DR, Burgeson RE & Morris NP (1992) Characterization of collagen types XII and XIV from fetal bovine cartilage. J Biol Chem 267: 20093-20099.

Wells WW, Xu DP, Yang YF & Rocque PA (1990) Mammalian thioltransferase (glutaredoxin) and protein disulfide isomerase have dehydroascorbate reductase activity. J Biol Chem 265: 15361-15364.

Wetterau JR, Combs KA, Spinner SN & Joiner BJ (1990) Protein disulfide isomerase is a component of the microsomal triglyceride transfer protein complex. J Biol Chem 265: 9801-9807.

Wijffels GL, Panaccio M, Salvatore L, Wilson L, Walker ID & Spithill TW (1994) The secreted cathepsin L-like proteinases of the trematode, *Fasciola hepatica*, contain 3-hydroxyproline residues. Biochem J 299: 781-790.Williams BD & Waterston RH (1994) Genes critical for muscle development and function in *Caenorhabditis elegans* identified through lethal mutations. J Cell Biol 124: 475-490.

Winter AD & Page AP (2000) 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: 4084-4093.

Wood WB (1988) Embryology. In: Wood WB (ed) The nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Plainview, NY, p 1-60 and p 215-241.

Wu JJ, Woods PE & Eyre DR (1992) Identification of cross-linking sites in bovine cartilage type IX collagen reveals an antiparallel type II-type IX molecular relationship and type IX to type IX bonding. J Biol Chem 267: 23007-23014.

Yao Y, Zhou Y & Wang C (1997) Both the isomerase and chaperone activities of protein disulfide isomerase are required for the reactivation of reduced and denatured acidic phospholipase α2. EMBO J 16: 651-658.

Yasothornsikul S, Davis WJ, Cramer G, Kimbrell DA & Dearolf CR (1997) *vinking*: identification and characterization of a second type IV collagen in *Drosophila*. Gene 198: 17-25.

Yu F, White SB, Zhao Q & Lee FS (2001) HIF-1α binding to VHL is regulated by stimulus-sensitive proline hydroxylation. Proc Natl Acad Sci USA 98: 9630-9635.

Zenteno-Ruiz JC, Kofman-Alfaro S & Méndez JB (2001) 46, XX Sex reversal. Arch Med Research 32: 559-566

Zhang ZH, Ren JS, Stammers DK, Baldwin JE, Harlos K & Schofield CJ (2000) Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. Nat Struct Biol 7: 127-133.

Zhu Y, Jong NC, Frazer KA, Gong E, Krauss RM, Jan-Fang C, Boffelli D & Rubin EM (2000) Genomic interval engineering of mice identifies a novel modulator of triglyceride production. Proc Natl Acad Sci USA 97: 1137-1142.