

Kati Takaluoma

LYSYL HYDROXYLASES

STUDIES ON RECOMBINANT LYSYL
HYDROXYLASES AND MOUSE LINES LACKING
LYSYL HYDROXYLASE 1 OR LYSYL HYDROXYLASE 3

FACULTY OF MEDICINE,
DEPARTMENT OF MEDICAL BIOCHEMISTRY AND MOLECULAR BIOLOGY,
BIOCENTER OULU,
UNIVERSITY OF OULU

D

MEDICA



ACTA UNIVERSITATIS OULUENSIS
D Medica 920

KATI TAKALUOMA

LYSYL HYDROXYLASES

Studies on recombinant lysyl hydroxylases and mouse lines lacking lysyl hydroxylase 1 or lysyl hydroxylase 3

Academic dissertation to be presented, with the assent of the Faculty of Medicine of the University of Oulu, for public defence in the Auditorium of the Medipolis Research Center (Kiviharjuntie 11), on May 25th, 2007, at 13 p.m.

OULUN YLIOPISTO, OULU 2007

Copyright © 2007
Acta Univ. Oul. D 920, 2007

Supervised by
Docent Johanna Myllyharju

Reviewed by
Doctor Susanne Grässel
Professor Nikolai Miosge

ISBN 978-951-42-8427-4 (Paperback)
ISBN 978-951-42-8428-1 (PDF)
<http://herkules.oulu.fi/isbn9789514284281/>
ISSN 0355-3221 (Printed)
ISSN 1796-2234 (Online)
<http://herkules.oulu.fi/issn03553221/>

Cover design
Raimo Ahonen

OULU UNIVERSITY PRESS
OULU 2007

Takaluoma, Kati, Lysyl hydroxylases. Studies on recombinant lysyl hydroxylases and mouse lines lacking lysyl hydroxylase 1 or lysyl hydroxylase 3

Faculty of Medicine, Department of Medical Biochemistry and Molecular Biology, Biocenter Oulu, University of Oulu, P.O. Box 5000, FI-90014 University of Oulu, Finland

Acta Univ. Oul. D 920, 2007

Oulu, Finland

Abstract

Lysyl hydroxylases (E.C. 1.14.11.4, LHs) have three isoenzymes that are found in humans and mice, and they hydroxylate lysine residues in collagens and other proteins containing collagenous sequences. The hydroxylysines formed are crucial for the intermolecular collagen crosslinks that stabilise collagen fibres, thereby providing the stiffness and stability required by various tissues. In addition, hydroxylysines serve as attachment sites for carbohydrates, whose functions on collagen molecules are not completely understood yet. In humans, lack of LH1 causes Ehlers-Danlos syndrome (EDS) VIA, which is characterised, for example, by severe progressive kyphoscoliosis and muscular hypotonia with joint laxity. Mutations in the LH2 gene are associated with Bruck syndrome, which is characterised by fragile bones with congenital joint contractures.

In the present work recombinant human lysyl hydroxylases were produced in insect cells and purified to homogeneity. Limited proteolysis revealed that LHs consist of at least three structural domains. The N-terminal domain plays no role in the lysyl hydroxylase activity, but instead, is responsible for the recently reported glucosyltransferase activity of LH3, and the galactosyltransferase activity reported here for the first time. The LH polypeptide lacking the N-terminal domain is a fully active LH with K_m values identical to those of full-length enzyme. In addition, direct evidence is shown that LH2, but not LH1 or LH3, hydroxylates the telopeptide lysine residues of fibrillar collagens. All three recombinant LHs were able to hydroxylate the synthetic peptides representing the helical hydroxylation sites in types I and IV collagens, with some differences in the V_{max} and K_m values. In addition, all three LHs hydroxylated the collagenous domain of coexpressed type I procollagen chain to similar extent.

In this study mouse lines lacking LH3 or LH1 were created and analysed. Unexpectedly, the LH3 null mice died during the embryonal period due to fragmentation of basement membranes. Type IV collagen, one of the major components in basement membranes, aggregates on its way to extracellular space and is absent from the basement membranes making them fragile. This is most probably caused by abnormal processing of type IV collagen due to decreased glucosyltransferase activity of the LH3 null embryos.

The first mouse model for human EDS VIA is presented here. The LH1 null mice did not have kyphoscoliosis characteristic of EDS VIA, but showed gait abnormalities due to muscular hypotonia and possible joint laxity, as also seen in EDS VIA patients. In addition, the null mice died occasionally from aortic ruptures. Ultra structural analysis revealed degradation of smooth muscle cells and abnormal collagen fibres even in non-ruptured aortas of LH1 null mice. The hydroxylation of lysine residues and crosslinking in LH1 null mice were also abnormal, as in human EDS VIA patients. The LH1 null mouse line provides an excellent tool for analysing several aspects of human EDS VIA, including muscular hypotonia, abnormalities in collagen fibres and their crosslinking.

Keywords: 2-oxoglutarate 5-dioxygenase, basement membrane, collagen, collagen crosslink, lysyl hydroxylase, procollagen-lysine, transgenic mice

Acknowledgements

This work was carried out at the Department of Medical Biochemistry and Molecular Biology and Biocenter Oulu, University of Oulu during the years 2001-2007. I wish to express my deepest gratitude to my supervisor Docent Johanna Myllyharju for providing excellent conditions for making research, for active and inspiring guidance and very effectively solving the problems encountered. Research Professor Emeritus Kari Kivirikko is acknowledged for his always enthusiastic attitude towards all the projects and for numerous invaluable ideas and comments. Docent Raija Soininen has been crucial for all the mouse experiments and is heartily acknowledged for her never failing positive attitude and for investing so much of time and effort in my projects. Professor Taina Pihlajaniemi, Acting Professor, Head of the department Peppi Karppinen, Professor Leena Ala-Kokko, Professor Seppo Vainio, and Professor Emeritus Ilmo Hassinen are acknowledged for creating positive and collaborative atmosphere in our department making working both effective and fun. All the people at the Department of Medical Biochemistry and Molecular Biology, as well as the Biocenter Oulu staff are warmly acknowledged for their help and effortless collaboration. Special thanks to Risto Helminen for his cheerful help with computer related matters.

I am very grateful to Professor Nicolai Miosge and Dr. Susan Grässler for their careful and fast revision of the manuscript. Dr. Peter Cura is thanked for reviewing the language of the manuscript and also for the comments on the content.

I want to thank all my coauthors for making this work possible. I am especially in debt to Kati Rautavuoma, Ph.D., for introducing me to lysyl hydroxylases and sharing the projects in the early days. Juha Lantto, M.D., is acknowledged for good collaboration and discussions and for making the hours in the lab enjoyable. Marjo Hyry, M. Sc., is thanked for her efforts in these projects and for pleasant and easy collaboration. Docent Raija Sormunen is acknowledged for sharing her vast knowledge with electron microscope. Kaisa Passoja, MD, Ph.D., and Docent Asta Pirskanen are acknowledged for their work on lysyl hydroxylases. Ari-Pekka Kvist, Ph.D., is acknowledged for his help in this work. I am very grateful to Professor Ruud Bank for the help with crosslink analysis and for hosting me when I visited Leiden. Ms. Jessica Snabel is acknowledged for her valuable help and the fun and productive time in the lab while I was in Leiden.

This work would not have been possible without excellent help from Anu Myllymäki, Outi Mänty and Liisa Äijälä. Anu is also acknowledged for helping me with all the practical matters in my early days. I am obliged also to the skilled people at Biocenter Oulu Transgenic core facility, particularly Sanna Kauppinen and Jonna Laakso, for valuable, smooth and productive collaboration. I am also grateful for the easy everyday collaboration with the Laboratory Animal Centre of the University of Oulu, especially with Tanja Sankala.

The current and former members of Johanna's group from Paradise to Hell are warmly acknowledged for collaboration and friendship, thanks to Outi, Laura, Reija, Kati, Päivi, Peppi, Hanna-Leena, Marjo, Maija, Antje, Katriina, Joni, Jaana, Annika, Tuomas, Ellinoora, as well as Ritva, Liisa, Tiina and Juha. It has been a privilege to work with you and I will miss you all. Cheers for Marjo, Outsa and Hanna-Leena as well as Laura, Kati and Tiina for sharing the office with me during the years and for tolerating me in both good and bad days. Particularly Outsa is thanked for numerous hours we have spent together working with research and other kinds of activities.

Finally I want to thank my family and my friends. Thanks to Mika Ilves and Miia Turpeinen for all the fun and relaxation we have had together, and Mika is acknowledged also for numerous advices and help concerning research. Special thanks to Anu Taivainen, Kaisa Posti, Päivi Kuuva, Sanna Marjavaara and Sarianna Åkerlund for their true friendship over the years. My mother Anneli and father Erkki deserve my deepest gratitude for all the help and support. My sisters Niina and Jenni and my brother Jukka with their families have always been very supportive and they are thanked for many enjoyable occasions and for help in many practical matters. Big hugs for Ville, Julius, Neo and Vertti. My beloved husband Jukka is acknowledged for discussions and advices concerning mysteries of proteins, constant help with computers and most importantly for making every day life such a pleasure.

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

Helsinki, February 2007,

Kati Takaluoma

Abbreviations

The standard one-letter and three-letter abbreviations are used for amino acids.

AA	Amino acid
BS	Bruck syndrome
BSA	bovine serum albumin
C	carboxy
cDNA	complementary deoxyribonucleic acid
C-P4H	collagen prolyl 4-hydroxylase
DNA	deoxyribonucleic acid
E	embryonic day
EDS	Ehlers-Danlos syndrome
ER	endoplasmic reticulum
HP	hydroxylysyl pyridinoline
kDa	kilodalton(s)
K_m	Michaelis constant
LH	lysyl hydroxylase
LO	lysine oxidase
LP	lysyl pyridinoline
mRNA	messenger RNA
N	amino
PAGE	polyacryl amide gel electrophoresis
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
<i>PLOD</i>	procollagen-lysine 2-oxoglutarate 5-dioxygenase; human gene
<i>Plod</i>	procollagen-lysine 2-oxoglutarate 5-dioxygenase; mouse gene
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
T_m	midpoint of thermal transition
UDP	uridine diphosphate
V_{max}	maximal velocity

List of original articles

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

- I Rautavuoma K, Takaluoma K, Passoja K, Pirskanen A, Kvist A-P, Kivirikko KI & Myllyharju J (2002) Characterization of three fragments that constitute the monomers of the human lysyl hydroxylase isoenzymes 1-3. *J Biol Chem* 277: 23084-23091.
- II Takaluoma K, Lantto J & Myllyharju J (2007) Lysyl hydroxylase 2 is a specific telopeptide hydroxylase, while all three isoenzymes hydroxylate collagenous sequences. Accepted for publication in *Matrix Biology*
- III Rautavuoma K*, Takaluoma K*, Sormunen R, Myllyharju J, Kivirikko KI & Soininen R (2004) Premature aggregation of type IV collagen and early lethality in lysyl hydroxylase 3 null mice. *Proc Natl Acad Sci USA* 101: 14120-14125.
- IV Takaluoma K, Hyry M, Lantto J, Sormunen R, Bank, RA, Kivirikko KI, Myllyharju J & Soininen R (2007) Tissue-specific changes in the hydroxylysine content and crosslinks of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice. *J Biol Chem* 282: 6588-6596.

*equal contribution

Contents

Abstract	
Acknowledgements	
Abbreviations	
List of original articles	
Contents	
1 Introduction	13
2 Review of the literature	15
2.1 Extracellular matrix and basement membranes	15
2.2 Collagens	16
2.2.1 Collagen types	17
2.2.2 Collagen biosynthesis	19
2.3 Hydroxylysines	22
2.3.1 Functions of hydroxylysines	23
2.3.1.1 Collagen cross-links	23
2.3.1.2 Glycosylation of hydroxylysine	26
2.4 Lysyl Hydroxylases	28
2.4.1 Lysyl hydroxylase isoenzymes	28
2.4.2 Lysyl hydroxylase expression	29
2.4.3 Molecular characteristics of lysyl hydroxylase isoenzymes	31
2.4.4 Reaction mechanism and cosubstrates	33
2.4.5 Substrates and substrate specificity	34
2.4.6 Diseases associated with lysyl hydroxylase deficiency	35
2.4.6.1 The kyphoscoliotic type of Ehlers-Danlos syndrome	35
2.4.6.2 Bruck syndrome	36
2.4.6.3 Fibrotic conditions	37
3 Aims of the study	38
4 Materials and Methods	39
5 Results	41
5.1 Purification of recombinant human lysyl hydroxylases and characterisation of their domain structure and functions of the domains (I)	41

5.2 Substrate specificity of recombinant human lysyl hydroxylases (II).....	43
5.3 Lack of lysyl hydroxylase 3 leads to fragmentation of basement membranes and embryonic lethality (III)	45
5.4 Lack of lysyl hydroxylase 1 in mice leads to increased mortality due to aortic rupture (IV).....	46
6 Discussion	49
6.1 Expression of recombinant lysyl hydroxylases in insect cells and characterisation of their domain structure and function (I)	49
6.2 Substrate specificity of lysyl hydroxylases (II)	51
6.3 Inactivation of <i>Plod3</i> leads to abnormal type IV collagen processing and fragmentation of basement membranes leading to embryonal lethality (III)	53
6.4 <i>Plod1</i> ^{-/-} mice as a disease model for human EDS VIA	55
References	
Original articles	

1 Introduction

Collagens form a large family of proteins that are major components of connective tissues and provide mechanical strength and structural stability for the tissues. Collagen biosynthesis is a complicated process involving several co- and posttranslational enzyme activities, including lysyl hydroxylases (LHs). LHs reside in the lumen of the endoplasmic reticulum and catalyse the hydroxylation of certain lysine residues in collagens and some other proteins with collagenous sequences. The hydroxylysines formed take part in the formation of intermolecular crosslinks that are crucial for the stability of collagen fibres, and serve as attachment sites for carbohydrate moieties. The exact role of the carbohydrate moieties is not yet understood. The extent of lysine hydroxylation and subsequent glycosylation of hydroxylysines varies between collagen types (Kivirikko *et al.* 1992).

There are three LH isoenzymes, LH1-3, in humans, mice and some other species. All three enzymes have been expressed as recombinant enzymes and they hydroxylate lysine residues *in vitro* with highly similar kinetic properties (Krol *et al.* 1996, Pirskanen *et al.* 1996, Valtavaara *et al.* 1997, 1998), and there seems to be no strict substrate sequence specificity with LHs (Risteli *et al.* 2004). The differences in the collagen hydroxylation could perhaps also be due to differences in expression, but currently it seems that all three LHs are quite ubiquitously expressed (Yeowell *et al.* 1994, Heikkinen *et al.* 1994, Valtavaara *et al.* 1997, 1998, 1999, Passoja *et al.* 1998b, Ruotsalainen *et al.* 1999, Yeowell & Walker 1999, Hjalt *et al.* 2001, Salo *et al.* 2006b). In addition to lysyl hydroxylase activity, LH3 also possesses glycosyltransferase activity (Heikkinen *et al.* 2000) and minor galactosyltransferase activity (Wang *et al.* 2002a).

In humans, lack of LH1 causes Ehlers-Danlos syndrome VIA, which is characterised, for example, by severe kyphoscoliosis, hypermobile joints, skin fragility and muscular hypotonia (Yeowell & Walker 2000, Steinmann *et al.* 2002, Myllyharju & Kivirikko 2004). Mutations in the gene coding for LH2 are associated with Bruck syndrome, which is characterised by fragile bones with congenital joint contractures (van der Slot *et al.* 2003, Ha-Vinh *et al.* 2004). In addition, increased expression of LH2 is associated with fibrotic conditions (Brinckmann *et al.* 1999, Uzawa *et al.* 2003, van der Slot *et al.* 2003, 2004, 2005, Pornprasertsuk *et al.* 2004, van der Slot-Verhoeven *et al.* 2005, Wu *et al.* 2006). No human diseases caused by mutations in LH3 are known at the moment.

Although LH was purified to homogeneity for the first time two decades ago (Turpeenniemi-Hujanen *et al.* 1980, 1981), there are still many aspects of the molecular properties and *in vivo* functions to be discovered. The aim of this work was to further increase our knowledge of LH isoenzymes. Recombinantly expressed LHs were used to study the domain structure of LHs and their molecular and catalytic properties, especially their substrate specificity. Two mouse lines lacking either LH1 or LH3 were also generated to identify the *in vivo* functions of LHs and are presented here.

2 Review of the literature

2.1 Extracellular matrix and basement membranes

The extracellular matrix is composed of a variety of proteins and glycosaminoglycans that are mainly attached to proteins thus forming proteoglycans. Glycosaminoglycans have a negative charge and bind cations, for example Na^+ , that in turn bind water molecules. This results in the formation of a gel-like structure that allows diffusion of nutrients, hormones and signalling molecules between blood and tissue cells. The major protein components of the extracellular matrix are collagens, other structural proteins like elastin, and adhesive glycoproteins like fibronectin, laminin, tenascin and nidogen. Connective tissues are specialised types of tissues that contain more extracellular matrix than the cells that they surround. The composition of connective tissue depends on its function: tendons have to resist extension and consist mostly of collagen fibres; cartilage is resistant to compression and contains 50% proteoglycans; bone is rigid and resistant to deformation, which is attained by calcium salts and collagen fibres; the aorta and lungs, on the other hand, are flexible and durable and consist of collagen that provides the tensile strength, and elastin that provides flexibility. Further diversity in different tissues is provided by the different collagen types with different molecular and supramolecular architecture, size and order that result in different functional properties. (Alberts *et al.* 1989, Kielty & Grant 2002.)

Basement membrane is a specialised form of extracellular matrix, a thin basal lamina between the connective tissue and the epithelium or the endothelium. Its function is to maintain the structural integrity of organs and tissues, divide tissues into compartments and regulate cell behaviour. Basement membrane allows diffusion of small molecules, e.g. nutrients and hormones, both into tissues and out of them. The self-assembly of basement membrane is initiated by deposition of a laminin polymer on the cell surface where integrins serve as an important interaction partner. The other major component of basement membranes, type IV collagen, forms network-like structures that are connected to laminins *via* nidogens (Figure 1). In addition, about 50 other proteins, including perlecan, type XV and XVIII collagens and fibulin, are found in basement membranes. The composition of basement membranes is organ-specific. The importance of the

function of basement membrane is clearly illustrated by many disorders caused by lack of basement membrane components. (Kalluri 2003, Sasaki *et al.* 2004, Yurchenco *et al.* 2004, Aszodi *et al.* 2006.)

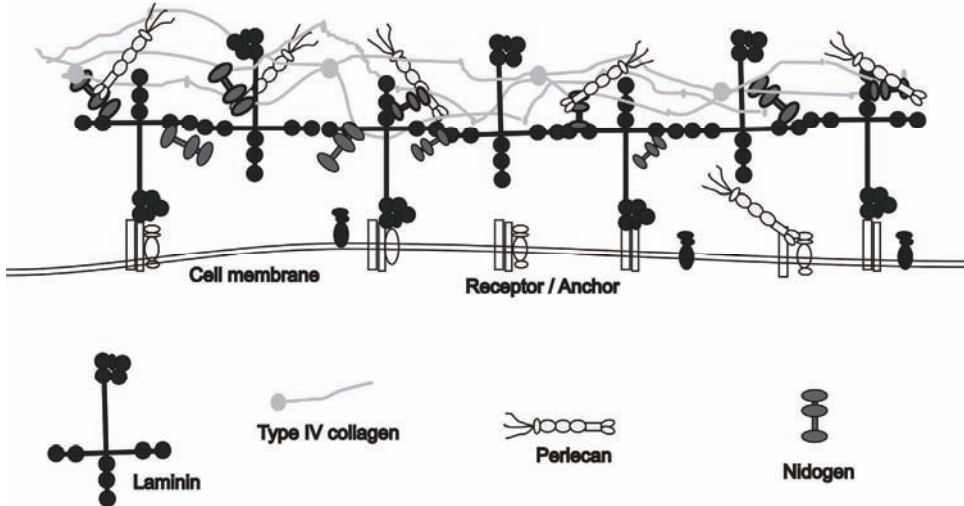


Fig. 1. Schematic illustration of a basement membrane structure. Modified from Kalluri 2003 and Yurchenco *et al.* 2004.

2.2 Collagens

Collagens are the most abundant proteins in the vertebrate body. They are located mainly outside the cells and are a prime component of the extracellular matrix and basement membranes. Most collagens form polymeric assemblies, e.g. fibrils, networks and filaments. Collagens are characterised by the presence of collagenous right-handed triple-helical domains, which are formed from three left-handed polypeptide chains (α chains) wound around each other. The collagenous domains consist of repeating -Gly-X-Y- triplets. The small size of glycine at every third position enables formation of the triple helix. Positions X and Y can be occupied by any amino acid, but proline is typically found in the X position and 4-hydroxyproline in the Y position. The prolines, with their rigid ring structures, stabilise the helix and limit rotation. The 4-hydroxyprolines are crucial for the thermal stability of collagen molecules and nonhydroxylated procollagen chains cannot form triple-helical molecules at vertebrate body temperatures. The midpoint of thermal transition from helix to coil (T_m) for hydroxylated collagen I is 39°C, but only 24°C in the absence of 4-hydroxyprolines (Berg & Prockop 1973, Jimenez *et al.* 1973). The mechanism by which 4-hydroxyprolines stabilise the triple helix is now believed to be caused by a stereoelectronic effect of the OH group. This effect leads to stabilisation of up-puckering of the pyrrolidine ring in the Y position. The up-puckering, in turn, pre-organises the range of all three of the main-chain torsion angles, enabling

optimal packing in the triple helix. Collagens also contain non-collagenous domains that have distinct functions. For example, the matricryptins derived proteolytically from collagen types IV, VIII, XV and XVIII inhibit angiogenesis and tumour growth. (Myllyharju & Kivirikko 2001, 2004, Kielty & Grant 2002, Engel & Bächinger 2005, Ricard-Blum *et al.* 2005.)

2.2.1 Collagen types

The vertebrate collagen superfamily currently consists of 28 different collagen types and 43 distinct α chains. This complex protein family can be divided into eight subfamilies according to the structural organisation and supramolecular assembly of the various collagens: (i) fibrillar collagens (I, II, III, V, XI, XXIV, XXVII); (ii) FACIT (fibril-associated collagens with interrupted triple helices) and related collagens (IX, XII, XIV, XVI, XIX, XX, XXI, XXII and XXVI) that are localised on the surface of collagen fibrils; (iii) collagens that form hexagonal networks (VIII and X); (iv) the family of type IV collagens; (v) type VI collagen, which forms beaded filaments; (vi) type VII collagen, which forms anchoring fibrils; (vii) collagens with transmembrane domains (XIII, XVII, XXIII, and XXV) that are type II transmembrane collagens with the N terminus inside the cell – such transmembrane collagens function as cell surface receptors and their ectodomain can be cleaved to produce a soluble ectodomain; and (viii) the multiplexin (multiple triple helix domains and interruptions) subfamily (XV and XVIII). Alternative splicing provides even more diversity in many collagen types. (Myllyharju & Kivirikko 2001, 2004, Kielty & Grant 2002, Veit *et al.* 2006.) The structure of collagen XXVIII seems to be unique and cannot clearly be assigned to any branch of the collagen family although sequence analysis shows a relationship to collagen VI (Veit *et al.* 2006). Type I collagen of the fibrillar collagens and type IV collagen are discussed below in more detail.

Fibrillar collagens were the first collagens to be discovered and are sometimes referred to as classical collagens. They all contain a large collagenous domain of about 300 nm flanked by non-collagenous N and C propeptides that are proteolytically cleaved to form a mature protein. Fibril-forming collagens are rigid molecules that self-assemble into highly organised cross-striated fibrils that are often heterotypic, i.e. they contain different types of fibrillar collagens. The exact composition of collagen fibrils is tissue-specific and they can thus provide the characteristic and diverse features required by specific tissues. (Myllyharju & Kivirikko 2001, 2004, Kielty & Grant 2002.)

Type I collagen generally occurs as a heterotrimer composed of two $\alpha 1(I)$ chains and one $\alpha 2(I)$ chain. However, low levels of a type I collagen homotrimer, $[\alpha 1(I)]_3$, also exist in connective tissues. Type I collagen is the most abundant collagen and is found throughout the body except in cartilaginous tissues. It comprises more than 95% of the total collagen in bone, corneas, dentin, fibrocartilage and tendons, 85% in the dermis, gingiva and heart valves, and 50-60% in intestinal, large vessel and uterine walls. The collagen fibres of connective tissues with low extensibility, such as bone, consist almost exclusively of type I collagen. The high amount of type I collagen in skin and tendons contributes to their great tensile strength, and in bone type I collagen defines the

biomechanical properties concerning load bearing, tensile strength and torsional stiffness. Most of collagen fibres contain more than one collagen type. In skin, collagen fibres contain, in addition to type I collagen, type III collagen and minor amounts of other collagen types, and in bone, tendons and corneas type V collagen. (Miller 1984, Kielty & Grant 2002, Gelse *et al.* 2003.)

About 250 mutations have been reported in the genes coding for the pro α 1 and pro α 2 chains of type I collagen. The majority of the reported mutations cause osteogenesis imperfecta I-IV, which is characterised by brittle bones and other abnormalities in other tissues rich in type I collagen. The phenotypes of osteogenesis imperfecta are heterogeneous, varying from embryonically lethal to mild forms with only few fractures. Mutations in the pro α 1 or pro α 2 chains that prevent cleavage of the N propeptide cause an arthrochalactic type of Ehlers-Danlos syndrome (EDS VIIA and VIIB, respectively). Mutations in type I collagen have also been associated with osteoporosis, and in some rare cases with the classical type of EDS. (Myllyharju & Kivirikko 2001, 2004, Byers & Cole 2002, Steinmann *et al.* 2002, Marini *et al.* 2006, OMIM 120150, 2006, OMIM 120160, 2006.)

Type IV collagen is found only in basement membranes, where it forms a filamentous network that serves as the major structural scaffold of basement membranes. In vertebrates, six genes code for the α 1(IV)- α 6(IV) chains of type IV collagen. The chains combine to form heterotypic collagen IV molecules with three compositions: α 1(IV) $_2$ α 2(IV), α 3(IV) α 4(IV) α 5(IV) and α 5(IV) $_2$ α 6(IV), the first one being the most common form. Collagen IV molecules contain a large globular domain at the C terminus (NC1), a short triple-helical 7S domain at the N terminus and a long (350 nm) collagenous domain that has several imperfections in the repeating -Gly-X-Y- sequence, which results in kinks or disruptions that provide flexibility to the network. The NC1 and 7S domains have important functions in the chain selection of the triple-helical molecules and their assembly into networks. (Kuhn 1995, Kielty & Grant 2002, Hudson *et al.* 2003, Ricard-Blum *et al.* 2005.)

The importance of the type IV collagen network in basement membranes is clearly demonstrated by naturally occurring or generated mutations in various organisms. Mutations in the human genes coding for the α 3(IV), α 4(IV) and α 5(IV) chains lead to a familial nephropathy, called Alport's syndrome (Tryggvason 1996, Hudson *et al.* 2003), and mutations in the human gene for α 1(IV) have been recently shown to cause perinatal cerebral haemorrhage and porencephaly (Gould *et al.* 2005, Breedveld *et al.* 2006, Gould *et al.* 2006). Mutations in various α (IV) chains or their complete absence lead to a wide spectrum of defects in mouse models, including renal failure, eye phenotypes, perinatal cerebral haemorrhage and porencephaly, haemorrhages or embryonal lethality, all caused by basement membrane defects (Cosgrove *et al.* 1996, Miner & Sanes 1996, Pöschl *et al.* 2004, Gould *et al.* 2005, Van Agtmael *et al.* 2005, Breedveld *et al.* 2006, Gould *et al.* 2006). The nematode *Caenorhabditis elegans* has two type IV collagen chains and mutations in either of them are embryonically lethal (Gupta *et al.* 1997). *Drosophila melanogaster* has two type IV α chains and mutation in the *Dcgl* gene encoding one of them is embryonically lethal as a result of defective muscular attachments (Borchiellini *et al.* 1996). *D. melanogaster* has also a type IV collagen-like protein pericardin that has been implicated in the morphogenesis and maintenance of the heart epithelium (Chartier *et al.* 2002).

In addition, there are more than 20 proteins that contain collagenous triple-helical domains but are not classified as collagens as they have no known structural role in the extracellular matrix. These are, for example, the surfactant proteins A and D and collectin 43 that are involved in innate immune defence, and the complement component C1q. Some of these proteins, like emilins (elastic fibre-associated glycoproteins), are associated with structural extracellular matrix proteins, and most of them also contain 4-hydroxyproline and hydroxylysine residues. (Myllyharju & Kivirikko 2001, 2004, Kielty & Grant 2002.)

2.2.2 Collagen biosynthesis

Collagen I is the most studied collagen type and is therefore used here as a model for collagen biosynthesis in general. Collagen polypeptides are translated on ribosomes and they translocate into the lumen of the endoplasmic reticulum (ER), where the first co-translational events occur while translation is still ongoing. The co- and posttranslational processing of collagens is a complicated process and involves several different enzyme activities (Figure 2). The signal peptides that efficiently target the nascent collagen polypeptides, the pre-pro α chains, into the lumen of the ER, are cleaved by a signal peptidase leading to pro α chains. Members of the 2-oxoglutarate dioxygenase family, collagen prolyl 4-hydroxylases, prolyl 3-hydroxylases and lysyl hydroxylases (LHs) are responsible for the hydroxylation of certain proline and lysine residues in the non-helical pro α chains. All three hydroxylases require Fe²⁺, 2-oxoglutarate, O₂, and ascorbate. (Kivirikko & Pihlajaniemi 1998, Myllyharju 2003, 2005.)

Collagen prolyl 4-hydroxylases (EC 1.14.11.2; C-P4H) catalyse the formation of 4-hydroxyprolines in the X-Pro-Gly sequences of collagens (Kivirikko & Pihlajaniemi 1998, Myllyharju 2003, 2005). 4-Hydroxyprolines are, as discussed above, crucial for the folding and the thermal stability of collagen triple helices. In vertebrates, the active C-P4H is a tetramer formed from two catalytic α subunits and two β subunits. At present, there are three known C-P4H α subunits; α (I), α (II) and α (III) in humans (Helaakoski *et al.* 1989, Annunen *et al.* 1997, Kukkola *et al.* 2003, Van Den Diepstraten *et al.* 2003) and rodents (Helaakoski *et al.* 1995, Kukkola *et al.* 2003), and the corresponding tetramers are referred to as type I, II, and III collagen prolyl 4-hydroxylases. The β subunit is identical to multifunctional enzyme and chaperone protein disulfide isomerase (EC 5.3.4.1, PDI) (Koivu *et al.* 1987, Pihlajaniemi *et al.* 1987), which keeps the α subunit in a soluble and active form, and in addition, its ER retention signal keeps the C-P4H tetramer inside the ER. PDI also catalyses intrachain and interchain disulfide bond formation in procollagens and binds to the C propeptide of unassembled pro α chains and thus prevents their premature and nonspecific association, and the secretion of unassembled pro α chains (Wilson *et al.* 1998, Bottomley *et al.* 2001). The C-P4H tetramer also participates in quality control by preventing secretion of procollagen molecules until the proper triple-helical conformation is achieved (Kivirikko & Pihlajaniemi 1998, Walmsley *et al.* 1999, Koide & Nagata 2005, Myllyharju 2005). Heat shock inducible chaperone Hsp47 is also essential to collagen processing and it binds to correctly folded triple helical collagen thus stabilising the folding intermediates and preventing premature procollagen

aggregation in the ER (Tasab *et al.* 2000, Koide & Nagata 2005, Myllyharju 2005). A lack of C-P4HI in mice is lethal and null mice die around embryonic day (E) 10.5 showing developmental delay. The basement membranes of the C-P4HI null embryos are fragmented and almost completely lack collagen IV, while fibrillar collagens seem to be quite normal with only slightly increased diameter (Holster *et al.* 2007).

Prolyl 3-hydroxylases (EC 1.14.11.7) catalyse the hydroxylation of proline residues in the X position of X-Y-Gly triplets, where Y must be a 4-hydroxyproline. It has been suggested that 3-hydroxyprolines modulate the stability of the collagen triple helix (Jenkins *et al.* 2003, Mizuno *et al.* 2004). Although prolyl 3-hydroxylase was partially characterised many years ago, it was cloned only recently, with three isoforms known in humans, mice, and chickens (Tryggvason *et al.* 1979, Kivirikko *et al.* 1992, Myllyharju & Kivirikko 2001, Vranka *et al.* 2004, Myllyharju 2005). Decreased prolyl 3-hydroxylation due to lack of CRTAP, a cartilage associated protein that interacts with prolyl 3-hydroxylases, causes recessive osteogenesis imperfecta (Morello *et al.* 2006).

The enzymes responsible for the hydroxylation and glycosylation of certain lysine and hydroxylysine residues, respectively, namely LHs and collagen glucosyl- and galactosyltransferases, are also located in the lumen of the ER and they are discussed in more detail below.

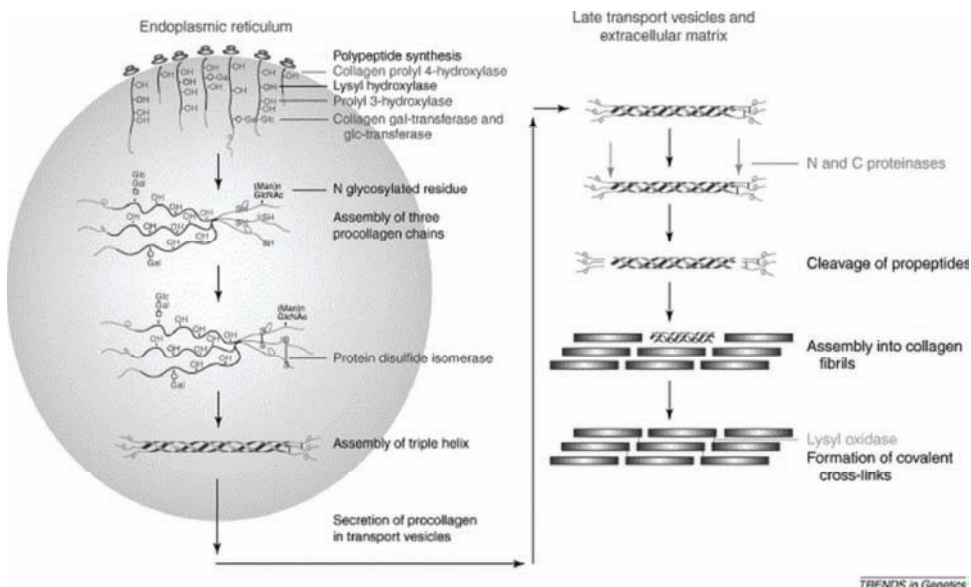


Fig. 2. Collagen biosynthesis. Reprinted from Trends in Genetics, 20, Myllyharju J., Kivirikko KI, Collagens, modifying enzymes and their mutations, p.36, Copyright (2004), with permission from Elsevier.

Association of the C-terminal propeptides in three pro α chains to form trimeric molecules initiates the folding of the triple helix of type I collagen. The C propeptides contain both intrachain and interchain disulfide bonds, the latter stabilising the association of the three C propeptides. The association of C propeptides is collagen type

specific and a so-called chain recognition region, a coiled-coil domain, has been identified in them that determines the correct chain assembly and stoichiometry. The C-terminal part of the collagenous domain then forms a nucleation point (correct alignment of the three chains), from where the triple helix folds from the C terminus to the N terminus like a zipper. To facilitate the folding process random *cis* peptide bonds preceding the proline residues are converted to the *trans* conformation by peptidyl-prolyl *cis-trans* isomerase (PPI). In addition to C-P4Hs and PDI, several chaperones, e.g. Hsp47, calreticulin and calnexin, are involved in the complex process of procollagen assembly, folding and quality control. (Lees *et al.* 1997, McLaughlin & Bulleid 1998, Lamande & Bateman 1999, Hendershot & Bulleid 2000, Hulmes 2002, Kielty & Grant 2002, McAlinden *et al.* 2003, Nagata 2003, Canty & Kadler 2005, Koide & Nagata 2005.) In non-fibrillar collagen polypeptide chains that lack the C-terminal propeptide coiled-coil recognition motifs are also employed for the determination of chain selection (Hulmes 2002, McAlinden *et al.* 2003, Khoshnoodi *et al.* 2006).

The folded triple-helical procollagen molecules are transported from the ER through the Golgi apparatus to the plasma membrane. It is generally thought that secreted proteins are transported through the secretory pathway in transport vesicles. In the case of procollagens, their dimensions exceed that of the vesicles and furthermore, they form aligned aggregates in the secretory pathway. The exact mechanism of how procollagens are secreted is still largely unknown. There is evidence that the ER to Golgi transport of procollagens takes place in saccular structures formed directly from protruding portions of specific ER exit sites (Mironov *et al.* 2003) and procollagens traverse the Golgi apparatus by cisternal maturation without ever actually leaving the lumen of the cisternae (Bonfanti *et al.* 1998). After secretion, the N- and C-terminal propeptides of procollagen molecules are cleaved by N and C proteinases, respectively, resulting in mature collagen molecules. Cleavage of the C propeptides allows spontaneous self-assembly of the collagen molecules into fibrils that aggregate and finally form a collagen fibre. In contrast, the N propeptide does not prevent assembly, but can influence the fibril shape and diameter. (Hulmes 2002, Kielty & Grant 2002, Koide & Nagata 2005.) Cleavage of the propeptides and assembly of the fibrils have been generally regarded as extracellular events. Interestingly, recent data show that in tendons membrane-bound collagen fibrils are also observed in the cytoplasm and in plasma membrane protrusions called fibripositors. Processing of procollagen has also been shown to occur inside tendon fibroblasts. (Canty *et al.* 2004, Canty & Kadler 2005.)

Non-fibrillar collagens are processed differently depending on their function. For example, in the case of collagen IV, which forms network-like structures, the newly assembled triple-helical molecules do not undergo proteolytic processing. Two type IV collagen NC1 trimers form a hexamer, i.e. two triple-helical collagen IV molecules unite via their NC1 domains head to head to form dimers containing six α chains, and four triple-helical 7S domains unite to form tetramers, leading to the formation of a network structure. (Hudson *et al.* 2003, Kalluri 2003.)

Lysyl oxidases (lysine 6-oxidase, EC 1.4.3.13) catalyse the oxidative deamination of specific lysine and hydroxylysine residues, both in the telopeptide and helical regions of collagen molecules. This leads to allysine and hydroxyallysine cross-linking pathways that are discussed later in the context of hydroxylysines and collagen cross-links. Lysyl oxidase is also crucial in cross-link formation in elastin. (Kielty & Grant 2002, Lucero &

Kagan 2006.) Lysyl oxidases form a family of copper-dependent amine oxidases that require molecular oxygen and lysyltyrosine quinone for activity. Mouse models have shown that the main lysyl oxidase is essential for the development of cardiovascular and respiratory systems and that it affects the integrity of collagen and elastic fibres (Mäki *et al.* 2002, Hornstra *et al.* 2003, Mäki *et al.* 2005). While mice lacking the lysyl oxidase-like 1 protein are viable, they show disturbed regeneration of elastic fibres in the postpartum intrauterine tract and develop pelvic prolapse, enlarged air spaces in the lung, laxity of the skin and vascular abnormalities, all caused by disruption of elastic fibre homeostasis (Liu *et al.* 2004).

2.3 Hydroxylysines

Hydroxylysines are mainly present in the X-Hyl-Gly triplets of the helical regions of collagens and collagen-like sequences of non-collagenous proteins, for example, subcomponent C1q of complement and acetyl cholinesterase. In addition, hydroxylysine residues in non-consensus sequences are found in the telopeptides of some fibrillar collagens, where hydroxylysines are found in the sequence X-Hyl-Ala or X-Hyl-Ser. The degree of hydroxylation of the lysines in the helical region varies markedly between different collagen types (Table 1). For example, in type IV collagen almost 90% of the lysines in the helical region are hydroxylated, while the corresponding number in collagen III is only about 17%. Furthermore, the hydroxylation level also varies within the same collagen type in different tissues and different developmental stages, and certain clinical conditions also affect the hydroxylation status within the same tissue. In general, the overall hydroxylation status of lysine residues is higher at young age. In contrast, the 4-hydroxyproline content of collagens in various tissues is quite constant. (Kivirikko *et al.* 1992, Myllyharju 2005.)

Table 1. Hydroxylation level of lysine residues in the helical region of various collagen chains

Collagen type	Collagen chain	Residues/1000 amino acids		% Hyl
		Hyl	Hyl + Lys	
I	$\alpha 1$	10	37	27
	$\alpha 2$	12	32	38
II	$\alpha 1$	18	38	47
III	$\alpha 1$	6	35	17
IV	$\alpha 1$	49	56	88
	$\alpha 2$	39	44	89
V	$\alpha 1$	35	55	64
	$\alpha 2$	24	42	57
	$\alpha 3$	43	58	74
VI	$\alpha 1$	48	59	81
	$\alpha 2$	67	83	81
	$\alpha 3$	48	69	70
VII	$\alpha 1$	41	59	69
VIII	$\alpha 1/\alpha 2$	22	45	49
IX	$\alpha 1$	33	53	62
	$\alpha 2$	46	58	79
X	$\alpha 1$	35	55	64
XI	$\alpha 1$	37	56	66
	$\alpha 2$	40	57	70
	$\alpha 3$	21	36	58

Modified from Kivirikko *et al.* 1992.

2.3.1 Functions of hydroxylysines

The hydroxylysine residues discussed have at least two important functions, i) they are important for the formation of stable cross-links between collagen molecules and ii) they serve as attachment sites for carbohydrate moieties (Kivirikko & Pihlajaniemi 1998). The hydroxylation level of lysine residues and subsequent glycosylation of hydroxylysine residues may affect fibril diameter; more extensive hydroxylation and glycosylation lead to a decrease in fibril diameter (Notbohm *et al.* 1999, Veit *et al.* 2006).

2.3.1.1 Collagen cross-links

Collagen fibrils provide strength and mechanical support to tissues. The tensile strength of collagen fibrils depends on the formation of covalent intramolecular and intermolecular cross-links. The importance of collagen cross-linking for the mechanical properties of tissues is clearly demonstrated by diseases caused by over- or under-hydroxylation of lysine residues (Kirsch *et al.* 1981, Bank *et al.* 1999, Steinmann *et al.*

2002, van der Slot *et al.* 2003) or by lack of lysyl oxidase activity (Kagan 1986, Kaler 1998, Smith-Mungo & Kagan 1998, Horn & Tümer 2002). The cross-linking sites have been identified in many collagen types. For example, collagens I, II and III each contain four cross-linking sites; one in each telopeptide and two in the triple-helical region at or near residues 87 and 930. The number and the types of cross-links appears to be more dependent on tissue type than collagen type. Sequences around the lysine or hydroxylysine residues at the cross-linking sites are quite similar and comparable motifs can be found in all fibrillar collagens. The flanking sequences may also affect cross-linking, as can glycosylation of the hydroxylysine residues. Two distinct cross-linking pathways exist, one employing hydroxylysine and the other lysine (see Figure 3), and the basic regulation of cross-linking is primarily determined by the hydroxylation pattern of telopeptide and helical lysines. Lysine derived cross-links are more common in tissues with a low level of telopeptide lysine hydroxylation, like skin and other soft tissues, while the hydroxyallysine pathway dominates in stiff connective tissues, such as bone, cartilage and tendons. (Knott & Bailey 1998, Eyre & Wu 2005.)

Cross-linking starts with extracellular oxidative deamination of the ϵ -amino group of hydroxylysine or lysine residues by lysyl oxidase, which leads to the formation of the corresponding aldehyde; hydroxyallysine or allysine, respectively. The later steps in the cross-linking pathways, which are spontaneous and do not require any enzymatic activity are illustrated in Figure 3. The telopeptide aldehyde will condense with a triple-helical lysine or hydroxylysine leading to immature divalent cross-links: i) lysinoketonorleucine, ii) hydroxylysinoketonorleucine, iii) hydroxylysino Δ norleucine or iv) intramolecular dimers, aldol condensation products. These can further mature into trivalent cross-links, namely lysyl pyrrole, lysyl pyridinoline (LP, also called deoxypyridinoline, Dpyr), hydroxylysyl pyridinoline (HP or Pyr) and hydroxylysyl pyrrole in the hydroxyallysine pathway. In the allysine pathway trivalent histidinyl hydroxy-lysinonorleucine is a condensation product of telopeptide allysine, glycosylated helical hydroxylysine and the helical histidine residue. (Eyre & Wu 2005.)

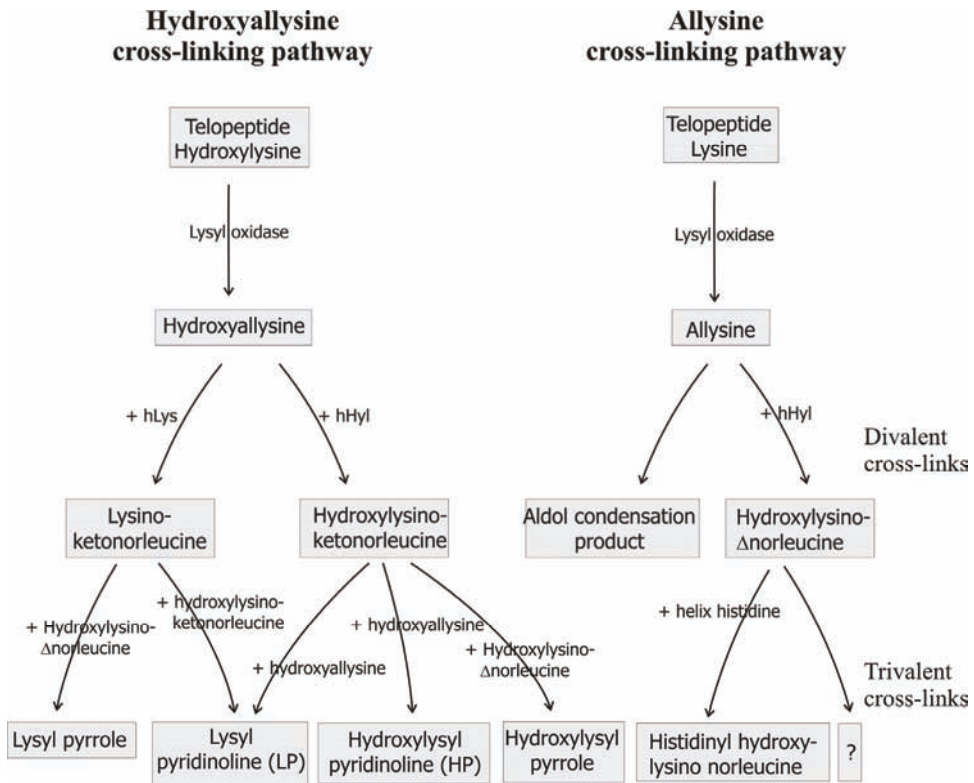


Fig. 3. Collagen crosslinking by the hydroxyallysine and allysine pathways. Modified from Eyre & Wu 2005.

Pyrroles are important cross-links in bone and their amount may correlate with the mechanical properties of the collagen fibre. Pyridinoline cross-links are equally distributed between the N and C telopeptide to helix cross-linking sites, while pyrroles are concentrated at the N telopeptide end. (Hanson & Eyre 1996, Knott & Bailey 1998, Bailey & Knott 1999.) Maturation of cross-links takes time and, in general, with age the amount of mature cross-links increases, while the amount of immature cross-links declines leading to increased insolubility of the fibres (Hanson & Eyre 1996, Knott & Bailey 1998, Eyre & Wu 2005). The level of pyridinolines that are not metabolised can be measured from urine and blood, and they are useful as markers of diseases or various physiological states, like EDS VI or bone remodelling, respectively. (Pasquali *et al.* 1994, Steinmann *et al.* 1995, Eyre & Wu 2005.)

Covalent cross-links are not found only in homopolymeric fibrils, but also in multi-component fibrils, where cross-links occur heterotypically. For example, in cartilage type II, V and IX collagens form multi-component fibrils, where these three collagen types are cross-linked both internally and to each other. (Eyre & Wu 2005.)

While cross-linking initiated by lysyl oxidase is common in fibrillar collagens, this mechanism may not be used in other collagen subgroups. For example, the network-like

structures formed by type IV collagen in basement membranes do not, in the view of recent studies, contain any lysyl oxidase-mediated cross-links, but rather stable disulfide bonds and strong hydrophobic interactions (Hudson *et al.* 2003, Eyre & Wu 2005). A novel type of cross-link has been recently detected in collagen IV, where methionine is covalently cross-linked to lysine (Than *et al.* 2002, 2005) or hydroxylysine (Vanacore *et al.* 2005) in the NC1 domain of collagen IV. In addition to disulfide bonds, other possible means for cross-linking of collagens are γ -glutamyl lysine cross-links and tyrosine derived cross-links. Over time collagens are also cross-linked through reactions of reducing sugars and lipid oxidation products, which leads to the increase in the stiffness of nonmineralising tissues observed with age. Pentoside, a difunctional cross-link formed from arginine, lysine and pentose, is one of the glycation, non-enzymatic glycosylation, end products present in bone and some other tissues. (Knott & Bailey 1998, Eyre & Wu 2005.)

2.3.1.2 Glycosylation of hydroxylysine

Collagen molecules contain carbohydrate units that are either monosaccharide galactose or disaccharide glucosyl-galactose moieties. The glucose and galactose in the disaccharide are linked by an unusual $\alpha 1 \rightarrow 2$ -O-glycosidic bond. The extent of glycosylation varies between collagen types and it is also tissue and age dependent. The ratio of monosaccharide to disaccharide depends on the collagen type. (Spiro & Spiro 1971, Spiro *et al.* 1971, Kivirikko & Myllylä 1979, Kivirikko *et al.* 1992.) The functions of these sugar moieties are not fully understood but they may have a role in the regulation of the lateral packing of collagen fibrils and supramolecular assembly (Kivirikko & Myllylä 1979, Kivirikko 1995, Notbohm *et al.* 1999), although contrary results have also been presented (Batge *et al.* 1997).

Glycosylation of hydroxylysine residues occurs in two steps and requires two different enzyme activities (Figure 4). In the first step hydroxylysine is galactosylated by hydroxylysyl galactosyltransferase (UDP galactose, 5-hydroxylysine-collagen galactosyltransferase, EC 2.4.1.50). Some of the galactosylhydroxylysines are further glucosylated by galactosylhydroxylysyl glucosyltransferase (UDP glucose, 5-hydroxylysine-collagen glucosyltransferase, EC 2.4.1.66). Both of these enzymes require a bivalent cation, preferably Mn^{2+} , and a free ϵ -amino acid in the hydroxylysine substrate. The carbohydrate is donated by an appropriate UDP-glycoside. The collagen polypeptides must be in a non-helical conformation for glycosylation to occur and the chain length also influences glycosylation efficiency (Risteli *et al.* 1976, Anttinen *et al.* 1977, Myllylä *et al.* 1977, Kivirikko & Myllylä 1979).

Active hydroxylysyl galactosyltransferase has been partially purified from chick embryo extracts with molecular weights of about 450, 200 and 50 kDa. The partially purified enzyme preferred denatured citrate soluble collagen as a substrate over gelatinised insoluble collagen, while native triple-helical collagen was not used as a substrate at all (Risteli *et al.* 1976). Galactosylhydroxylysyl glucosyltransferase has also been purified to homogeneity from chick embryos with a molecular weight of about 72-78 kDa (Myllylä *et al.* 1977). No genes have been cloned for these enzymes, but LH3

was recently reported to possess galactosylhydroxylysyl glucosyltransferase activity and a low level of hydroxylysyl galactosyltransferase activity (Heikkinen *et al.* 2000, Wang *et al.* 2002a, article I of this thesis).

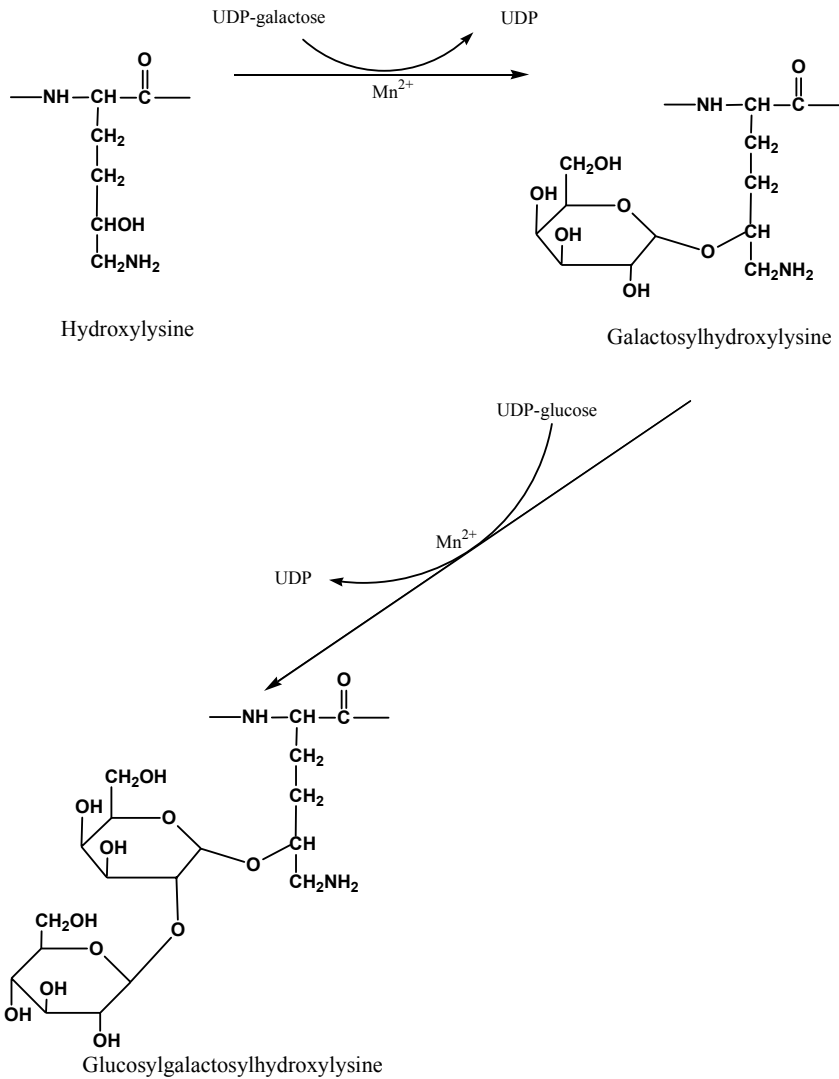


Fig. 4. Formation of galactosylhydroxylysine and glucosylgalactosyl hydroxylysine. Modified from Kivirikko & Myllylä 1982.

2.4 Lysyl Hydroxylases

2.4.1 Lysyl hydroxylase isoenzymes

LH was first cloned from chickens (Myllylä *et al.* 1991). Subsequently, the LH family was found to consist of three members in vertebrates, LH1, LH2 and LH3, which have been cloned from humans (Hautala *et al.* 1992, Valtavaara *et al.* 1997, 1998, Passoja *et al.* 1998b), mice (Ruotsalainen *et al.* 1999) and rats (Armstrong & Last 1995, Mercer *et al.* 2003). Putative orthologs are also found in several other species, zebrafish having three LH isoenzymes (Schneider & Granato 2006, 2007) and *Drosophila melanogaster* (Wang *et al.* 2002b) and *Caenorhabditis elegans* (Norman & Moerman 2000) at least one functional LH. The genes encoding human LHs, *PLOD1*, *PLOD2*, and *PLOD3* (procollagen-lysine 2-oxoglutarate 5-dioxygenase) have been mapped into the chromosome regions 1p36.2-p36.3 (Hautala *et al.* 1992), 3q23-q24 (Szpirer *et al.* 1997) and 7q22 (Valtavaara *et al.* 1998), respectively, and the corresponding mouse genes *Plod1*, *Plod2* and *Plod3* into chromosomes 4, 9 and 5 (Sipilä *et al.* 2000).

The second isoenzyme, LH2, was first cloned from human fetal kidney and pancreas cDNA libraries (Valtavaara *et al.* 1997). It was later reported to be expressed as two alternatively spliced isoforms, LH2 short and LH2 long, also called LH2a and LH2b, respectively (Valtavaara 1999, Yeowell & Walker 1999, Walker 2005a). The longer form has an additional exon 13A coding for a unique stretch of 21 amino acids (Yeowell & Walker 1999). The significance of the two LH2 isoforms is not completely understood yet. Besides differences in the expression pattern they have also been reported to have different K_m values for ascorbate and a synthetic (Ile-lys-Gly)₃ peptide substrate, and to have different peptide binding properties (Valtavaara 1999, Risteli *et al.* 2004). It has been suggested that LH2 is the telopeptide LH, and enhanced expression of LH2 long has been shown to be associated with altered collagen crosslinking and fibrosis (Uzawa *et al.* 1999, Mercer *et al.* 2003, Pornprasertsuk *et al.* 2004, van der Slot *et al.* 2004, 2005, van der Slot-Verhoeven *et al.* 2005). In addition, mutations in *PLOD2* are associated with Bruck syndrome, which is characterised by a lack of telopeptide lysine hydroxylation in bone collagen (van der Slot *et al.* 2003). The association of LH isoenzymes with different syndromes and disease states is discussed below in more detail.

Surprisingly, the most recently identified LH isoenzyme, LH3, also possesses collagen glucosyltransferase activity and galactosyltransferase activity, in addition to LH activity (Heikkinen *et al.* 2000, Wang *et al.* 2002a, article I of this thesis). Phylogenetic analysis has revealed that LH3 is the most conserved isoenzyme (Ruotsalainen *et al.* 1999) and the only LH of *C. elegans*, *let-268* (Norman & Moerman 2000), has also been shown to have collagen glucosyltransferase activity (Wang *et al.* 2002b). Furthermore, a zebrafish LH3 homologue, Diwanka, has a glycosyltransferase domain (Schneider & Granato 2007) and we have shown in collaboration with Drs. Schneider and Granato that Diwanka has collagen glucosyltransferase activity (unpublished).

2.4.2 *Lysyl hydroxylase expression*

Expression of LH isoenzymes in human and mouse tissues has been studied at mRNA level by Northern blot analysis of commercial RNA filters, conventional PCR and RT-PCR methodologies, and *in situ* hybridisation. In a few cases expression has been also analysed at the protein level with isoenzyme-specific antibodies. The expression patterns of LHs in human and mouse tissues are summarised in Tables 2 and 3, respectively. Differences in the reported expression patterns are most likely due to differences in the sensitivity of the assays used in the various studies. Analysis of LH mRNA expression in rat tissues has shown that all three LH isoenzymes, including both LH2 splicing variants, are constitutively expressed in all tissues studied (liver, heart, lung, spleen, brain, kidney, hind limb bone, bone marrow, skin, muscle, colon and cartilage), with the exception that LH2 short and LH3 are absent in stem cells originating from the bone marrow and LH2 short in differentiated hepatic stellate cells (Mercer *et al.* 2003).

In the human tissues analysed both LH2 splice variants are expressed in the frontal lobe, spleen, kidneys, liver, cartilage and leucocytes (Yeowell & Walker 1999, Ha-Vinh *et al.* 2004), while LH2 long is the only or at least the major isoform in the skin, dura, lungs and aorta (Yeowell & Walker 1999, Ha-Vinh *et al.* 2004, Salo *et al.* 2006b). LH2 short is the major form in human kidneys and spleen (Yeowell & Walker 1999) and in mouse kidneys and testis (Salo *et al.* 2006b).

In cultured human skin fibroblasts splicing of LH2 is regulated by cell density and cycloheximide, an inhibitor of protein synthesis, in such a way that low cell density or cycloheximide treatment leads to an increased relative expression of the LH2 short isoform (Walker *et al.* 2005a). The effect of cycloheximide on the relative expression levels of the LH2 splice variants indicates that the exclusion of exon 13A, which leads to LH2 short, requires a newly synthesised factor that is absent in high cell densities. The same phenomenon is seen in human kidney cells that express both splice variants. In kidney cells cycloheximide treatment decreases the LH2 long levels suggesting that cycloheximide suppresses a factor that inhibits exclusion of 13A (Walker *et al.* 2005a).

Table 2. Expression of LH isoenzymes in human tissues

Tissue	LH1	LH2	LH3
Adrenal gland	n.d.	n.d.	yes ⁵
Aorta	yes ¹	yes ⁷	n.d.
Artery	yes ¹	n.d.	n.d.
Bone marrow	n.d.	n.d.	yes ⁵
Brain	yes ²	no ⁶ ; yes ^{3,7}	yes ^{4,5}
Cartilage	yes ¹	yes ⁷	n.d.
Chorionic villus	n.d.	yes ⁷	n.d.
Diaphragm	yes ¹	n.d.	n.d.
Dura	n.d.	yes ⁷	n.d.
Oesophagus	yes ¹	n.d.	n.d.
Frontal lobe	n.d.	yes ⁷	n.d.
Gall bladder	yes ¹	n.d.	n.d.
Heart	yes ²	yes ^{3,6}	yes ^{4,5}
Intestine	yes ¹	n.d.	n.d.
Kidney	n.d.	no ⁶ ; yes ^{3,7}	yes ^{4,5}
Liver	yes ²	yes ^{3,6,7}	yes ^{4,5}
Lung	yes ^{1,2}	no ⁶ ; yes ^{3,7}	yes ^{4,5}
Lymph node	n.d.	n.d.	yes ⁵
Pancreas	n.d.	yes ^{3,6}	yes ^{4,5}
Placenta	yes ^{1,2}	yes ^{3,6,7}	yes ^{4,5}
Skeletal muscle	yes ^{2,8}	yes ^{3,6,8}	yes ^{4,5}
Skin	n.d.	yes ⁷	n.d.
Spinal cord	n.d.	n.d.	yes ⁵
Spleen	yes ¹	yes ⁷	n.d.
Stomach	n.d.	n.d.	yes ⁵
Thyroid	n.d.	n.d.	yes ⁵
Trachea	n.d.	n.d.	yes ⁵
Vein	yes ¹	n.d.	n.d.

1 (Yeowell *et al.* 1994); 2 (Heikkinen *et al.* 1994); 3 (Valtavaara *et al.* 1997); 4 (Valtavaara *et al.* 1998); 5 (Passoja *et al.* 1998b); 6 (Valtavaara 1999); 7 (Yeowell & Walker 1999); 8 (Hjalt *et al.* 2001); n.d.= not determined.

Table 3. Expression of LH isoenzymes in mouse tissues

Tissue	LH1	LH2	LH3
Brain	yes ^{2,3}	no ¹ ; yes ^{2,3,4,5}	yes ^{2,4}
Eye	n.d.	yes ²	n.d.
Heart	yes ^{2,3}	yes ^{1,2,3,4}	yes ²
Kidney	no ³ ; yes ²	yes ^{1,2,3,4}	yes ^{2,4}
Liver	yes ^{2,3}	no ^{2,3,4} ; yes ¹	yes ^{2,4}
Lung	yes ^{2,3}	no ^{1,3} ; yes ^{2,4}	yes ^{2,4}
Ovary	n.d.	yes ²	n.d.
Pancreas	n.d.	yes ⁴	yes ⁴
Skeletal muscle	yes ^{2,3}	no ³ ; yes ^{1,2,4}	yes ^{2,4}
Spleen	no ² ; yes ³	no ^{1,2} ; yes ^{3,4}	no ² ; yes ⁴
Testis	yes ^{2,3}	no ³ ; yes ^{1,2,4}	yes ²
Uterus	n.d.	yes ²	n.d.

1 (Valtavaara 1999); 2 (Ruotsalainen *et al.* 1999); 3 (Hjalt *et al.* 2001); 4 (Salo *et al.* 2006b); n.d.=not determined.

Expression of LH1 mRNA is detected by RT-PCR in mouse embryos at E7 and E11 and that of LH2 at E7 and E17 (Hjalt *et al.* 2001). The early embryonic expression of LHs was confirmed by a very recent study using both *in situ* hybridisation and RT-PCR (Salo *et al.* 2006b). All three LHs were found to be expressed from the earliest stage studied, E7, onwards. Expression was found in the mesoderm at E7.5, and in mesoderm-derived tissues and the hind- and mid-brain at the later stages. In early mouse embryos LH2 short was more abundant, while LH2 long became the major form from E11.5 onwards. In contrast, in human foetal skin splicing of LH2 is not developmentally regulated and expression of LH2 long is seen from at least 8 weeks of gestation onwards (Yeowell & Walker 1999). LH3 mRNA expression is detected in human embryonic lung, liver and kidney tissue, but not in brain tissue (Passoja *et al.* 1998b).

Expression of LH isoenzymes has also been studied in several non-transformed and transformed cell lines. It has been shown that expression of LH3 is quite uniform in all cell lines studied, while the mRNA levels of LH2 can even show 20-30-fold changes from one cell line to another (Wang *et al.* 2000). In another study, LH1 and LH3 were found to be highly expressed in normal skin fibroblasts and bone marrow stromal cells, while LH2 expression was increased only in fully differentiated bone marrow stromal cells, coinciding with matrix mineralization and hydroxylation of collagen telopeptides (Uzawa *et al.* 1999).

2.4.3 Molecular characteristics of lysyl hydroxylase isoenzymes

The newly translated LHs contain a signal peptide (18-25 amino acids in humans) that targets them into the ER. The processed human LH1, LH2 short, LH2 long and LH3 polypeptides contain 709, 712, 733 and 714 amino acids, respectively (Hautala *et al.* 1992, Valtavaara *et al.* 1997, 1998, Passoja *et al.* 1998b). LHs are lumenally orientated membrane associated ER proteins (Kellokumpu *et al.* 1994), although they do not contain

either of the known ER retention signals, a -Lys-Asp-Glu-Leu or a double lysine motif. Instead, the localisation is mediated by a C-terminal peptide segment in the iron-binding domain of the polypeptides (Kellokumpu *et al.* 1994, Suokas *et al.* 2000, 2003). Interestingly, it has been shown that a fraction of LH3 is secreted in cells overexpressing it leading to the presence of an active LH3 in the extracellular space (Heikkinen *et al.* 2000, Salo *et al.* 2006a). *In vivo*, LH3 was found to be present in the serum (Salo *et al.* 2006a). Furthermore, LH3 was also found to be located both inside the ER and in the extracellular space in various tissues, the distribution of LH3 between the two compartments being tissue specific (Salo *et al.* 2006a). For example, in mouse tissues LH3 is found in both compartments in muscle and spleen, while in the kidneys it is mainly inside the cells and in the liver and pancreas it is found only in the extracellular space (Salo *et al.* 2006a, 2006b). Localisation between the ER and extracellular space may be developmentally regulated at least in mouse liver (Salo *et al.* 2006b).

The amino acid sequence identity between the human LH1 and LH2 polypeptides is 75%, that between LH2 and LH3 57%, and that between LH1 and LH3 59% (Valtavaara *et al.* 1997, 1998, Passoja *et al.* 1998b). In mice the corresponding identities are 59%, 58% and 58% (Ruotsalainen *et al.* 1999). The mouse LHs are about 91% identical to their human counterparts at the amino acid level (Ruotsalainen *et al.* 1999). The identity is highest within the C-terminal region that contains the catalytically critical amino acids typical of the 2-oxoglutarate family (Myllyharju & Kivirikko 1997, Kivirikko & Pihlajaniemi 1998, Myllyharju 2003); two histidines and an aspartate (His638, Asp640 and His690 in the processed human LH1) that are required for binding of Fe²⁺ to the catalytic site (Pirskanen *et al.* 1996) and an arginine (Arg700 in LH1) that binds the C-5 carboxyl group of 2-oxoglutarate (Passoja *et al.* 1998a). In addition, the human LHs contain several conserved cysteine residues that may be structurally important. For example, LH1 contains 10 cysteine residues, nine of which are conserved between the three isoenzymes. Individual mutations of the cysteines to serines revealed that four of them (Cys351Ser, Cys357Ser, Cys534Ser and Cys669Ser) eliminated LH activity, while Cys249Ser, Cys252Ser and Cys662Ser mutations reduced the activity somewhat and Cys186Ser, Cys466Ser and Cys558Ser mutations did not decrease the activity (Yeowell *et al.* 2000a).

The collagen glycosyltransferase activities of LH3 reside in the N terminus of the polypeptide (Heikkinen *et al.* 2000, Wang *et al.* 2002b, article I of this thesis). This region contains an Asp-X-Asp motif characteristic of many glycosyltransferases, and also a cysteine and a leucine residue that were found to be important for the glucosyltransferase activity of LH3 (Wang *et al.* 2002b). The same Asp-X-Asp motif and cysteine were also found to be important for galactosyltransferase activity (Wang *et al.* 2002a).

The LH isoenzymes have several potential asparagine-linked glycosylation sites, some of which have been shown to be required for optimal enzyme activity (Myllylä *et al.* 1988, Pirskanen *et al.* 1996). Site-directed mutagenesis studies have shown that two of the four potential glycosylation sites in LH1, Asn179 and Asn520, are effectively glycosylated: mutation of Asn179 to serine decreased the enzyme activity to about 25% (Pirskanen *et al.* 1996). The catalytic α subunit of C-P4H contains separate peptide-substrate-binding and catalytic domains (Myllyharju & Kivirikko 1999, Hieta *et al.* 2003,

Pekkala *et al.* 2004). So far, no peptide-substrate-binding domain or region has been identified in LHs and structural data of the full-length enzymes is still missing.

2.4.4 Reaction mechanism and cosubstrates

LHs belong to a family of 2-oxoglutarate dioxygenases as indicated by the names of their genes, *Plod* (procollagen-lysine 2-oxoglutarate 5-dioxygenase). Like other collagen hydroxylases, namely collagen prolyl 4-hydroxylases and prolyl 3-hydroxylases, LHs require Fe^{2+} , 2-oxoglutarate, molecular oxygen and ascorbate for their activity. The reaction mechanisms of the three collagen hydroxylases are quite similar and involve an ordered binding of Fe^{2+} , 2-oxoglutarate, O_2 and the peptide substrate; the order of involvement of the last two reactants was somewhat uncertain until recently. (Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998, Myllyharju 2003, 2005.) In fact, mechanistic studies based on several crystal structures of 2-oxoglutarate dioxygenases have shown that the substrate is bound before O_2 (Clifton *et al.* 2006). The reaction products are also released in an ordered manner; first the hydroxylated peptide, followed by CO_2 and succinate. Fe^{2+} is not generally released between the cycles. In addition, bovine serum albumin, catalase and dithiothreitol are required for maximal *in vitro* activity. (Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998, Myllyharju 2003, 2005.)

2-Oxoglutarate is stoichiometrically decarboxylated during the hydroxylation and one atom from O_2 is incorporated into the hydroxyl group formed in the lysine residue, while the other is transferred to the succinate (Figure 5). In contrast, ascorbate is not consumed stoichiometrically and the reaction can proceed for several cycles even in its absence. Ascorbate is required for the uncoupled decarboxylation of 2-oxoglutarate, a reaction which occurs without subsequent hydroxylation of the substrate, where it serves as an alternative oxygen acceptor. In the uncoupled reaction ascorbate is used stoichiometrically. The rate of the uncoupled reaction is about 1-10% of the complete reaction at saturating concentrations of peptide substrates, and peptides that do not serve as substrates have been shown to enhance the rate of the uncoupled reaction. (Myllylä *et al.* 1984, Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998, Myllyharju 2003, 2005.)

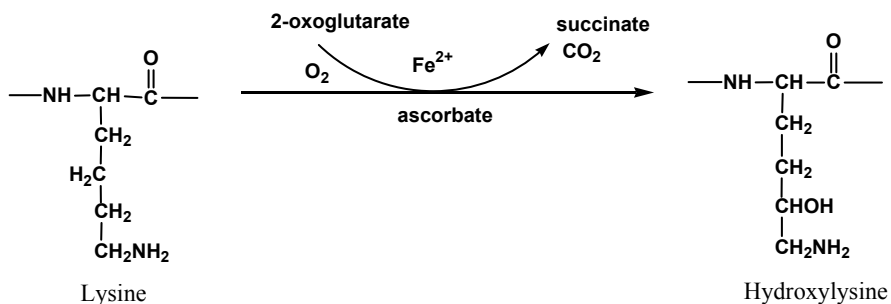


Fig. 5. Hydroxylation of a lysine residue by LH.

The K_m values of vertebrate collagen hydroxylases for the cosubstrates are fairly similar (Table 4), major differences being seen in the case of 2-oxoglutarate for which LH has about a 5 times higher K_m than C-P4H. This suggests that collagen hydroxylases may have differences in their 2-oxoglutarate binding sites.

Table 4. K_m values of vertebrate collagen hydroxylases for cosubstrates

Enzyme and substrate	K_m (μ M)			
	Fe ²⁺	2-oxoglutarate	O ₂	Ascorbate
Lysyl hydroxylase				
Peptide substrate ¹	2	100	45	300-350
Collagen prolyl 4-hydroxylase				
Peptide substrate ²	4	22	43	330
Biological substrate ³	2	5	n.d.	100
Collagen prolyl 3-hydroxylase				
Biological substrate ⁴	2	3	30	120

1 (Turpeenniemi-Hujanen *et al.* 1981, Pirskanen *et al.* 1996); 2 (Myllylä *et al.* 1977, Helaakoski *et al.* 1995); 3 (Kivirikko & Prockop 1967); 4 (Tryggvason *et al.* 1979); n.d. not determined. Modified from Kivirikko & Pihlajaniemi 1998.

2.4.5 Substrates and substrate specificity

LHs hydroxylate lysine residues in collagens *in vivo* in the lumen of the ER while the collagen polypeptide chains are still being translated and before the formation of the triple helix. The triple-helical conformation completely prevents lysine hydroxylation by LH. *In vitro* studies have established that the minimum substrate requirement for LH is an X-Lys-Gly triplet and that LH does not hydroxylate free lysine. Lysine residues in X-Lys-Ala and X-Lys-Ser sequences are also hydroxylated in the telopeptides of fibrillar collagens and in arginine-rich histones. (Kivirikko & Myllylä 1980, Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998, Myllyharju & Kivirikko 2004.) It has been shown that the amino acids adjacent to the lysine as well as the length of the peptide substrate, affect the lysine hydroxylation, while the chain length of the peptide substrate affects the K_m value, but not the rate of the reaction. (Kivirikko & Myllylä 1980, Kivirikko *et al.* 1992, Kivirikko & Pihlajaniemi 1998.)

It currently seems that the LH isoenzymes do not have strict substrate sequence specificities, but they may have certain preferences (Risteli *et al.* 2004). Analysis of the mRNA expression levels has also revealed no correlation between different LHs and collagen types indicating that LHs have no collagen type specificity (Wang *et al.* 2000). In contrast, differences in the hydroxylation level of helical lysines and the telopeptide lysines of fibrillar collagens under different physiological conditions and in different tissues suggest that LHs should have differences in either their substrate or tissue specificity (Royce & Barnes 1985, Gerriets *et al.* 1993, Eyre *et al.* 2002).

Based on observations with EDS VIA patients (see below) LH1 is thought to be responsible for hydroxylation of the lysine residues in the triple-helical region of

collagens. These patients lack LH1 activity and the hydroxylation level of the lysines in the helical region is decreased, while that in the telopeptides is unaltered (Steinmann *et al.* 2002). However, some of the helical lysines are still hydroxylated in EDS VIA patients, the hydroxylation status being tissue and collagen type dependent (Ihme *et al.* 1984). This implies that either LH2 or LH3, or both, are also able to hydroxylate these residues. Contradicting results about the hydroxylation of telopeptides by LH1 have also been reported (Royce & Barnes 1985, Mercer *et al.* 2003).

Several studies indicate that LH2 is the main LH isoenzyme responsible for hydroxylation of telopeptides. Mutations in *PLOD2* have been found in Bruck syndrome (see below), a disease characterised by pyridinoline deficiency and thus a lack of telopeptide hydroxylysines in bone collagen (van der Slot *et al.* 2003), while increased LH2 expression has been shown to be associated with an increased amount of hydroxylysine-derived crosslinks in keloids, systemic scleroderma and in fibrotic disorders in general (Uzawa *et al.* 1998, van der Slot *et al.* 2003, 2004, 2005, Brinckmann *et al.* 2005, van der Slot-Verhoeven *et al.* 2005). The same phenomenon is also seen in cell cultures overexpressing LH2 (Mercer *et al.* 2003, Pornprasertsuk *et al.* 2004). On the other hand, the presence of normal levels of pyridinoline cross-links in the cartilage and tendons of Bruck syndrome patients indicates that some other enzyme besides LH2 can also hydroxylate telopeptides (Bank *et al.* 1999).

2.4.6 Diseases associated with lysyl hydroxylase deficiency

LH activity is required for proper post-translational modifications of collagen molecules, and the hydroxylysines formed are important for the collagen cross-links and serve as attachment sites for glycosyl groups. In general, lack of LH activity leads to more fragile tissues while overexpression of LH can lead to excessively stiff tissues, as in fibrotic conditions. Mutations in the human genes coding for LH1 and LH2 have been associated with the specific syndromes discussed below, while no human diseases caused by lack of LH3 activity are currently known. For other diseases related to collagens and collagen modifying enzymes or their mouse models, see recent reviews Myllyharju & Kivirikko 2004 and Aszodi *et al.* 2006.

2.4.6.1 The kyphoscoliotic type of Ehlers-Danlos syndrome

EDS VIA, also called the kyphoscoliotic type of EDS (OMIM 225400), is an autosomal recessive disease caused by lack of LH1 activity. This disease leads to changes in collagens, altering their solubility and tensile strength as a result of changes in the hydroxylation of helical lysines and collagen cross-links. This syndrome is characterised by severe muscular hypotonia at birth, generalised joint laxity with occasional hypermobility and luxation, progressive kyphoscoliosis, soft and fragile skin with a tendency to severe scarring, scleral fragility and rupture of the ocular globe. Other symptoms observed in EDS VIA patients are tissue fragility, easy bruising, arterial

ruptures, a Marfanoid habitus, microcornea and osteopenia without tendency to fractures. (Steinmann *et al.* 2002, OMIM 225400, 2006).

At least 20 mutations in the *PLOD1* gene causing EDS VIA have been identified and they include duplications, deletions, insertions and point mutations (Yeowell & Walker 2000, Steinmann *et al.* 2002, Walker *et al.* 2005b). The two most common *PLOD1* mutations, which have been found in five or more unrelated patients, are a large duplication of seven exons caused by an Alu-Alu recombination in introns 9 and 16 (Heikkinen *et al.* 1997) and a nonsense mutation in exon 14 leading to a premature termination codon (Yeowell & Walker 1997, Walker *et al.* 1999, Pousi *et al.* 2000, Yeowell *et al.* 2000b). At least two other mutations have been found in more than one unrelated patient indicating that there might be an ancestral gene involved in certain mutations. Several alternative splicing pathways have been reported, which bypass mutations that lead to a premature termination codon in cases when the mutations are not in the catalytically active C-terminal region (Yeowell & Walker 2000, Yeowell *et al.* 2000b). Nevo syndrome was recently reported to be caused by mutations in the *PLOD1* gene and can thus be considered to be identical to EDS VIA (Giunta *et al.* 2005a). Although EDS VIA is a rare disease, many mutations causing it are now known and a molecular diagnosis strategy for identification of mutations causing EDS VI has been recently proposed (Giunta *et al.* 2005b).

The clinical diagnosis of EDS VIA based on the symptoms can be confirmed by assaying LH activity in cultured skin fibroblasts from the patients, which is usually less than 25% of the normal activity. Another typical biochemical finding in patient fibroblasts is an increased mobility of the collagen α chains in gel electrophoresis due to under-hydroxylation of the lysine residues. Another valuable, non-invasive assay for diagnosis is the determination of the total urinary LP to HP ratio. In EDS VIA patients the LP/HP-ratio is markedly increased; 5.97 ± 0.99 (range 4.3-8.1) when compared to control values of 0.20 ± 0.05 (range 0.10-0.38) (Pasquali *et al.* 1994, 1995, Yeowell & Walker 2000, Steinmann *et al.* 2002, Giunta *et al.* 2005b). Patients with partly overlapping symptoms but with normal LH activity are classified as having EDS VIB (OMIM 229200). No *PLOD1* mutations have been detected in EDS VIB but some of these patients have decreased LH2 or LH3 mRNA levels (Walker *et al.* 2004) and in some cases a decrease was seen in the hydroxylation of both helical and telopeptide lysines (Uzawa *et al.* 2003).

2.4.6.2 Bruck syndrome

Bruck syndrome (BS, OMIM 259450), an autosomal recessive disease affecting the bones, was recently associated with mutations in exon 17 of the *PLOD2* gene (van der Slot *et al.* 2003, Ha-Vinh *et al.* 2004). BS patients suffer from congenital joint contractures, pterygia, osteoporosis and fragile bones leading to multiple fractures and short stature. (McPherson & Clemens 1997, Blacksin *et al.* 1998, Leroy *et al.* 1998, OMIM 259450, 2006.)

Biochemical analyses of BS patients have shown that the telopeptide lysine residues of type I collagen, but not helical lysines, are under-hydroxylated in the bone, which leads to

a decreased amount of pyridinoline cross-links in this tissue. In contrast, the cross-linking pattern of collagens in the cartilage and ligaments is normal in these patients (Bank *et al.* 1999). BS patients have an increased collagen turnover, which is shown by an increased urinary hydroxyproline (Hyp) level. The best diagnostic marker in BS patients seems to be a HP+LP/Hyp ratio that has been below 1 in the few patients tested so far, compared to the ratio of 2.80 (range 1.83-4.31) in healthy individuals (Ha-Vinh *et al.* 2004). The known *PLOD2* mutations in BS are located in a conserved region in exon 17, but the effects of these mutations on the function of LH2 are yet unknown.

2.4.6.3 Fibrotic conditions

In fibrotic diseases an excess amount of extracellular proteins, mainly collagens, accumulate due to elevated collagen synthesis and decreased collagen degradation (Trojanowska *et al.* 1998). Several studies have indicated that cross-linking of collagens is changed in fibrotic diseases. Elevation of cross-linking derived from the hydroxyallysine route, which employs a telopeptide hydroxylysine, leads to an irreversible fibrotic process, this being often associated with overexpression of LH2 (Brinckmann *et al.* 1999, Uzawa *et al.* 2003, van der Slot *et al.* 2003, 2004, 2005, Pornprasertsuk *et al.* 2004, van der Slot-Verhoeven *et al.* 2005, Wu *et al.* 2006). For example, an increase in both pyridinolines and dehydro-histidinohydroxymerodesmosine has been observed in patients with localised scleroderma (Brinckmann *et al.* 2001), and in skin fibroblasts from systemic sclerosis patients an increase in the HP+LP/collagen triple helix ratio was observed due to an increase in the amount of HP cross-links - these are almost absent in normal skin (van der Slot *et al.* 2003, Brinckmann *et al.* 2005). In the two latter studies mRNA levels in fibroblasts were analysed and the expression of LH2 long was highly increased, while the expression levels of the other LHs and lysyl oxidase were normal. An increase in both pyridinoline formation and expression of LH2 long is also seen in other fibrotic diseases (van der Slot *et al.* 2004) and in the hydroxylation of the telopeptides in keloid fibroblasts (Uzawa *et al.* 2003). Inflammation and hypoxia caused by microvascular damage are major pathological processes during the early and acute progressive stages of systemic scleroderma (Brinckmann *et al.* 2005). In an *in vitro* model of systemic sclerosis expression of LH2 was found to be induced by the inflammatory cytokine interleukin 4 and prolonged hypoxia (Brinckmann *et al.* 2005). In normal fibroblasts, an increase in both the expression of LH2 long and formation of pyridinoline cross-links is seen as a response to stimulation with profibrotic cytokines, such as interleukin-4, transforming growth factor β , activin A and tumour necrosis factor α (van der Slot *et al.* 2005).

3 Aims of the study

LH1 was first purified to homogeneity and partially characterised more than 20 years ago (Turpeenniemi-Hujanen *et al.* 1980, 1981). Recently two novel isoenzymes, namely LH2 and LH3, were cloned from humans, mice and other species. The aim of this study was to further characterise this enzyme family, with special emphasis on their molecular and catalytic properties, substrate specificity, and *in vivo* functions. This study utilises modern methodologies from recombinant expression of LHs in insect cells to generation of mouse lines lacking LH1 or LH3.

By using recombinant LHs the following aspects were studied: (i) the domain structure of LHs and the functions of each domain; (ii) the kinetic parameters of LHs; and (iii) the substrate specificity of LH isoenzymes, which was studied using either *in vitro* hydroxylation of synthetic peptides representing collagen sequences by purified recombinant LHs, or *in vivo* hydroxylation of recombinant procollagen chains by simultaneously coexpressed LH in insect cells.

By using of mice that lack LH1 or LH3 activity the following aspects were analysed: (i) the *in vivo* function/role of the specific LH; (ii) the expression pattern of LH1 and LH3 isoenzymes; and (iii) the ability of other LH isoenzymes to compensate for the lack of one isoenzyme. An additional original aim in the case of LH1 null mice was to generate an animal model for the human disease EDS VIA.

4 Materials and Methods

The methods used in this thesis are summarised in Table 5, and the *Plod1* and *Plod3* gene structures and the targeting constructs used to create the corresponding null mouse lines are illustrated below in Figure 6. The experimental procedures are described in detail in original articles I-IV.

Table 5. Methods used in original publications I-IV

Method	Original publication
Amino acid analysis	II, IV
Cell culture	I, II, III, IV
Circular dichroism spectroscopy	I
Collagen cross-link analysis	IV
Determination of LH activity using 2-oxo[1- ¹⁴ C]glutarate	I, II
Determination of galactosyl- and glucosyltransferase activity	I, III
Determination of LH activity with [¹⁴ C]lysine-labelled procollagen	III, IV
Expression of recombinant proteins in insect cells	I, II
Generation of gene-targeted mouse lines	III, IV
Light and immunofluorescence microscopy	III, IV
N-terminal sequencing	I, II
Protease digestion	I
Purification and characterisation of recombinant proteins	I, II
SDS-PAGE, Tris-Tricine PAGE and urea gradient electrophoresis, and Western blot analysis	I, II, III
Southern analysis and RT-PCR	III, IV
Transmission and immunoelectron microscopy	III, IV

All the blastocyst injections and most of the embryonic stem (ES) cell cultures were carried out in the Transgenic core facility of Biocenter Oulu and the gene targeted mice were maintained in the Laboratory Animal Centre of the University of Oulu. Animal experiments were approved by the Animal Research Committee of the University of Oulu. The samples for electron microscopy studies were processed and analysed in the

Imaging Core facility of Biocenter Oulu. Cross-link analysis was performed during my visit to TNO, Leiden, Netherlands. N-terminal sequencing and amino acid analysis was done in the Protein sequencing and amino acid analysis core facility of Biocenter Oulu. The echocardiographic analysis was done at the Biocenter Oulu Core Facility of Physiological Analysis of Transgenic animals.

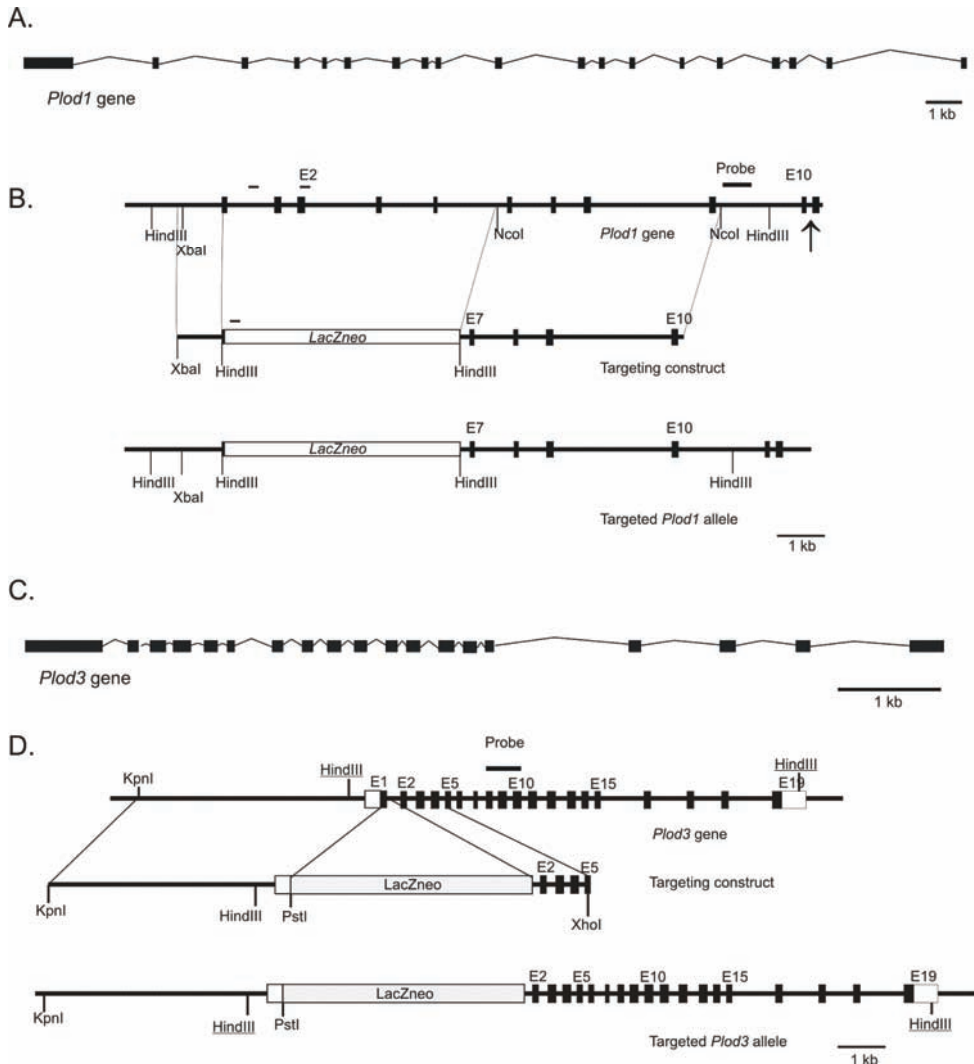


Fig. 6. Schematic representation of the *Plod1* (A) and *Plod3* (C) genes and the targeting constructs used to generate the *Plod1*^{-/-} (B) and *Plod3*^{-/-} (D) mouse lines. The insertion site of the gene trap construct in the LH1:RST531 mouse ES cell line is indicated by an arrow in B.

5 Results

5.1 Purification of recombinant human lysyl hydroxylases and characterisation of their domain structure and functions of the domains (I)

As ER resident proteins LHs have a tendency to aggregate easily in various buffers. In order to obtain reasonable yields of recombinant LHs and to make their purification easier, the cDNAs coding for human LH1, LH2 long and LH3 were cloned into a pAcGP67A expression vector. This vector contains a sequence for the baculoviral GP67 signal peptide, which leads to efficient secretion of the fused protein into the culture medium. In addition, to facilitate the purification procedure a histidine-tag was inserted in-frame between the GP67 signal sequence and the LH cDNA. High five insect cells were infected with a baculovirus coding for the desired polypeptide and used as host cells to produce recombinant LHs. The LH polypeptides were purified from the culture medium 48 h after infection using metal chelate affinity chromatography with imidazole elution. Imidazole inhibits LH activity and it was removed by using PD-10 columns. The purified LHs were analysed by SDS-PAGE under reducing conditions and by non-reducing PAGE (Figure 1 in I). The purified polypeptides had a molecular weight of 80-85 kDa in SDS-PAGE corresponding to full-length LH1, LH2 and LH3. N-terminal sequencing was used to confirm the identities of the purified polypeptides (not shown). The purified LHs were further analysed in a calibrated gel filtration column to determine the molecular composition of the native LHs. All three LHs eluted in positions corresponding to a molecular weight of about 180 kDa indicating that they were dimers. All purified LHs had LH activity (as shown for LH1 in Table I in I).

Limited proteolysis of the purified recombinant LHs with thermolysin, trypsin and proteinase K was employed to study their possible domain structure. The purified LHs were incubated with the proteases for 30-90 min at 37°C and the digested samples were analysed by 12% Tris-Tricine PAGE followed by Coomassie blue staining (as shown for LH1 digested with thermolysin in Figure 2 in I). The five most prominent peptides of about 68, 37, 33, 18 and 16 kDa were found after digestion of all three LHs with each of

the proteases, except that the 16 kDa peptide was absent in the LH3 digest. N-terminal sequencing was used to identify the peptides originating from LH1 and LH3, and based on this data two protease sensitive regions in the LHs were identified and are indicated in the sequence alignment of the LH polypeptides in Figure 3 in I. This indicates that LH polypeptides contain at least three structural domains, termed here fragments A, B and C from the N to C terminus (Figure 3B in I). The molecular masses of these fragments, when calculated based on the amino acids sequences, are 29.6, 36.2 and 15.9 kDa for nonglycosylated LH1 and 30.4, 36.2, and 15.8 kDa for the LH3 fragments. Based on N-terminal sequencing, the protease-resistant 68, 37, 33 and 16 kDa polypeptides in the SDS-PAGE correspond to the A-B fragment pair, and the single B, A and C fragments, respectively. The 18 kDa band observed in the LH1 digest most likely also corresponds to fragment C, the higher molecular weight probably being due to N-glycosylation of residue 668. The protease resistant fragments of LH2 were not sequenced because their sizes were similar to those of the corresponding fragments of LH1 and LH3, and the domain structure of LH2 was assumed to be similar.

In order to study if it was possible to express and purify recombinant LH1 and LH3 fragments A, B and C in folded forms in insect cells, the cDNAs coding for each fragment were cloned into the pAcGP67A vector with a histidine-tag. LH1 fragment A contained amino acids 1-262, fragment B amino acids 255-572 and fragment C amino acids 565-709 and the corresponding fragments of LH3 the amino acids 1-266, 261-577 and 570-714. The fragments were expressed and purified as described above. Only the A fragments of LH1 and LH3 were efficiently secreted into the culture medium and could be purified to homogeneity. In contrast, fragments B and C were found mainly inside the cell pellets and could be efficiently solubilised only in 1% SDS. Therefore, vectors coding for the LH1 and LH3 fragment pairs A-B (amino acids 1-572 and 1-577, respectively) and B-C (amino acids 255-709 and 261-714) with histidine-tags were constructed and insect cells were infected with the corresponding viruses. A significant proportion of these fragment pairs was found in the culture medium and could be easily purified to homogeneity. Figure 4A in I shows SDS-PAGE analysis of purified recombinant A fragments and fragment pairs A-B and B-C of LH1 and LH3. Folding of the purified fragments and fragments pairs was analysed with transverse urea-gradient gel electrophoresis (Figure 4B in I) and far-UV CD-spectroscopy (Figure 5 in I). Both methods indicated that these fragments and fragment pairs were folded, this was also supported by the data from activity measurements.

Purified recombinant LH1 and LH3 and their A fragments and fragment pairs A-B and B-C were analysed for LH activity. Full length recombinant LH1 and LH3 were active (Table I in I) after removal of imidazole, which rapidly inactivates the enzymes, especially LH3, if stored in its presence. Unsurprisingly, the A fragments and fragment pairs A-B had no LH activity (Table I in I), since they lack the C terminal domain that contains the residues critical for LH activity. In contrast, fragment pairs B-C of LH1 and LH3 were fully active LHs. Even the K_m values of the LH1 B-C domain pair for cosubstrates and the peptide substrate were essentially identical to those of full-length LH1 (Table II in I).

At the time of this study, it had been reported that full-length recombinant LH3, but not LH1, possessed collagen glucosyltransferase activity (Heikkinen *et al.* 2000). Therefore, the purified recombinant full-length human LH1 and LH3 and their A

fragments and fragment pairs A-B and B-C were analysed for collagen glucosyltransferase activity. As reported previously (Heikkinen *et al.* 2000), the recombinant LH3 had collagen glucosyltransferase activity, while LH1 did not (Table III in I). The glucosyltransferase activity of LH3 was found to be located in its N terminus as its A fragment and fragment pair A-B both had glucosyltransferase activity, while the fragment pair B-C did not, nor did any fragment of LH1 (Table III in I). In addition, recombinant LH3 and its A fragment were found for the first time to also have a minor amount of collagen galactosyltransferase activity. This activity level was quite low, however, and in the case of some preparations even below the reliable detection limit. In any case, the galactosyltransferase activity of LH3 seems to be specific, as none of the LH1 fragments nor fragment pairs or the full-length LH1 had any of this activity, nor did the B-C fragment pair of LH3.

5.2 Substrate specificity of recombinant human lysyl hydroxylases (II)

In order to study the substrate specificity of LHs, the recombinant full-length human LHs 1-3 were expressed in High five insect cells as a fusion protein with a GP67 signal peptide and his-tag. The recombinant LHs were purified from the culture medium with metal chelate affinity chromatography and eluted with either imidazole or histidine (Figure 1 in II). LH activity was analysed using a method based on the hydroxylation-coupled decarboxylation of 2-oxo-[1-¹⁴C]glutarate with the synthetic peptide (Ile-Lys-Gly)₃ as a substrate. The activities of the enzyme preparations varied considerably and thus accurate comparison of the specific activities of LHs 1-3 was not possible. Therefore, in all further analyses the (Ile-Lys-Gly)₃ peptide was used as an internal control to obtain the relative hydroxylation efficiencies of each LH isoenzyme with various peptide substrates.

Synthetic peptides representing sequences of collagen I and IV were used as substrates to find out if the LHs showed substrate sequence specificity. The peptides were: 1) $\alpha 1(\text{I})$ -Ntelo and 2) $\alpha 1(\text{I})$ -Ctelo, representing the telopeptide crosslinking sites in the $\alpha 1$ chain of type I collagen, both contained one lysine residue that is known to be hydroxylated *in vivo*; 3) $\alpha 1(\text{I})$ -1, a peptide representing the triple-helical crosslinking site in the $\alpha 1(\text{I})$ chain that contains lysine 87, known to be hydroxylated *in vivo*, as well as two other potentially hydroxylatable lysines; 4) $\alpha 1(\text{I})$ -2, another peptide from the triple-helical region of the $\alpha 1(\text{I})$ chain that contains two potential hydroxylation sites; 5) $\alpha 1(\text{IV})$ -1 and 6) $\alpha 1(\text{IV})$ -2, two peptides from the triple-helical region of the $\alpha 1$ chain of type IV collagen that contain four and three potential hydroxylation sites, respectively; and 7) $\alpha 2(\text{IV})$ -7S, a peptide from the 7S domain of the $\alpha 2$ chain of type IV collagen with three potential hydroxylation sites (Figure 2 in II). The sequences of the synthetic peptides are shown in Table 1 in II.

The initial hydroxylation assays were done at a 200 μM peptide concentration. None of the LH isoenzymes catalysed hydroxylation of the $\alpha 1(\text{I})$ -Ntelo and $\alpha 1(\text{I})$ -Ctelo peptides (Figure 3 in II). *In vitro* hydroxylation of these peptides was also studied at

higher and lower peptide concentrations in case their K_m was particularly high or in case high peptide concentration caused product inhibition, respectively, but no activity was obtained with any of the isoenzymes. All three isoenzymes hydroxylated the $\alpha 1(I)-1$ peptide most efficiently, the activities being about 2-3 times higher than those obtained with the control (Ile-Lys-Gly)₃ peptide (Figure 3 in II). The $\alpha 1(I)-2$ peptide was hydroxylated less efficiently, especially by LH1, the activity being only about 20% of that obtained with the control, while the activities obtained with LH2 and LH3 were about 80% (Figure 3 in II). When compared with the activities obtained with the $\alpha 1(I)-1$ peptide, the relative activities obtained with $\alpha 1(I)-2$ were even lower, about 5% for LH1 and 40% for LH2 and LH3.

The peptides representing type IV collagen sequences were hydroxylated by all three isoenzymes. The $\alpha 1(IV)-1$ peptide representing the collagenous domain of type IV collagen was hydroxylated by all three LHs with activities of about 50-70% of those with (Ile-Lys-Gly)₃ (Figure 3 in II). The $\alpha 1(IV)-2$ peptide was hydroxylated more efficiently by LH2 and LH3, the relative activities being about 130-140% of that for the control peptide, while the activity of LH1 was markedly lower, about 60% (Figure 3 in II). LH1 had the highest relative activity with the $\alpha 2(IV)-7S$ peptide, about 60% of the control, while the activities of LH2 and LH3 were only about 30-35% (Figure 3 in II).

The K_m and V_{max} values of the LH isoenzymes for each synthetic peptide were also determined (Table 2 in II). The V_{max} values are expressed as relative values with respect to the V_{max} obtained with the (Ile-Lys-Gly)₃ control peptide (Table 2 in II). All three isoenzymes had about the same K_m values for the control (Ile-Lys-Gly)₃ peptide, 310-470 μM , but some differences in the values for the other peptide substrates were seen. LH1 had a K_m of 230 μM for the $\alpha 1(I)-1$ peptide and a V_{max} that was comparable to that obtained with the control peptide. LH2 and LH3 on the other hand had lower K_m values for $\alpha 1(I)-1$, 100 and 80 μM , respectively, but their V_{max} values were only 60-70% of the control. The K_m values of all three LHs were nearly identical for the $\alpha 1(I)-2$ peptide, 80-120 μM , with V_{max} values of 40% for LH2 and LH3, but only 5% for LH1 indicating that despite the relatively low K_m of LH1, it is inefficient at hydroxylating this peptide. LH2 and LH3 had nearly identical K_m values for the peptides representing collagen IV, in the range of 40-70 μM , but there was more diversity in the V_{max} values, which were about 10-60% of the control. The K_m of LH1 for $\alpha 1(IV)-1$ was 300 μM , i.e. 6-7.5-fold higher than that of LH2 and LH3, but the V_{max} values of all three isoenzymes for this peptide were quite similar, about 20-30% of the control. Also, in case of the $\alpha 2(IV)-7S$ peptide, LH1 had a higher K_m than LH2 and LH3, 240 μM when compared to 50-60 μM , but still LH1 hydroxylated this peptide more efficiently with a V_{max} value of 40%, while LH2 and LH3 had V_{max} values of about 10 and 20%. The K_m values of all LHs for the $\alpha 2(IV)-2$ peptide were quite similar, 40-80 μM , LH1 having the highest. LH1 was also the least effective isoenzyme in hydroxylating this peptide, with a relative V_{max} of only 7%, while the values for LH2 and LH3 were about 50-60%. It should be also noted that the K_m values of LH2 and LH3 are lower for the peptides representing collagen IV than for those representing collagen I (40-70 μM compared to 80-120 μM), while LH1 showed no such differences.

As the N and C telopeptides of type I collagen are known to be hydroxylated *in vivo* and as no activity was obtained with the short synthetic peptides, the hydroxylation of N telopeptides was analysed *in cellulose*. Each of the LH isoenzymes was coexpressed in

insect cells with full-length pro α 1(I) chains and the two subunits of C-P4H (Myllyharju *et al.* 1997, Nokelainen *et al.* 1998) in 3-4 independent experiments. The recombinant type I procollagen homotrimers were digested with pepsin, the collagen homotrimers were purified and analysed by N-terminal sequencing to determine if the lysine in the N telopeptide was hydroxylated. Amino acid analysis was employed to determine the overall lysine hydroxylation status of the recombinant collagen I homotrimers. The LH virus was omitted in control experiments.

N-terminal sequencing showed that when coexpressed with LH2, about 25% of the telopeptide Lys9 was hydroxylated, while the other isoenzymes were not able to hydroxylate this residue (Table 3 in II). The highest overall hydroxylation level [Hyl/(Hyl+Lys)] obtained with all three isoenzymes was about 30% and no hydroxylysine was detected when the procollagen chains were expressed without a recombinant LH, indicating that while all three LHs were active and capable of hydroxylating the helical lysines to a similar extent, only LH2 was able to hydroxylate the lysine in the N telopeptide (Table 3 in II). We further verified that LH1 and LH2 were equally efficient at hydroxylating the helical lysines by analysing the *in vitro* hydroxylation of recombinant human collagen α 1(I) chains purified from the yeast *Pichia pastoris* (Nokelainen *et al.* 2001). Equal amounts of LH1 and LH2 hydroxylated about 18% and 16% of the total lysines present in the chains, respectively.

5.3 Lack of lysyl hydroxylase 3 leads to fragmentation of basement membranes and embryonic lethality (III)

In order to analyse the *in vivo* roles of LH3 a knock out mouse line was generated by targeting *Plod3* with a construct containing a *LacZ* gene in-frame with exon 1 of *Plod3* followed by a neomycin resistance gene for a selection of ES cells (Figure 1 in III). ES cells positive for the correct recombination of the targeting construct were injected into blastocysts and transferred into pseudopregnant females, and the best chimeras obtained were mated with wild type C57BL/6 mice to obtain mice heterozygous for the targeted *Plod3* allele. The heterozygous mice appeared normal. Surprisingly, mating of the *Plod3*^{+/-} mice gave only wild type and heterozygous offspring and lack of functional LH3 was found to be lethal around E9 (Table 1 in III).

The *Plod3*^{-/-} embryos were clearly smaller than their littermates and very fragile (Figure 2A in III). To study if the early embryonic death was caused by a placental defect, placentas were analysed but no obvious defects were found, although staining for β -galactosidase activity showed strong LH3 expression in the vasculature of the placenta (Figure 2B, C in III). Histological and transmission EM examination of the embryos showed that the *Plod3*^{-/-} embryos have a markedly lower mesenchymal cell density than the wild type controls (Figure 2D-G and 3A, B in III). Transmission EM of the null embryos showed fragmentation and detachment of basement membranes (Figure 3D-E in III) and immunofluorescence analysis with an antibody against collagen IV revealed weak or almost absent staining in the basement membranes, and abnormal localisation of this network forming collagen in the null embryos (Figure 4D-F in III). The basement

membranes of the null embryos stained for the laminin $\gamma 1$ chain (Figure 4B in III) and perlecan, but they appeared thinner and fragmented. Immuno EM clearly showed that collagen IV is not localised into the basement membranes in the null embryos, but instead is found in unusual aggregates in the extracellular space, some of these aggregates being visible inside the cells in the secretory vesicles (Figure 5B-C). Collagen IV staining is also seen in the dilated ER (Figure 5D).

Further analysis of type IV collagen isolated from the null embryos revealed that the quantities of $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains were similar to those in the wild type embryos, but they migrated faster in SDS-PAGE (Figure 6 in III). The difference in the migration could be explained by lack of post-translational modification, like glycosylation, as the hydroxylysines in type IV collagen are known to be efficiently glycosylated. When the LH and glycosyltransferase activities of the null embryos were analysed it was found that in the *Plod3*^{-/-} embryos LH and galactosyltransferase activities were normal, but the glucosyltransferase activity was only about 15% of the wild type values indicating that the glucosyltransferase activity of LH3 is essential for proper collagen IV processing and embryonal development.

The expression pattern of LH3 was analysed by staining for β -galactosidase activity driven by the *Plod3* promoter. Staining was ubiquitous and intense in the null and heterozygous embryos (Figure 2A and 7A in III) and the staining pattern became more specific as the heterozygous embryos developed further (Figure 2B in III). In adult *Plod3*^{+/-} mice β -galactosidase staining could be seen in the chondrocytes, capillary walls and perimysia of the muscles, and alveolar capillaries and blood vessel walls of the lungs (Figure 7H, J, L in III). In the kidneys, intense staining was seen in the vascular poles and mesangia of the glomeruli (Figure 7I, K in III), and scattered staining in the tubuli. Intense staining was also visible in the adrenal gland and single hepatocytes of the liver (not shown). In the developing eye the lens was intensely stained at E12.5 (Figure 7C in III), staining was weaker at E14.5 (Figure 7D in III), and in the adult lens staining was seen only in the lens epithelial cells (Figure 7E in III). In the adult eye staining was also visible in the ciliary body, capillaries of the retina and the inner segment of the photoreceptors (Figure 7E, F in III).

5.4 Lack of lysyl hydroxylase 1 in mice leads to increased mortality due to aortic rupture (IV)

Mouse lines lacking LH1 activity were generated as a model for human EDS VIA and in order to analyse the *in vivo* roles of LH1 and to study its expression pattern. The *Plod1* gene was inactivated by homologous recombination of a targeting construct (Figure 6 and Figure 1A in IV) that led to deletion of exons 3-6 and in-frame fusion of the *LacZNeo* reporter cassette to exon 2. Two separate mouse lines were generated from the targeted ES cells by routine methods (Figure 1B in IV) and backcrossed into the C57BL/6 mouse line. The lack of LH1 mRNA in the null mouse lines was confirmed by RT-PCR with two oligo pairs downstream of the recombination site, i.e. from exons 10 and 12 and 18 and 19, the latter exons encoding catalytically critical amino acids required for the binding of

Fe²⁺ and the C-5 carboxyl group of 2-oxoglutarate. As no β -galactosidase staining in these mouse lines could be detected, an additional mouse line was generated from the ES-cells LH1:RST531 obtained from the Gene Trap consortium. This mouse line was identical with the two other targeted *Plod1*^{-/-} mouse lines in all aspects analysed, and tissues originating from this mouse line were used to study the expression pattern of the *Plod1* gene (Figures 2 and 4B in IV).

Staining for β -galactosidase activity in the tissue samples of the LH1:RST531 mouse line showed that expression of LH1 was associated with cells producing fibrillar collagens. For example, intense staining was seen in the fibroblasts of the lungs, skin and tendons (Figure 2A, B, D in IV). In the heart a very intense staining was seen in the auricle (Figure 2C in IV), while in the heart muscle the staining was weaker (not shown). In the muscle tissue the epimysium was stained (Figure 2D in IV), as well as some nuclei (not shown). Intense staining was also observed in the periosteum, chondrocytes and osteoblasts of the bone (Figure 2E in IV). In addition, keratocytes arranged between collagen fibres in the cornea (Figure 2F in IV) and papilla of hair follicles and arrector pili muscles of the skin (Figure 2A in IV) were stained.

The *Plod1*^{-/-} mice were viable and could reproduce normally and they showed no signs of the kyphoscoliosis that is characteristic of the human EDS VIA syndrome. However, the null mice appeared to be slower than their wild type littermates and they were passive and hypotonic when handled. As the mice got older and gained weight difficulties in moving became more obvious. When tested for walking on a rod or on top of the wall of a plastic container the wild type mice had no difficulties in moving and ran fast (Figure 3A, C in IV). In contrast, the null mice either did not move much or got tired very quickly (Figure 3B, D in IV).

Although the null mice were viable an increased mortality rate was observed. About 17% of the *Plod1*^{-/-} male and 9% of the *Plod1*^{-/-} females were found dead in the morning in their cages. Mice are most active during the night and these sudden deaths occurred during the active period. Autopsies revealed haemorrhages mainly in the thoracic or abdominal cavities (Table 1 in IV). Histological analysis revealed aortic dissections (Figure 4C in IV) and in one case rupture of the media and a scar like accumulation of cells subjacent to the adventitia (Figure 4D in IV) possibly due to a previous injury. The ruptures were located between the external elastic lamellae, while the elastic layers were intact (Figure 4C-D, F in IV).

The thickness and the structure of the aortic wall in the healthy *Plod1*^{-/-} mice was comparable to that of their wild type littermates (Figure 4A, B in IV) and echocardiographic analysis showed no alterations in the structure or function of the heart (not shown). The vascular smooth muscle cells appeared less organised in the *Plod1*^{-/-} mice, however (Figure 4B in IV). More detailed analysis with transmission EM revealed degenerative changes in the aorta of the apparently healthy *Plod1*^{-/-} mice as well, such as vacuolisation and mitochondrial swelling in the smooth muscle cells of the aortic walls and even some completely degenerated cells (Figure 4F in IV).

The skin of the *Plod1*^{-/-} mice appeared normal and no hyperextensibility or abnormal scars were detected, although they are common findings in human EDS VIA. The histology of the skin at light microscopy level was also normal. However, EM analysis revealed changes in the collagen fibres. In the *Plod1*^{-/-} mice the variation in the collagen fibril diameters was larger than in the wild type in both the aorta and skin and the mean

fibril diameters were increased (Figure 5 in IV). The mean fibril diameter was 67 nm in the *Plod1*^{-/-} aorta and 44 nm in the wild type, the corresponding values in the skin being 91 nm and 62 nm. In addition to the increased diameter, the shape of the fibrils in the *Plod1*^{-/-} tissues was irregular (Figures 5B, F) and the fibril diameter also varied within a single fibril (not shown). The diameters of the collagen fibrils of the cornea were the same in both genotypes (not shown).

In human EDS VIA patients one diagnostic method is to measure the LH activity of cultured fibroblasts. Here the LH activity was analysed from skin and aorta homogenates using [¹⁴C]-lysine labelled non-hydroxylated procollagen as a substrate. The LH activity level of the skin samples of *Plod1*^{-/-} mice was about 35% of that of the wild type samples, and in the aorta samples about 45% (Table 2 in IV). The *Plod*^{+/-} mouse samples had intermediate activities. These surprisingly high activity levels indicate that either LH2 or LH3 or both are active in the tissues analysed but are not able to fully compensate for the lack of LH1 activity.

The total hydroxylysine content in various tissues of the *Plod1*^{-/-} and wild type mice were analysed (Table 3 in IV). Whole tissue samples were used without any further purification of the different collagen types and the differences observed between various tissues thus reflect their different collagen compositions. In the wild type animals the highest hydroxylysine content (55.43 residues per collagen triple helix, i.e. three α chains) was found in the lungs and the lowest (21.13 residues) in the skin. The amount of hydroxylysine was decreased in all *Plod1*^{-/-} tissues analysed when compared to the wild type mice. However, the decrease in the total hydroxylysine content varied markedly between different tissues, the lowest amount being seen in the skin, which also had the most dramatic decrease in the amount of hydroxylysine, being only 22% of that of the wild type skin. Tail tendons also had a markedly decreased (24%) hydroxylysine content while in the lungs and femurs the decrease was less dramatic, being 86% and 75% of that of the wild type tissues, respectively.

The amounts of trifunctional crosslinks, HP and LP, were analysed in five tissues without further purification of the different collagen types (Table 4 in IV). These pyridinolines could not be detected reliably in either the wild type or mutant skin due to the low amounts in this tissue. The collagen HP content varied markedly between the different wild type tissues, the highest value, 0.7 residues per collagen triple helix, being obtained from the aorta, and the lowest, 0.05 residues per triple helix, from tail tendons and corneas. The HP content was decreased in all the *Plod1*^{-/-} tissues analysed, the decrease being largest in the aorta (28% of the wild type), corneas (33%) and lungs (34%); while a more moderate decrease was seen in the tail tendons (59%) and femurs (47%) (Table 4 in IV). Interestingly, there was no correlation between the changes in the hydroxylysine and the HP content. For example, in the tail tendons of *Plod1*^{-/-} mice the hydroxylysine content was only 24% of the wild type and still the amount of HP crosslinks was 59% of the wild type values.

As expected, the LP content of all LH1 null tissues was increased relative to the wild type tissues, and thus the LP/HP ratio was also increased markedly, ranging from about 0.12 to 0.39 in the wild type tissues and from 3.8 to 12.5 in the null tissues (Table 4 in IV). The total amount of trifunctional crosslinks (HP + LP) was increased in the null tissues to 140-235% in the aorta, femurs, lungs and corneas and to even as much as 720% in the tail tendons (Table 4 in IV).

6 Discussion

6.1 Expression of recombinant lysyl hydroxylases in insect cells and characterisation of their domain structure and function (I)

LHs are lumenally oriented peripheral ER membrane proteins and their complete solubilisation from tissues is difficult (Kivirikko and Myllylä 1982, Kivirikko *et al.* 1992). The most commonly used methods of purifying endogenous vertebrate LH1 involve the use of high salt buffers containing detergents, affinity chromatography in concanavalin A and/or denatured collagen columns and gel filtration. LH1 has been shown to bind strongly to various chromatographic materials through hydrophobic interactions and its efficient elution is greatly facilitated by the use of up to 60% ethylene glycol. Additional problems in the purification of endogenous LH1 are caused by its tendency to become easily inactivated and aggregated. Therefore, the recoveries of pure LH1 from tissues have been low, typically only 4%. (Turpeenniemi *et al.* 1977, Turpeenniemi-Hujanen *et al.* 1980, 1981, Kivirikko and Myllylä 1982.) Recombinant LH1 has been expressed at high levels in the ER of insect cells, but its purification was impeded by difficulties similar to those described above (Pirkanen *et al.* 1996). Purification of recombinant LH3 from the ER of insect cells to homogeneity utilising an N-terminal histidine-tag and metal affinity chromatography has been reported, but the purified enzyme had no LH activity due to inactivation caused by the imidazole used in the elution (Heikkinen *et al.* 2000). Furthermore, it has been reported that efficient secretion of recombinant LH1 from insect cells to the culture medium is obtained by the use of a signal peptide from the baculoviral envelope protein GP67 (Krol *et al.* 1996).

In order to be able to study the structural properties of a protein, relatively large amounts of pure starting material are needed. In this work we set out to develop an efficient and simple method for the purification of the three full length recombinant human LHs from insect cells. We decided to use the GP67 signal sequence, which leads to the secretion of the produced proteins into the culture medium. There they were present in a membrane-free form and were less prone to aggregation, and an N-terminal histidine-tag was used to facilitate their rapid purification. As previous studies have

shown that imidazole has a deleterious effect on LH activity, it was rapidly removed after elution of the recombinant LHs with a PD-10 gel filtration column. By using this system, all three recombinant LHs were, for the first time, purified to homogeneity as active enzymes. Considerable variation in the LH activity in different purified preparations of the same isoenzyme was observed (Table I in I), however, the specific activity of each isoenzyme ranged from 10 to 65 mol/mol/min. Similar variability has also been observed between different samples of purified non-recombinant LH1, the highest specific activities measured being 60 to 100 mol/mol/min (Turpeenniemi-Hujanen *et al.* 1980, 1981). Therefore, it is important to note that it is not possible to make exact comparisons between the specific activities of the three LHs, and a suitable internal control is required when comparing them, for example, the substrate preference of the different isoenzymes (see 6.2). Based on gel filtration data all three recombinant active LHs were homodimers with an apparent molecular weight of about 180 kDa, indicating that they had a similar molecular composition as reported previously for the LH1 isolated from chick embryos and human placenta (Turpeenniemi-Hujanen *et al.* 1980, 1981, Myllylä *et al.* 1988).

A structural domain in a protein is usually defined as a self-stabilising element in the overall structure that often folds independently of the rest of the protein. A convenient biochemical tool to study the domain composition of a protein is limited protease digestion, which can reveal protease sensitive non-folded hinge regions between folded protease resistant structural domains. (Hubbard 1998.) Limited proteolysis of the recombinant LHs with different proteases revealed the existence of at least two protease sensitive regions, suggesting that the LHs contained three different domains, termed A, B and C, from the N to C terminus. Only the N-terminal A domains could be expressed in insect cells as secretable and soluble recombinant polypeptides, however, the B and C domains were mainly found to accumulate into insoluble aggregates within the cells. In contrast, the domain pairs A-B and B-C were secreted and could be purified to homogeneity. The purified recombinant LH1-3 domains A and domain pairs A-B and B-C were folded based on transverse urea-gradient gel electrophoresis and CD spectrum analysis.

When the LH activity of these domains and domain pairs was analysed it was found that the N-terminal A domain played no role in the LH activity, as the B-C double domains were fully active LHs with K_m values for the peptide substrate (Ile-Lys-Gly)₃ identical to those of the full length proteins (Tables I and II in I). In the other main collagen hydroxylase, prolyl 4-hydroxylase, a peptide substrate binding domain separate from the catalytic domain has been identified in the α subunit (Myllyharju & Kivirikko 1999, Hieta *et al.* 2003, Pekkala *et al.* 2004). No amino acid sequence similarity was observed between the peptide substrate binding domain of the collagen prolyl 4-hydroxylases and the LHs. Obviously, the A domain of LHs does not function as a substrate binding site for lysyl hydroxylation. Whether this site is present in domain B or C or whether both domains are required for the binding to occur, is currently unknown.

The surprising finding that LH3, but not LH1 or LH2, also possessed collagen glucosyltransferase activity (Heikkinen *et al.* 2000), was confirmed here. We also showed that the N-terminal A domain of LH3, consisting of residues 1-266, was responsible for this activity as it had essentially the same amount of glucosyltransferase activity as full length LH3 (Table III in I). Site-directed mutagenesis studies by others have since shown that the A domain residues Cys144, Leu208 and Asp187-191 are critical for the

glucosyltransferase activity (Wang *et al.* 2002b). Interestingly, others have shown that an N-terminal LH3 fragment containing residues 1-232 has no glucosyltransferase activity, while fragments containing residues 1-520, 1-401 and 1-388 possess this activity, but the amount is reduced when compared to that of full length LH3 (Heikkinen *et al.* 2000, Wang *et al.* 2002b). These results may be caused by incorrect folding of the LH3 fragments as their C termini do not conform with the domain boundaries. According to our studies the A domain and the B-C domain pair of LH3 have essentially the same amounts of glucosyltransferase and LH activity, respectively, as the full length LH3. This indicates that distinct binding sites for the collagen polypeptide substrate are utilised in LH3 for the two reactions.

We showed for the first time that in addition to glucosyltransferase activity, LH3 also has a minor amount of collagen galactosyltransferase activity. This has since been confirmed by others, and the amino acids shown to be critical for the glucosyltransferase activity were also found to be required for the galactosyltransferase activity (Wang *et al.* 2002a). The galactosyltransferase activity of LH3 in the *in vitro* assay system was much lower, up to 63 mol/mol/min, than the glucosyltransferase activity of LH3, which was up to 920 mol/mol/min, however. The glucosyltransferase activity of the purified recombinant LH3 obtained here was only about 2% of that obtained with purified collagen glucosyltransferase from chick embryos (Anttinen *et al.* 1977, Kivirikko & Myllylä, 1982). Also, the K_m values for UDP-glucose and denatured citrate-soluble rat skin collagen were 4 to 5 fold higher than the highest values reported for collagen glucosyltransferases obtained from other sources (Kivirikko & Myllylä 1982). The galactosyltransferase activities of recombinant LH3 or its domains were so low that no K_m values could be determined. To control our assay system, we measured the glucosyltransferase activities of human serum samples simultaneously with recombinant LH3 and they were within the normal range, indicating that the assay was functional. Based on gel filtration studies the active galactosyltransferases purified from chick embryos had molecular weights of about 450, 200 and 50 kDa (Risteli *et al.* 1976) and the glucosyltransferase had a molecular weight of about 80 kDa (Myllylä *et al.* 1977). Here we determined that the full length purified recombinant LH3 had a molecular weight of about 180 kDa, indicating that the protein should be a dimer. The glucosyltransferase purified from chick embryos with a molecular weight of 80 kDa could possibly represent a monomeric LH3, but the partially purified galactosyltransferase clearly had a different molecular weight from the purified LH3. Although the glycosyltransferase activities of LH3 are specific and real, as no such activities could be detected with LH1 and LH2, in light of the previous results with partly purified collagen glycosyltransferases the physiological roles of these LH3 activities are not clear, nor is it clear whether additional collagen glycosyltransferases exist.

6.2 Substrate specificity of lysyl hydroxylases (II)

The lysine hydroxylation level varies considerably between collagen types and even within the same collagen type in different tissues (Kivirikko *et al.* 1992, Myllyharju 2005). Furthermore, lack of LH1 and LH2 activity in two heritable human diseases has

been shown to cause changes in lysine hydroxylation and collagen crosslinking in a tissue, collagen type and hydroxylation site specific manner. In EDS VIA, which is caused by a lack of LH1 activity, the amount of helical hydroxylysine is decreased in type I and III collagens, while types II, IV and V have a normal hydroxylysine content in the triple helix (Ihme *et al.* 1984). Additionally, the type I collagen isolated from skin completely lacked helical hydroxylysine residues, while that from bone, tendons, lungs and kidneys had a less dramatic decrease in the hydroxylysine content (Ihme *et al.* 1984). In patients suffering from BS caused by mutations in the gene for LH2, the telopeptide lysines of the type I collagen from bone are severely underhydroxylated, while type II collagen from cartilage and type I collagen from ligaments have a normal telopeptide hydroxylation pattern, hydroxylation of the helical lysines being normal in each case (Bank *et al.* 1999). These findings suggest that the LH isoenzymes may show tissue and/or substrate specificity. However, no obvious differences have been found in the expression patterns of the three LH mRNAs in various tissues, all are widely expressed (see Tables 2 and 3), and the expression patterns of the mRNAs of different LHs and collagen types show no correlation in cultured cells (Wang *et al.* 2000).

In this study we used several short synthetic peptides representing potential lysyl hydroxylation sites in type I and IV collagens to analyse the proposed substrate specificity of the three LHs *in vitro*. Of special interest were the proposed roles of LH1 and LH2 as the specific helical and telopeptide LHs, respectively. Surprisingly, none of the recombinant purified LHs, not even LH2, was able to hydroxylate the lysine in the peptides representing the N and C telopeptides of the collagen $\alpha 1(I)$ chain. Previously it has been reported that LH1 purified from chick embryos was not able to hydroxylate the CNBr cleavage-derived $\alpha 1(I)$ and $\alpha 2(I)$ peptides containing the N telopeptide hydroxylation site. The N telopeptide lysines were not hydroxylated either when nonhydroxylated type I procollagen chains were used as a substrate (Royce & Barnes 1985). In the same study about 37% of the helical lysines were hydroxylated indicating that the purified LH1 was active (Royce & Barnes 1985). However, when we coexpressed each of the LHs with the full-length pro $\alpha 1(I)$ chains *in vivo* in insect cells, LH2 was able to hydroxylate the N telopeptide lysine, while LH1 and LH3 were not. This is the first direct evidence that LH2 is indeed responsible for hydroxylation of the telopeptide lysines. Due to the technical difficulty of C-terminal sequencing, hydroxylation of the C telopeptide lysine was not analysed. In addition to the findings in BS, the role of LH2 as a specific telopeptide hydroxylase is supported by many reports associating overexpression of LH2 and increased telopeptide hydroxylation, and the subsequent increased formation of pyridinoline crosslinks in fibrotic conditions (Uzawa *et al.* 1999, Mercer *et al.* 2003, Pornprasertsuk *et al.* 2004, van der Slot *et al.* 2004, 2005, van der Slot-Verhoeven *et al.* 2005). In our study LH2 hydroxylated the N telopeptide lysine only when present in a longer peptide, in this case, the full length procollagen chain. The minimum length of the peptide substrate required for the hydroxylation of the N telopeptide lysine by LH2 still needs to be determined, however, as well as whether the increase in length must be in the form of a repeated -X-Y-Gly- sequence. LHs in general do not hydroxylate free lysine residues but the tripeptide Ile-Lys-Gly already serves as a substrate (Kivirikko *et al.* 1992). Recently a novel hydroxylysine in a Leu-Hyl-Ala sequence was identified in the NC1 domain of the $\alpha 1$ chain of type IV collagen, where it takes part in a novel Hyl-Met crosslink stabilising the collagen IV network (Vanacore *et*

al. 2005). It will be interesting to find out if this motif is also preferentially hydroxylated by LH2 and if so what length of peptide is required for hydroxylation.

All three LHs were able to hydroxylate all of the synthetic peptides used to represent the helical hydroxylation sites of the type I and IV collagen chains. Some differences in their K_m and relative V_{max} values for a given peptide were observed, but they were not substantial enough to draw the conclusion that one isoenzyme was more specific for helical hydroxylation sites in general than another. Furthermore, all three recombinant LHs were able to hydroxylate up to 30% of the lysine residues of full length human pro α 1(I) chains when coexpressed in insect cells, and LH1 and LH2 were able to hydroxylate about 18% and 16%, respectively, of the lysine residues of recombinant human α 1(I) chains *in vitro*. However, as we analysed only the total hydroxylation level in these experiments, our data do not exclude the possibility that particular lysine residues in the triple helical region might be specifically or preferentially hydroxylated by only a certain LH. Similarly, when binding and hydroxylation of 727 short collagenous peptides was analysed, some preferences but no strict sequence specificities between the LHs were identified (Risteli *et al.* 2004). A positive charge on the peptide and the presence of some specific amino acids close to the lysine residue were found to enhance the binding of all LHs (Risteli *et al.* 2004). It has previously been speculated that positively charged sequence motifs, including histidines, may be required for efficient hydroxylation by LH1 (Eyre *et al.* 1984).

6.3 Inactivation of *Plod3* leads to abnormal type IV collagen processing and fragmentation of basement membranes leading to embryonal lethality (III)

Hydroxylysines are not required for collagen fibril formation as recombinant fibrillar collagens produced *in vitro* with no hydroxylysines can form native-type fibrils (Myllyharju *et al.* 2000, Nokelainen *et al.* 2001). However, the amount of hydroxylysines and their associated glycosylation have been shown to be involved in the control of fibril diameter (Notbohm *et al.* 1999). Additionally, the human diseases caused by lack of LH1 activity and mutations in the genes for LH2, EDS VIA and BS, respectively, are not lethal (Bank *et al.* 1999, Steinmann *et al.* 2002, van der Slot *et al.* 2003). In this study we generated the first mouse line lacking the activity of a specific LH isoenzyme, i.e. LH3. To our surprise lack of LH3 was found to be lethal during early embryogenesis due to abnormal supramolecular assembly and localisation of the network forming type IV collagen, resulting in fragmentation of basement membranes. The null embryos developed in a seemingly normal way until E8.5 and after that showed growth retardation and were very fragile when handled. No obvious defects were found in the placentas, but we cannot completely rule out the possibility of abnormalities in the vasculature of the placenta that, based on staining for β -galactosidase activity, expresses high amounts of LH3. The total LH activity and galactosyltransferase activity in the *Plod3*^{-/-} embryos was normal when compared to that of their wild type littermates, but in contrast, the glucosyltransferase activity was markedly decreased. Furthermore, the increased mobility

of collagen IV chains in the *Plod3*^{-/-} embryos was consistent with a lack of hydroxylysine-linked carbohydrates. The data obtained in article II of this thesis indicate that LH2 and LH3 may be slightly more efficient in hydroxylating collagenous sequences present in type IV collagen than LH1. Interestingly, it has been recently shown that when the LH activity, but not the glycosyltransferase activities of LH3, is missing in mice, a reduction in the hydroxylysine content to about 70% of that of the wild type occurs in the type IV and V collagens in kidneys and lungs, while the hydroxylation of lysine residues in the type I, II and III collagens is normal (Ruotsalainen *et al.* 2006). Furthermore, these mutant mice are viable and the type IV collagen is localised normally to basement membranes that are only slightly thinner than in the wild type mice (Ruotsalainen *et al.* 2006). Altogether, our data and the subsequent data indicate that it is the lack of glucosyltransferase activity in LH3 that is critical for the assembly of collagen IV into normal network-like supramolecular structures. The importance of glucosylation is also demonstrated by LH3 mutant mouse lines that have varying levels of glucosyltransferase activity. Embryos with glucosyltransferase activities of 11-20% of the wild type survive until E10.5-14.5 and their survival directly correlates with the remaining activity level (Ruotsalainen *et al.* 2006).

Our data indicate that a portion of the type IV collagen synthesised in the *Plod3*^{-/-} mice is retained in the dilated ER. This is consistent with previous findings about the quality control of collagen synthesis, in that abnormal collagen molecules are retained within the ER through chaperones and degraded via an ER-associated lysosomal or proteasomal degradation pathway (Lamandé *et al.* 1995, Fitzgerald *et al.* 1999, Lamandé and Bateman 1999, Gotkin *et al.* 2004, Wilson *et al.* 2005). In our Western blot analysis with antibodies against the type IV collagen, the migration of $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains from the null embryos was faster than in the control mice indicating changes in their post-translational modification, i.e. lack of hydroxylysine-linked glycosylation. In the immunofluorescence analysis, abnormal type IV collagen aggregates were seen both inside the cells and in the extracellular space. In the immuno-EM analysis these collagen IV aggregates were seen in the secretory vesicles, but not in the ER, indicating that aggregation of the abnormal type IV collagen molecules occurs in the secretory pathway.

LH1 and LH2, which lack collagen glycosyltransferase activity, were not able to compensate for the lack of LH3, although they are expressed in mice at least from E7.5 onwards (Hjalt *et al.* 2001, Salo *et al.* 2006b). In evolutionary terms, LH3 is the most conserved LH (Ruotsalainen *et al.* 1999). Disruption of the gene for the only LH present in *C. elegans*, *let-268*, is also lethal due to impaired processing and secretion of type IV collagen, which leads to abnormal basement membranes and developmental arrest at the two cell stage (Norman & Moerman 2000). The *C. elegans* LH also has glucosyltransferase activity as well as some galactosyltransferase activity (Wang *et al.* 2002b). The zebrafish LH3, *diwanka*, is required for motor growth cone migration from the spinal cord to periphery and, interestingly, *diwanka* mutant embryos have defects in type II collagen secretion (Schneider & Granato 2006). We have shown in collaboration with Drs. Schneider and Granato that *diwanka* also has glucosyltransferase activity (unpublished), and the N-terminal domain of *diwanka* containing the glucosyltransferase activity is able to rescue the motor axonal phenotype (Schneider & Granato 2006). Mice lacking type IV collagen $\alpha 1$ and $\alpha 2$ chains survive longer than the *Plod3*^{-/-} mice, until E10.5-E11.5 (Pöschl *et al.* 2004). These data indicate that in the early steps of basement

membrane assembly collagen IV is not required for the correct deposition of other basement membrane components and the role of type IV collagen becomes essential only at later stages when mechanical demands increase (Pöschl *et al.* 2004). Type IV collagen is not the sole substrate of LH3 and abnormalities in other basement membrane collagens, like the type XVIII collagen, may affect the basement membrane phenotype in the *Plod3*^{-/-} mice. Our preliminary data also show differences in the distribution of type XVIII collagen in the *Plod3*^{-/-} mice (unpublished) and this collagen was recently identified as a substrate for LH3 in zebrafish (Schneider & Granato 2006). Type XVIII collagen is found in the epithelial and endothelial basement membranes and its deficiency leads to different ocular defects, abnormal maturation of the retinal vasculature and abnormalities in the basement membranes (Fukai *et al.* 2002, Marneros & Olsen 2003, Ylikärppä *et al.* 2003a, Ylikärppä *et al.* 2003b, Hurskainen *et al.* 2005). Lack of type XVIII collagen in mice also leads to broadening of basement membranes, which in the case of choroid plexuses leads to hydrocephalus in a mouse line susceptible for it (Utriainen *et al.* 2004).

The laminin family of basement membrane proteins is crucial for both the assembly and the maintenance of basement membranes. This is clearly illustrated by several mouse lines which lack one of the laminin family members with phenotypes varying from early post-implantation lethality to mild muscular dystrophy and impaired microvessel maturity (reviewed in Miner & Yurchenco 2004, Yurchenco *et al.* 2004). Nidogen-1 and nidogen-2 are ubiquitous basement membrane proteins and it has been proposed that they act as linkers between the laminin and collagen IV networks of basement membranes and integrate other basement membrane components into it. Surprisingly, mice lacking nidogen-1 or -2 are healthy (Murshed *et al.* 2000, Miosge *et al.* 2002, Schymeinsky *et al.* 2002), suggesting complementary functions, but double null mice die at the perinatal stage and show basement membrane defects that abolish further development of the lungs and disrupt the integrity of cardiac tissues (Bader *et al.* 2005). Another ubiquitous basement membrane component is perlecan, a heparan sulfate proteoglycan (Knox & Whitelock 2006). Perlecan null mice develop embryonic basement membranes, but they are unstable and discontinuous leading to lethality, mainly at E10-E12.5, due to disruption of basement membranes in the contracting myocardium and cardiac arrest. A minority of the perlecan null mice survive to the perinatal period, but die because of severe brain and skeletal defects and respiratory failure. (Arikawa-Hirasawa *et al.* 1999, Costell *et al.* 1999.) The heparan sulfate side chains of perlecan have also been shown to have important roles in the proper structure and function of basement membranes as their absence in mice leads to abnormalities in the eyes, enhanced carotid artery intimal hyperplasia, impaired angiogenesis, delayed wound healing, proteinuria and retarded tumour growth (Rossi *et al.* 2003, Tran *et al.* 2004, Zhou *et al.* 2004, Morita *et al.* 2005).

6.4 *Plod1*^{-/-} mice as a disease model for human EDS VIA

The human EDS VIA is characterised by progressive kyphoscoliosis, severe muscular hypotonia at birth, generalised joint laxity with occasional hypermobility and luxation, and sometimes eye phenotypes (Steinmann *et al.* 2002). We created *Plod1*^{-/-} mouse lines

to study the *in vivo* functions of LH1, and also as a disease model for EDS VIA. *Plod1*^{-/-} mice are viable and reproduce normally and no scoliosis was detected. This may be caused by differences in the anatomy, size and posture between mice and humans. Muscular hypotonia was detected in the *Plod1*^{-/-} mice and it became more obvious as the mice gained age and weight. The muscular hypotonia was not as severe as in humans, where some infants have been reported to even have difficulties in feeding (Steinmann *et al.* 2002). The *Plod1*^{-/-} mice were flaccid when lifted and had clear gait abnormalities. So far no muscular hypotonia has been reported in the animal models of EDS, and *Plod1*^{-/-} mice could potentially serve as a model to study the role of the extracellular matrix in muscular diseases. Mutations in the genes coding for the three α chains of type VI collagen cause hereditary myopathies, an autosomal dominant Bethlem myopathy or an autosomal recessive Ullrich disease (Steinmann *et al.* 2002). *Col6a1* null mice display muscular abnormalities including irregular muscle fibre diameter, fibre necrosis and impaired contractile strength of the muscles (Bonaldo *et al.* 1998, Irwin *et al.* 2003). It will be of interest to study whether the *Plod1*^{-/-} mice show similar abnormalities.

EDS VIA patients have a soft and fragile skin with a tendency to severe scarring. Other symptoms observed in EDS VIA patients are tissue fragility and easy bruising, a Marfanoid habitus, microcornea and osteopenia without tendency to fractures. Arterial ruptures are also found in EDS VIA patients and they are the major life-threatening complication in this syndrome. (Steinmann *et al.* 2002.) One EDS VIA patient is known to have died because of arterial rupture at the age of 14 (Yeowell *et al.* 2000a). In the *Plod1*^{-/-} mice increased mortality was observed due to aortic ruptures occurring during the most active period of the day, i.e. during the night. Male null mice were more susceptible to aortic ruptures than females and also at a younger age. The heart function and macroscopic structure were normal in the null mice as was the histology of the heart tissue at the light microscopy level. Using electron microscopy, degenerative changes in the smooth muscle cells and collagen fibrils of the aorta were observed in samples from young null mice without ruptures, these changes probably predisposed them to aortic rupture because of compromised mechanical strength.

In EDS VIA patients the LH activity of cultured skin fibroblasts is typically less than 25% of the control values. The LH activity in skin homogenates of *Plod1*^{-/-} mice was about 35% of the wild type values indicating clear differences between mice and humans. The lysine hydroxylation level of skin collagens in the *Plod1*^{-/-} mice was also clearly higher, about 20% higher than the wild type, than in EDS VIA patients, who have a typical hydroxylysine content of about 5% in their skin (Steinmann *et al.* 2002). These differences may explain the lack of any obvious skin phenotypes in the *Plod1*^{-/-} mice, although the collagen fibres of the skin had an increased diameter and an abnormal shape in the null mice.

The amount of hydroxylysines in the *Plod1*^{-/-} mice was decreased in all the tissues studied. However, the degree of the decrease varied markedly between tissues. The highest remaining hydroxylysine amounts were 86% of the wild type in the lungs and 75% in the femurs, while the lowest hydroxylysine amount, about 20% of the wild type, was observed in the skin and tail tendons. This indicates that the two remaining LHs, LH2 and LH3, are able to compensate for the lack of LH1 activity to varying degrees in different tissues, which may reflect differences in their tissue distribution. All LHs were expressed in all tissues studied, with the exception of mouse liver, which lacks LH2

expression, but the relative expression levels of the LH isoenzymes are as yet unknown (Valtavaara *et al.* 1997, 1998, Passoja *et al.* 1998, Yeowell & Walker 1999, Ruotsalainen *et al.* 1999, Salo *et al.* 2006b).

Collagen fibres are stabilised by intermolecular crosslinks that are derived from the allysine or hydroxyallysine pathways. The allysine pathway dominates in the skin and corneas, while the hydroxyallysine pathway occurs mainly in the bone, cartilage, ligaments and tendons. The latter pathway leads to trivalent HP and LP crosslinks. (Eyre & Wu 2005.) In the *Plod1*^{-/-} mice the HP level was decreased in all the tissues analysed, and again clear differences were observed between tissues. Interestingly, there was no correlation between the level of decrease in the total lysine hydroxylation and the HP content. This indicates that although LH2 and LH3 are able to hydroxylate collagenous sequences, there are differences in the abilities of the LHs to hydroxylate the specific helical lysines that take part in the crosslinks, and this also seems to be tissue dependent. It has been proposed that LH1 is the LH responsible for the hydroxylation of the lysines at the helical crosslinking sites due to a HP to LP shift in EDS VIA patients and our results from the LH1 null mice support this hypothesis. In addition, LH2 is thought to be responsible for the hydroxylation of telopeptide lysines, which is a prerequisite for the formation of both LP and HP crosslinks, as also indicated by our results (article II in this thesis) and as already discussed above. Interestingly, the total amount of HP + LP crosslinks was clearly higher in the *Plod1*^{-/-} mice than in the wild type. The reason for this increase is unknown and may be a way to respond to changes in the crosslinking pattern. It could be caused, for example, by the increased activity of LH2 or slower assembly of the newly synthesised collagen chains into triple-helical molecules, which would allow more time for LH2 to function.

EM analysis revealed changes in the diameter of the collagen fibres of the skin and aorta of the *Plod1*^{-/-} mice, the mean diameter being increased, while those of the corneas were indistinguishable from the wild type fibrils. In the skin and aorta of the null mice the shape of the collagen fibres was also abnormal. Collagen I fibrils contain small amounts of type III and V collagens that have been shown to regulate the thickness of the fibrils (Kielty & Grant 2002). Mutations in collagen III or collagen V cause vascular and classical types of EDS, respectively, the mouse models of which also show abnormal collagen fibrils (Liu *et al.* 1997, Wenstrup *et al.* 2006). Similar abnormalities have been observed in mice lacking small leucine-rich proteoglycans such as biglycan, decorin, fibromodulin and lumican that associate with collagen fibrils (Danielson *et al.* 1997, Liu *et al.* 1997, Chakravarti *et al.* 1998, Svensson *et al.* 1999, Corsi *et al.* 2002, Wenstrup *et al.* 2006). Perlecan has also been shown to affect type II collagen fibril assembly in the cartilage (Kvist *et al.* 2006). It is possible that the altered posttranslational modifications of the fibril-forming collagens in *Plod1*^{-/-} mice could affect the binding of various associated proteins and thus lead to abnormal assembly of the collagen fibres.

Collagen fibril assembly and determination of the thickness of fibrils can be influenced by many factors, including those mentioned above, and is not completely understood yet. The hydroxylysine-linked glycosylation itself affects the lateral growth of fibrils (Notbohm *et al.* 1999). In the *Plod1*^{-/-} mice the hydroxylation level of collagens is altered, which is likely to lead to altered glycosylation as well. The mechanisms causing changes in collagen fibril diameter are open for further studies and will probably benefit

from the availability of various mouse lines lacking one or more of the proteins involved in fibrillogenesis.

References

- Alberts B, Bray D, Lewis J, Raff M, Roberts K & Watson JD (1989) *Molecular Biology of the Cell*. 2nd ed, Garland Publishing Inc, New York.
- 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 α (I) α (II) β_2 tetramer. *J Biol Chem* 272: 17342-17348.
- Anttinen H, Myllylä R & Kivirikko KI (1977) Hydrophobic and carbohydrate-recognition chromatographies of collagen glucosyltransferase. *Eur J Biochem* 78: 11-17.
- Arikawa-Hirasawa E, Watanabe H, Takami H, Hassell JR, Yamada Y. (1999) Perlecan is essential for cartilage and cephalic development. *Nat Genet* 23: 354-8.
- Armstrong LC & Last JA (1995) Rat lysyl hydroxylase: molecular cloning, mRNA distribution and expression in a baculovirus system. *Biochim Biophys Acta* 1264: 93-102.
- Aszodi A, Legate KR, Nakchbandi I & Fässler R (2006) What mouse mutants teach us about extracellular matrix function. *Annu Rev Cell Dev Biol* 22: 591-621.
- Bader BL, Smyth N, Nedbal S, Miosge N, Baranowsky A, Mokkapati S, Murshed M & Nischt R (2005) Compound genetic ablation of nidogen 1 and 2 causes basement membrane defects and perinatal lethality in mice. *Mol Cell Biol* 25: 6846-6856.
- Bailey AJ & Knott L (1999) Molecular changes in bone collagen in osteoporosis and osteoarthritis in the elderly. *Exp Gerontol* 34: 337-351.
- Bank RA, Robins SP, Wijmenga C, Breslau-Siderius LJ, Bardoel AF, van der Sluijs HA, Pruijs HE & TeKoppele JM (1999) Defective collagen crosslinking in bone, but not in ligament or cartilage, in Bruck syndrome: indications for a bone-specific telopeptide lysyl hydroxylase on chromosome 17. *Proc Natl Acad Sci U.S.A.* 96(3): 1054-1058.
- Batge B, Winter C, Notbohm H, Acil Y, Brinckmann J & Müller PK (1997) Glycosylation of human bone collagen I in relation to lysylhydroxylation and fibril diameter. *J Biochem (Tokyo)* 122: 109-115.
- 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.
- Blacksin MF, Pletcher BA & David M (1998) Osteogenesis imperfecta with joint contractures: Bruck syndrome. *Pediatr Radiol* 28: 117-119.
- Bonaldo P, Braghetta P, Zanetti M, Piccolo S, Volpin D, Bressan GM. (1998) Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for Bethlem myopathy. *Hum Mol Genet* 7: 2135-40.

- Bonfanti L, Mironov AA, Jr, Martinez-Menarguez JA, Martella O, Fusella A, Baldassarre M, Buccione R, Geuze HJ, Mironov AA & Luini A (1998) Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell* 95: 993-1003.
- Borchiellini C, Coulon J & Le Parco Y (1996) The function of type IV collagen during *Drosophila* muscle development. *Mech Dev* 58: 179-191.
- Bottomley MJ, Batten MR, Lumb RA & Bulleid NJ (2001) Quality control in the endoplasmic reticulum: PDI mediates the ER retention of unassembled procollagen C-propeptides. *Curr Biol* 11: 1114-1118.
- Breedveld G, de Coo IF, Lequin MH, Arts WF, Heutink P, Gould DB, John SW, Oostra B & Mancini GM (2006) Novel mutations in three families confirm a major role of *COL4A1* in hereditary proencephaly. *J Med Genet* 43: 490-495.
- Brinckmann J, Kim S, Wu J, Reinhardt DP, Batmunkh C, Metzen E, Notbohm H, Bank RA, Krieg T & Hunzelmann N (2005) Interleukin 4 and prolonged hypoxia induce a higher gene expression of lysyl hydroxylase 2 and an altered cross-link pattern: important pathogenetic steps in early and late stage of systemic scleroderma? *Matrix Biol* 24: 459-468.
- Brinckmann J, Neess CM, Gaber Y, Sobhi H, Notbohm H, Hunzelmann N, Fietzek PP, Müller PK, Risteli J, Gebker R & Scharffetter-Kochanek K (2001) Different pattern of collagen cross-links in two sclerotic skin diseases: lipodermatosclerosis and circumscribed scleroderma. *J Invest Dermatol* 117: 269-273.
- Brinckmann J, Notbohm H, Tronnier M, Acil Y, Fietzek PP, Schmeller W, Müller PK & Batge B (1999) Overhydroxylation of lysyl residues is the initial step for altered collagen cross-links and fibril architecture in fibrotic skin. *J Invest Dermatol* 113: 617-621.
- Byers BH & Cole WG (2002) Osteogenesis imperfecta. In: Royce PM & Steinmann B (eds) *Connective Tissue and Its Heritable Disorders. Molecular, Genetic, and Medical Aspects*. 2nd ed. Wiley-Liss Inc., New York, 385-430.
- Canty EG & Kadler KE (2005) Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci* 118: 1341-1353.
- Canty EG, Lu Y, Meadows RS, Shaw MK, Holmes DF & Kadler KE (2004) Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *J Cell Biol* 165: 553-563.
- Chakravarti S, Magnuson T, Lass JH, Jepsen KJ, LaMantia C & Carroll H (1998) Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J Cell Biol* 141: 1277-1286.
- Chartier A, Zaffran S, Astier M, Semeriva M & Gratecos D (2002) Pericardin, a *Drosophila* type IV collagen-like protein is involved in the morphogenesis and maintenance of the heart epithelium during dorsal ectoderm closure. *Development* 129: 3241-3253.
- Clifton IJ, McDonough MA, Ehrismann D, Kershaw NJ, Granatino N & Schofield CJ (2006) Structural studies on 2-oxoglutarate oxygenases and related double-stranded β -helix fold proteins. *J Inorg Biochem* 100: 644-669.
- Corsi A, Xu T, Chen XD, Boyde A, Liang J, Mankani M, Sommer B, Iozzo RV, Eichstetter I, Robey PG, Bianco P & Young MF (2002) Phenotypic effects of biglycan deficiency are linked to collagen fibril abnormalities, are synergized by decorin deficiency, and mimic Ehlers-Danlos-like changes in bone and other connective tissues. *J Bone Miner Res* 17: 1180-1189.
- Cosgrove D, Meehan DT, Grunkemeyer JA, Kornak JM, Sayers R, Hunter WJ & Samuelson GC (1996) Collagen *COL4A3* knockout: a mouse model for autosomal Alport syndrome. *Genes Dev* 10: 2981-2992.
- Costell M, Gustafsson E, Aszodi A, Morgelin M, Bloch W, Hunziker E, Addicks K, Timpl R & Fässler R (1999) Perlecan maintains the integrity of cartilage and some basement membranes. *J Cell Biol* 147: 1109-1122.

- Danielson KG, Baribault H, Holmes DF, Graham H, Kadler KE & Iozzo RV (1997) Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J Cell Biol* 136: 729-743.
- Engel J & Bächinger HP (2005) Structure, stability and folding of the collagen triple helix. In: Brinckmann J, Notbohm H & Müller PK (eds) *Topics in Current Chemistry, Collagen: Primer in Structure, Processing and Assembly*. Springer-Verlag, Berlin Heidelberg, 247: 7-33.
- Eyre DR, Paz MA & Gallop PM (1984) Cross-linking in collagen and elastin. *Annu Rev Biochem* 53: 717-748.
- Eyre D, Shao P, Weis MA & Steinmann B (2002) The kyphoscoliotic type of Ehlers-Danlos syndrome (type VI): differential effects on the hydroxylation of lysine in collagens I and II revealed by analysis of cross-linked telopeptides from urine. *Mol Genet Metab* 76: 211-216.
- Eyre DR & Wu JJ (2005) Collagen cross-links. In: Brinckmann J, Notbohm H & Müller PK (eds) *Topics In Current Chemistry, Collagen: Primer in Structure, Processing and Assembly*. Springer-Verlag, Berlin Heidelberg, 247: 207-229.
- Fitzgerald J, Lamandé SR, Bateman JF. (1999) Proteasomal degradation of unassembled mutant type I collagen pro- α 1(I) chains. *J Biol Chem* 274: 27392-8.
- Fukai N, Eklund L, Marneros AG, Oh SP, Keene DR, Tamarkin L, Niemelä M, Ilves M, Li E, Pihlajaniemi T & Olsen BR (2002) Lack of collagen XVIII/endostatin results in eye abnormalities. *EMBO J* 21: 1535-1544.
- Gelse K, Pöschl E & Aigner T (2003) Collagens - structure, function, and biosynthesis. *Adv Drug Deliv Rev* 55: 1531-1546.
- Gerriets JE, Curwin SL & Last JA (1993) Tendon hypertrophy is associated with increased hydroxylation of nonhelical lysine residues at two specific cross-linking sites in type I collagen. *J Biol Chem* 268: 25553-25560.
- Giunta C, Randolph A, Al-Gazali LI, Brunner HG, Kraenzlin ME & Steinmann B (2005a) Nevo syndrome is allelic to the kyphoscoliotic type of the Ehlers-Danlos syndrome (EDS VIA). *Am J Med Genet A* 133: 158-164.
- Giunta C, Randolph A & Steinmann B (2005b) Mutation analysis of the *PLOD1* gene: An efficient multistep approach to the molecular diagnosis of the kyphoscoliotic type of Ehlers-Danlos syndrome. *Mol Genet Metab* 86: 269-276.
- Gotkin MG, Ripley CR, Lamande SR, Bateman JF, Bienkowski RS. (2004) Intracellular trafficking and degradation of unassociated pro α 2 chains of collagen type I. *Exp Cell Res* 296:307-16.
- Gould DB, Phalan FC, Breedveld GJ, van Mil SE, Smith RS, Schimenti JC, Aguglia U, van der Knaap MS, Heutink P & John SW (2005) Mutations in *Col4a1* cause perinatal cerebral hemorrhage and porencephaly. *Science* 308: 1167-1171.
- Gould DB, Phalan FC, van Mil SE, Sundberg JP, Vahedi K, Massin P, Bousser MG, Heutink P, Miner JH, Tournier-Lasserre E, John SW. (2006) Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *N Engl J Med*. 354:1489-96.
- Gupta MC, Graham PL & Kramer JM (1997) Characterization of α 1(IV) collagen mutations in *Caenorhabditis elegans* and the effects of α 1 and α 2(IV) mutations on type IV collagen distribution. *J Cell Biol* 137: 1185-1196.
- Hanson DA & Eyre DR (1996) Molecular site specificity of pyridinoline and pyrrole cross-links in type I collagen of human bone. *J Biol Chem* 271: 26508-26516.
- Hautala T, Byers MG, Eddy RL, Shows TB, Kivirikko KI & Myllylä R (1992) Cloning of human lysyl hydroxylase: complete cDNA-derived amino acid sequence and assignment of the gene (*PLOD*) to chromosome 1p36.3-p36.2. *Genomics* 13: 62-69.
- Ha-Vinh R, Alanay Y, Bank RA, Campos-Xavier AB, Zankl A, Superti-Furga A & Bonafe L (2004) Phenotypic and molecular characterization of Bruck syndrome (osteogenesis imperfecta with contractures of the large joints) caused by a recessive mutation in *PLOD2*. *Am J Med Genet A* 131: 115-120.

- 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 type VI patients. *Genomics* 24: 464-471.
- Heikkinen J, Risteli M, Wang C, Latvala J, Rossi M, Valtavaara M & Myllylä R (2000) Lysyl hydroxylase 3 is a multifunctional protein possessing collagen glucosyltransferase activity. *J Biol Chem* 275: 36158-36163.
- 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 a repetitive sequence within the gene and is a common cause for the type VI variant of Ehlers-Danlos syndrome. *Am J Hum Genet* 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 $\alpha_2\beta_2$ tetramer with the protein disulfide-isomerase/ β subunit. *Proc Natl Acad Sci USA* 92: 4427-4431.
- 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.
- Hendershot LM & Bulleid NJ (2000) Protein-specific chaperones: the role of hsp47 begins to gel. *Curr Biol* 10: R912-5.
- Hieta R, Kukkola L, Permi P, Pirilä P, Kivirikko KI, Kilpeläinen I & Myllyharju J (2003) The peptide-substrate-binding domain of human collagen prolyl 4-hydroxylases. Backbone assignments, secondary structure, and binding of proline-rich peptides. *J Biol Chem* 278: 34966-34974.
- Hjalt TA, Amendt BA & Murray JC (2001) PITX2 regulates procollagen lysyl hydroxylase (PLOD) gene expression: implications for the pathology of Rieger syndrome. *J Cell Biol* 152: 545-552.
- Holster T, Pakkanen O, Soininen R, Sormunen R, Nokelainen M, Kivirikko KI & Myllyharju J (2007) Loss of assembly of the main basement-membrane collagen, type IV, but not fibril-forming collagens and embryonic death in collagen prolyl 4-hydroxylase I null mice. *J Biol Chem* 282: 2512-2519.
- Horn N & Tümer Z (2002) Menkes disease and the occipital horn syndrome. In: Royce PM & Steinmann B (eds) *Connective Tissue and Its Heritable Disorders. Molecular, Genetic, and Medical Aspects*. 2nd ed. Wiley-Liss Inc., New York, 651-685.
- Hornstra IK, Birge S, Starcher B, Bailey AJ, Mecham RP & Shapiro SD (2003) Lysyl oxidase is required for vascular and diaphragmatic development in mice. *J Biol Chem* 278: 14387-14393.
- Hubbard SJ, Beynon RJ, Thornton JM. (1998) Assessment of conformational parameters as predictors of limited proteolytic sites in native protein structures. *Protein Eng* 11:349-59.
- Hudson BG, Tryggvason K, Sundaramoorthy M & Neilson EG (2003) Alport's syndrome, Goodpasture's syndrome, and type IV collagen. *N Engl J Med* 348: 2543-2556.
- Hulmes DJ (2002) Building collagen molecules, fibrils, and suprafibrillar structures. *J Struct Biol* 137: 2-10.
- Hurskainen M, Eklund L, Hägg PO, Fruttiger M, Sormunen R, Ilves M, Pihlajaniemi T. (2005) Abnormal maturation of the retinal vasculature in type XVIII collagen/endostatin deficient mice and changes in retinal glial cells due to lack of collagen types XV and XVIII. *FASEB J* 19: 1564-6
- Ihme A, Krieg T, Nerlich A, Feldmann U, Rauterberg J, Glanville RW, Edel G & Müller PK (1984) Ehlers-Danlos syndrome type VI: Collagen type specificity of defective lysyl hydroxylation in various tissues. *J Invest Dermatol* 83: 161-165.

- Irwin WA, Bergamin N, Sabatelli P, Reggiani C, Megighian A, Merlini L, Braghetta P, Columbaro M, Volpin D, Bressan GM, Bernardi P, Bonaldo P. (2003) Mitochondrial dysfunction and apoptosis in myopathic mice with collagen VI deficiency. *Nat Genet* 35: 367-71.
- Jenkins CL, Bretscher LE, Guzei IA & Raines RT (2003) Effect of 3-hydroxyproline residues on collagen stability. *J Am Chem Soc* 125: 6422-6427.
- Jimenez S, Harsch M & Rosenbloom J (1973) Hydroxyproline stabilizes the triple helix of chick tendon collagen. *Biochem Biophys Res Commun* 52: 106-114.
- Kagan HM (1986) Characterisation and regulation of lysyl oxidase. In: Mecham RP (ed) *Biology of Extracellular Matrix*. 1. Academic Press, Orlando, 321-389.
- Kaler SG (1998) Metabolic and molecular bases of Menkes disease and occipital horn syndrome. *Pediatr Dev Pathol* 1: 85-98.
- Kalluri R (2003) Basement membranes: structure, assembly and role in tumour angiogenesis. *Nat Rev Cancer* 3: 422-433.
- Kellokumpu S, Sormunen R, Heikkinen J & Myllylä R (1994) Lysyl hydroxylase, a collagen processing enzyme, exemplifies a novel class of luminally-oriented peripheral membrane proteins in the endoplasmic reticulum. *J Biol Chem* 269: 30524-30529.
- Khoshnoodi J, Sigmundsson K, Cartiailler JP, Bondar O, Sundaramoorthy M & Hudson BG (2006) Mechanism of chain selection in the assembly of collagen IV: A prominent role for the $\alpha 2$ chain. *J Biol Chem* 281: 6058-6069.
- Kielty CM & Grant ME (2002) The collagen family: structure, assembly, and organization in the extracellular matrix. In: Royce PM & Steinmann B (eds) *Connective Tissue and Its Heritable Disorders. Molecular, Genetic, and Medical Aspects*. 2nd ed. Wiley-Liss Inc., New York, 159-221.
- Kirsch E, Krieg T, Remberger K, Fendel H, Bruckner P & Müller PK (1981) Disorder of collagen metabolism in a patient with osteogenesis imperfecta (lethal type): increased degree of hydroxylation of lysine in collagen types I and III. *Eur J Clin Invest* 11: 39-47.
- Kivirikko KI (1995) Posttranslational processing of collagens. In: Bittar EE, Bittar N & Greenwich CT (eds) *Principles of Medical Biology*. JAI Press, London, 233-254.
- Kivirikko KI & Myllylä R (1979) Collagen glycosyltransferases. *Int.Rev.Connect.Tissue Res.* 8: 23-72.
- Kivirikko KI & Myllylä R (1980) The hydroxylation of prolyl and lysyl residues. In: Friedman DB & Hawkins H (eds) *The Enzymology of Post-Translational Modification of Proteins*. Academic Press, New York, 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 JJ & Crabbe JC (eds) *Post-Translational Modifications of Proteins*, CRC Press, Boca Raton, 1-51.
- 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) Enzymatic hydroxylation of proline and lysine in protocollagen. *Proc Natl Acad Sci USA* 57: 782-789.
- Knott L & Bailey AJ (1998) Collagen cross-links in mineralizing tissues: A review of their chemistry, function and clinical relevance. *Bone* 22: 181-187.
- Knox SM & Whitelock JM. (2006) Perlecan: how does one molecule do so many things? *Cell Mol Life Sci* 63:2435-45.
- Koide T & Nagata K (2005) Collagen Biosynthesis. In: Brinckmann J, Notbohm H & Müller PK (eds) *Topics In Current Chemistry, Collagen: Primer in Structure, Processing and Assembly*. Springer-Verlag, Berlin Heidelberg, 247: 85-114.

- 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.
- Krol BJ, Murad S, Walker LC, Marshall MK, Clark WL, Pinnell SR, Yeowell HN. (1996) The expression of a functional, secreted human lysyl hydroxylase in a baculovirus system. *J Invest Dermatol* 106:11-6.
- Kuhn K (1995) Basement membrane (type IV) collagen. *Matrix Biol* 14: 439-445.
- Kukkola L, Hieta R, Kivirikko KI & Myllyharju J (2003) Identification and characterization of a third human, rat, and mouse collagen prolyl 4-hydroxylase isoenzyme. *J Biol Chem* 278: 47685-47693.
- Kvist AJ, Johnson AE, Mörgelin M, Gustafsson E, Bengtsson E, Lindblom K, Aszodi A, Fässler R, Sasaki T, Timpl R & Aspberg A (2006) Chondroitin sulfate perlecan enhances collagen fibril formation. Implications for perlecan chondrodysplasias. *J Biol Chem* 281: 33127-33139.
- 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.
- Lamandé SR, Chessler SD, Golub SB, Byers PH, Chan D, Cole WG, Silience DO, Bateman JF. (1995) Endoplasmic reticulum-mediated quality control of type I collagen production by cells from osteogenesis imperfecta patients with mutations in the pro α 1(I) chain carboxyl-terminal propeptide which impair subunit assembly. *J Biol Chem* 270: 8642-9.
- Lees JF, Tasab M & Bulleid NJ (1997) Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. *EMBO J* 16: 908-916.
- Leroy JG, Nuytincq L, De Paepe A, De Rammelaere M, Gillerot Y, Verloes A, Loeys B & De Groote W (1998) Bruck syndrome: neonatal presentation and natural course in three patients. *Pediatr Radiol* 28: 781-789.
- Liu X, Wu H, Byrne M, Krane S & Jaenisch R (1997) Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. *Proc Natl Acad Sci USA* 94: 1852-1856.
- Liu X, Zhao Y, Gao J, Pawlyk B, Starcher B, Spencer JA, Yanagisawa H, Zuo J & Li T (2004) Elastic fiber homeostasis requires lysyl oxidase-like 1 protein. *Nat Genet* 36: 178-182.
- Lucero HA & Kagan HM (2006) Lysyl oxidase: an oxidative enzyme and effector of cell function. *Cell Mol Life Sci* 63: 2304-2316.
- Marini JC, Forlino A, Cabral WA, Barnes AM, San Antonio JD, Milgrom S, Hyland JC, Körkkö J, Prockop DJ, De Paepe A, Coucke P, Symoens S, Glorieux FH, Roughley PJ, Lund AM, Kuurila-Svahn K, Hartikka H, Cohn DH, Krakow D, Mottes M, Schwarze U, Chen D, Yang K, Kuslich C, Troendle J, Dagleish R & Byers PH (2006) Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Hum Mutat* 28: 208-221.
- Marneros AG & Olsen BR (2003) Age-dependent iris abnormalities in collagen XVIII/endostatin deficient mice with similarities to human pigment dispersion syndrome. *Invest Ophthalmol Vis Sci* 44: 2367-2372.
- McAlinden A, Smith TA, Sandell LJ, Ficheux D, Parry DA & Hulmes DJ (2003) α -helical coiled-coil oligomerization domains are almost ubiquitous in the collagen superfamily. *J Biol Chem* 278: 42200-42207.
- McLaughlin SH & Bulleid NJ (1998) Thiol-independent interaction of protein disulphide isomerase with type X collagen during intra-cellular folding and assembly. *Biochem J* 331: 793-800.
- McPherson E & Clemens M (1997) Bruck syndrome (osteogenesis imperfecta with congenital joint contractures): review and report on the first North American case. *Am J Med Genet* 70: 28-31.
- Mercer DK, Nicol PF, Kimbembe C & Robins SP (2003) Identification, expression, and tissue distribution of the three rat lysyl hydroxylase isoforms. *Biochem Biophys Res Commun* 307: 803-809.

- Miller JE (1984) Chemistry of collagens and their distribution. In: Piez KA & Reddi AH (eds) *Extracellular Matrix Biochemistry*. Elsevier, New York, 41-81.
- Miner JH & Sanes JR (1996) Molecular and functional defects in kidneys of mice lacking collagen $\alpha 3(\text{IV})$: implications for Alport syndrome. *J Cell Biol* 135: 1403-1413.
- Miner JH & Yurchenco PD (2004) Laminin functions in tissue morphogenesis. *Annu Rev Cell Dev Biol* 20: 255-284.
- Miosge N, Sasaki T & Timpl R (2002) Evidence of nidogen-2 compensation for nidogen-1 deficiency in transgenic mice. *Matrix Biol* 21: 611-621.
- Mironov AA, Mironov AA Jr, Beznoussenko GV, Trucco A, Lupetti P, Smith JD, Geerts WJ, Koster AJ, Burger KN, Martone ME, Deerinck TJ, Ellisman MH & Luini A (2003) ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains. *Dev Cell* 5: 583-594.
- Mizuno K, Hayashi T, Peyton DH & Bächinger HP (2004) The peptides acetyl-(Gly-3(S)Hyp-4(R)Hyp)₁₀-NH₂ and Acetyl-(Gly-Pro-3(S)Hyp)₁₀-NH₂ do not form a collagen triple helix. *J Biol Chem* 279: 282-287.
- Morello R, Bertin TK, Chen Y, Hicks J, Tonachini L, Monticone M, Castagnola P, Rauch F, Glorieux FH, Vranka J, Bächinger HP, Pace JM, Schwarze U, Byers PH, Weis M, Fernandes RJ, Eyre DR, Yao Z, Boyce BF & Lee B (2006) CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell* 127: 291-304.
- Morita H, Yoshimura A, Inui K, Ideura T, Watanabe H, Wang L, Soinen R, Tryggvason K. (2005) Heparan sulfate of perlecan is involved in glomerular filtration. *J Am Soc Nephrol* 16:1703-10.
- Murshed M, Smyth N, Miosge N, Karolat J, Krieg T, Paulsson M & Nischt R (2000) The absence of nidogen 1 does not affect murine basement membrane formation. *Mol Cell Biol* 20: 7007-7012.
- Myllyharju J (2003) Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol* 22: 15-24.
- Myllyharju J (2005) Intracellular post-translational modifications of collagens. In: Brinckmann J, Notbohm H & Müller PK (eds) *Topics In Current Chemistry, Collagen: Primer in Structure, Processing and Assembly*. Springer-Verlag, Berlin Heidelberg, 247: 115-147.
- 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.
- Myllyharju J & Kivirikko KI, (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* 20: 33-43.
- Myllyharju J, Lamberg A, Notbohm H, Fietzek PP, Pihlajaniemi T & Kivirikko KI (1997) Expression of wild-type and modified pro α chains of human type I procollagen in insect cells leads to the formation of stable [$\alpha 1(\text{I})$]₂ $\alpha 2(\text{I})$ collagen heterotrimers and [$\alpha 1(\text{I})$]₃ homotrimers but not [$\alpha 2(\text{I})$]₃ homotrimers. *J Biol Chem* 272: 21824-21830.
- Myllyharju J, Nokelainen M, Vuorela A & Kivirikko KI (2000) Expression of recombinant human type I-III collagens in the yeast *pichia pastoris*. *Biochem Soc Trans* 28: 353-357.
- Myllylä R, Anttinen H, Risteli L & Kivirikko KI (1977) Isolation of collagen glucosyltransferase as a homogeneous protein from chick embryos. *Biochim Biophys Acta* 480: 113-121.
- 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, Pihlajaniemi T, Pajunen L, Turpeenniemi-Hujanen T & Kivirikko KI (1991) Molecular cloning of chick lysyl hydroxylase. Little homology in primary structure to the two types of subunit of prolyl 4-hydroxylase. *J Biol Chem* 266: 2805-2810.
- 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.
- Mäki JM, Räsänen J, Tikkanen H, Sormunen R, Mäkikallio K, Kivirikko KI & Soininen R (2002) Inactivation of the lysyl oxidase gene *Lox* leads to aortic aneurysms, cardiovascular dysfunction, and perinatal death in mice. *Circulation* 106: 2503-2509.
- Mäki JM, Sormunen R, Lippo S, Kaarteenaho-Wiik R, Soininen R & Myllyharju J (2005) Lysyl oxidase is essential for normal development and function of the respiratory system and for the integrity of elastic and collagen fibers in various tissues. *Am J Pathol* 167: 927-936.
- Nagata K (2003) HSP47 as a collagen-specific molecular chaperone: function and expression in normal mouse development. *Semin Cell Dev Biol* 14: 275-282.
- Nokelainen M, Helaakoski T, Myllyharju J, Notbohm H, Pihlajaniemi T, Fietzek PP & Kivirikko KI (1998) Expression and characterization of recombinant human type II collagens with low and high contents of hydroxylysine and its glycosylated forms. *Matrix Biol* 16: 329-338.
- Nokelainen M, Tu H, Vuorela A, Notbohm H, Kivirikko KI & Myllyharju J (2001) High-level production of human type I collagen in the yeast *Pichia pastoris*. *Yeast* 18: 797-806.
- 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.
- Notbohm H, Nokelainen M, Myllyharju J, Fietzek PP, Müller PK & Kivirikko KI (1999) Recombinant human type II collagens with low and high levels of hydroxylysine and its glycosylated forms show marked differences in fibrillogenesis *in vitro*. *J Biol Chem* 274: 8988-8992.
- Online Mendelian Inheritance in Man, OMIM (TM). Johns Hopkins University, Baltimore, MD. MIM Number: 120150; last edited: 10/2/2006. World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>.
- Online Mendelian Inheritance in Man, OMIM (TM). Johns Hopkins University, Baltimore, MD. MIM Number: 120160; last edited: 9/20/2006. World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>.
- Online Mendelian Inheritance in Man, OMIM (TM). Johns Hopkins University, Baltimore, MD. MIM Number: 225400; last edited: 6/15/2005. World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>.
- Online Mendelian Inheritance in Man, OMIM (TM). Johns Hopkins University, Baltimore, MD. MIM Number: 229200; last edited: 3/29/2006. World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>.
- Online Mendelian Inheritance in Man, OMIM (TM). Johns Hopkins University, Baltimore, MD. MIM Number: 259450; last edited: 2/25/2005. World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>.
- Pasquali M, Dembure PP, Still MJ & Elsas LJ (1994) Urinary pyridinium cross-links: a noninvasive diagnostic test for Ehlers-Danlos syndrome type VI. *N Engl J Med* 331: 132-133.
- Pasquali N, Still MJ, Dembure PP & Elsas LJ (1995) Pyridinium cross-links in heritable disorders of collagen. *Am J Hum Genet* 57: 1508-1510.
- Passoja K, Myllyharju J, Pirskanen A & Kivirikko KI (1998a) Identification of arginine-700 as the residue that binds the C-5 carboxyl group of 2-oxoglutarate in human lysyl hydroxylase 1. *FEBS Lett* 434: 145-148.

- Passoja K, Rautavuoma K, Ala-Kokko L, Kosonen T & Kivirikko KI (1998b) Cloning and characterization of a third human lysyl hydroxylase isoform. *Proc Natl Acad Sci USA* 95: 10482-10486.
- Pekkala M, Hieta R, Bergmann U, Kivirikko KI, Wierenga RK & Myllyharju J (2004) The peptide-substrate-binding domain of collagen prolyl 4-hydroxylases is a tetratricopeptide repeat domain with functional aromatic residues. *J Biol Chem* 279: 52255-52261.
- Pihlajaniemi T, Helaakoski T, Tasanen K, Myllylä R, Huhtala M-L, Koivu J & Kivirikko KI (1987) Molecular cloning of the β -subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. *EMBO J* 6: 643-649.
- Pirkanen A, Kaimio A-M, Myllylä R & Kivirikko KI (1996) Site-directed mutagenesis of human lysyl hydroxylase expressed in insect cells. Identification of histidine residues and an aspartic acid residue critical for catalytic activity. *J Biol Chem* 271: 9398-9402.
- Pornprasertsuk S, Duarte WR, Mochida Y & Yamauchi M (2004) Lysyl hydroxylase-2b directs collagen cross-linking pathways in MC3T3-E1 cells. *J Bone Miner Res* 19: 1349-1355.
- Pousi B, Heikkinen J, Schröter J, Pope M & Myllylä R (2000) A nonsense codon of exon 14 reduces lysyl hydroxylase mRNA and leads to aberrant RNA splicing in a patient with Ehlers-Danlos syndrome type VI. *Mutat Res* 432: 33-37.
- Pöschl E, Schlötzer-Schrehardt U, Brachvogel B, Saito K, Ninomiya Y & Mayer U (2004) Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development* 131: 1619-1628.
- Ricard-Blum S, Ruggiero F & van der Rest M (2005) The collagen superfamily. In: Brinckmann J, Notbohm H & Müller PK (eds) *Topics In Current Chemistry, Collagen: Primer in Structure, Processing and Assembly*. Springer-Verlag, Berlin Heidelberg, 247: 35-84.
- Risteli L, Myllylä R & Kivirikko KI (1976) Partial purification and characterization of collagen galactosyltransferase from chick embryos. *Biochem J* 155: 145-153.
- Risteli M, Niemitalo O, Lankinen H, Juffer AH & Myllylä R (2004) Characterization of collagenous peptides bound to lysyl hydroxylase isoforms. *J Biol Chem* 279: 37535-37543.
- Rossi M, Morita H, Sormunen R, Airene S, Kreivi M, Wang L, Fukai N, Olsen BR, Tryggvason K, Soininen R. (2003) Heparan sulfate chains of perlecan are indispensable in the lens capsule but not in the kidney. *EMBO J* 22: 236-45.
- Royce PM & Barnes MJ (1985) Failure of highly purified lysyl hydroxylase to hydroxylate lysyl residues in the non-helical regions of collagen. *Biochem J* 230: 475-480.
- Ruotsalainen H, Sipilä L, Kerkelä E, Pospiech H & Myllylä R (1999) Characterization of cDNAs for mouse lysyl hydroxylase 1, 2 and 3, their phylogenetic analysis and tissue-specific expression in the mouse. *Matrix Biol* 18: 325-329.
- Ruotsalainen H, Sipilä L, Vapola M, Sormunen R, Salo AM, Uitto L, Mercer DK, Robins SP, Risteli M, Aszodi A, Fässler R & Myllylä R (2006) Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J Cell Sci* 119: 625-635.
- Salo AM, Sipilä L, Sormunen R, Ruotsalainen H, Vainio S & Myllylä R (2006b) The lysyl hydroxylase isoforms are widely expressed during mouse embryogenesis, but obtain tissue- and cell-specific patterns in the adult. *Matrix Biol* 25: 475-483.
- Salo AM, Wang C, Sipilä L, Sormunen R, Vapola M, Kervinen P, Ruotsalainen H, Heikkinen J & Myllylä R (2006a) Lysyl hydroxylase 3 (LH3) modifies proteins in the extracellular space, a novel mechanism for matrix remodeling. *J Cell Physiol* 207: 644-653.
- Sasaki T, Fässler R & Hohenester E (2004) Laminin: the crux of basement membrane assembly. *J Cell Biol* 164: 959-963.
- Schneider VA & Granato M (2006) The myotomal diwanka (lh3) glycosyltransferase and type XVIII collagen are critical for motor growth cone migration. *Neuron* 50: 683-695.
- Schneider VA & Granato M (2007) Genomic structure and embryonic expression of zebrafish lysyl hydroxylase 1 and lysyl hydroxylase 2. *Matrix Biol* 26: 12-19.

- Schymeinsky J, Nedbal S, Miosge N, Pöschl E, Rao C, Beier DR, Skarnes WC, Timpl R & Bader BL (2002) Gene structure and functional analysis of the mouse nidogen-2 gene: nidogen-2 is not essential for basement membrane formation in mice. *Mol Cell Biol* 22: 6820-6830.
- Sipilä L, Szatanik M, Vainionpää H, Ruotsalainen H, Myllylä R & Guenet JL (2000) The genes encoding mouse lysyl hydroxylase isoforms map to chromosomes 4, 5, and 9. *Mamm Genome* 11: 1132-1134.
- Smith-Mungo LI & Kagan HM (1998) Lysyl oxidase: properties, regulation and multiple functions in biology. *Matrix Biol* 16: 387-398.
- Spiro RG, Lucas F & Rudall KM (1971) Glycosylation of hydroxylysine in collagens. *Nat New Biol* 231: 54-55.
- Spiro MJ & Spiro RG (1971) Studies on the biosynthesis of the hydroxylysine-linked disaccharide unit of basement membranes and collagens. II. Kidney galactosyltransferase. *J Biol Chem* 246: 4910-4918.
- Steinmann B, Eyre DR & Shao P (1995) Urinary pyridinoline cross-links in Ehlers-Danlos syndrome type VI. *Am J Hum Genet* 57: 1505-1508.
- Steinmann B, Royce PM & Superti-Furga A (2002) The Ehlers-Danlos syndrome. In: Royce PM & Steinmann B (eds) *Connective Tissue and Its Heritable Disorders. Molecular, Genetic, and Medical Aspects*. 2nd ed. Wiley-Liss, Inc., New York, 431-523.
- Suokas M, Lampela O, Juffer AH, Myllylä R & Kellokumpu S (2003) Retrieval-independent localization of lysyl hydroxylase in the endoplasmic reticulum via a peptide fold in its iron-binding domain. *Biochem J* 370: 913-920.
- Suokas M, Myllylä R & Kellokumpu S (2000) A single C-terminal peptide segment mediates both membrane association and localization of lysyl hydroxylase in the endoplasmic reticulum. *J Biol Chem* 275: 17863-17868.
- Svensson L, Aszodi A, Reinholt FP, Fässler R, Heinegård D & Oldberg A (1999) Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered lumican deposition in tendon. *J Biol Chem* 274: 9636-9647.
- Sziper C, Sziper J, Riviere M, Vanvooren P, Valtavaara M & Myllylä R (1997) Localization of the gene encoding a novel isoform of lysyl hydroxylase. *Mamm Genome* 8: 707-708.
- Tasab M, Batten MR & Bulleid NJ (2000) Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen. *EMBO J* 19: 2204-2211.
- Than ME, Bourenkov GP, Henrich S, Mann K & Bode W (2005) The NC1 dimer of human placental basement membrane collagen IV: does a covalent crosslink exist? *Biol Chem* 386: 759-766.
- Than ME, Henrich S, Huber R, Ries A, Mann K, Kühn K, Timpl R, Bourenkov GP, Bartunik HD & Bode W (2002) The 1.9-Å crystal structure of the noncollagenous (NC1) domain of human placenta collagen IV shows stabilization via a novel type of covalent Met-Lys cross-link. *Proc Natl Acad Sci USA* 99: 6607-6612.
- Tran PK, Tran-Lundmark K, Soininen R, Tryggvason K, Thyberg J, Hedin U. (2004) Increased intimal hyperplasia and smooth muscle cell proliferation in transgenic mice with heparan sulfate-deficient perlecan. *Circ Res* 94: 550-8.
- Trojanowska M, LeRoy EC, Eckes B & Krieg T (1998) Pathogenesis of fibrosis: type 1 collagen and the skin. *J Mol Med* 76: 266-274.
- Tryggvason K (1996) Mutations in type IV collagen genes and Alport phenotypes. *Contrib Nephrol* 117: 154-171.
- Tryggvason K, Majamaa K, Risteli J & Kivirikko KI (1979) Partial purification and characterization of chick-embryo prolyl 3-hydroxylase. *Biochem J* 183: 303-307.
- Turpeenniemi TM, Puistola U, Anttinen H, Kivirikko KI. (1977) Affinity chromatography of lysyl hydroxylase on concanavalin A-agarose. *Biochim Biophys Acta* 483: 215-9.

- Turpeenniemi-Hujanen TM, Puistola U & Kivirikko KI (1980) Isolation of lysyl hydroxylase, an enzyme of collagen synthesis, from chick embryos as a homogeneous 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.
- Utriainen A, Sormunen R, Kettunen M, Carvalhaes LS, Sajanti E, Eklund L, Kauppinen R, Kitten GT & Pihlajaniemi T (2004) Structurally altered basement membranes and hydrocephalus in a type XVIII collagen deficient mouse line. *Hum Mol Genet* 13: 2089-2099.
- Uzawa K, Grzesik WJ, Nishiura T, Kuznetsov SA, Robey PG, Brenner DA & Yamauchi M (1999) Differential expression of human lysyl hydroxylase genes, lysine hydroxylation, and cross-linking of type I collagen during osteoblastic differentiation *in vitro*. *J Bone Miner Res* 14: 1272-1280.
- Uzawa K, Marshall MK, Katz EP, Tanzawa H, Yeowell HN & Yamauchi M (1998) Altered posttranslational modifications of collagen in keloid. *Biochem Biophys Res Commun* 249: 652-655.
- Uzawa K, Yeowell HN, Yamamoto K, Mochida Y, Tanzawa H & Yamauchi M (2003) Lysine hydroxylation of collagen in a fibroblast cell culture system. *Biochem Biophys Res Commun* 305: 484-487.
- Valtavaara M (1999) Novel lysyl hydroxylase isoforms. *Acta Univ Oulu A* 334.
- Valtavaara M, Papponen H, Pirttilä AM, Hiltunen K, Helander H & Myllylä R (1997) Cloning and characterization of a novel human lysyl hydroxylase isoform highly expressed in pancreas and muscle. *J Biol Chem* 272: 6831-6834.
- Valtavaara M, Szpirer C, Szpirer J & Myllylä R (1998) Primary structure, tissue distribution, and chromosomal localization of a novel isoform of lysyl hydroxylase (lysyl hydroxylase 3). *J Biol Chem* 273: 12881-12886.
- Van Agtmael T, Schlötzer-Schrehardt U, McKie L, Brownstein DG, Lee AW, Cross SH, Sado Y, Mullins JJ, Pöschl E & Jackson IJ (2005) Dominant mutations of *Col4a1* result in basement membrane defects which lead to anterior segment dysgenesis and glomerulopathy. *Hum Mol Genet* 14: 3161-3168.
- Van Den Diepstraten C, Papay K, Bolender Z, Brown A & Pickering JG (2003) Cloning of a novel prolyl 4-hydroxylase subunit expressed in the fibrous cap of human atherosclerotic plaque. *Circulation* 108: 508-511.
- van der Slot AJ, van Dura EA, de Wit EC, De Groot J, Huizinga TW, Bank RA & Zuurmond AM (2005) Elevated formation of pyridinoline cross-links by profibrotic cytokines is associated with enhanced lysyl hydroxylase 2b levels. *Biochim Biophys Acta* 1741: 95-102.
- van der Slot AJ, Zuurmond AM, Bardoel AF, Wijmenga C, Pruijs HE, Sillence DO, Brinckmann J, Abraham DJ, Black CM, Verzijl N, DeGroot J, Hanemaaijer R, TeKoppele JM, Huizinga TW & Bank RA (2003) Identification of *PLOD2* as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J Biol Chem* 278: 40967-40972.
- van der Slot AJ, Zuurmond AM, van den Bogaerd AJ, Ulrich MM, Middelkoop E, Boers W, Karel Runday H, DeGroot J, Huizinga TW & Bank RA (2004) Increased formation of pyridinoline cross-links due to higher telopeptide lysyl hydroxylase levels is a general fibrotic phenomenon. *Matrix Biol* 23: 251-257.
- van der Slot-Verhoeven AJ, van Dura EA, Attema J, Blauw B, Degroot J, Huizinga TW, Zuurmond AM & Bank RA (2005) The type of collagen cross-link determines the reversibility of experimental skin fibrosis. *Biochim Biophys Acta* 1740: 60-67.

- Vanacore RM, Friedman DB, Ham AJ, Sundaramoorthy M & Hudson BG (2005) Identification of S-hydroxylysyl-methionine as the covalent cross-link of the noncollagenous (NC1) hexamer of the $\alpha 1(\alpha 1)\alpha 2$ collagen IV network: a role for the post-translational modification of lysine 211 to hydroxylysine 211 in hexamer assembly. *J Biol Chem* 280: 29300-29310.
- Veit G, Kobbe B, Keene DR, Paulsson M, Koch M & Wagener R (2006) Collagen XXVIII, a novel VWA-domain-containing protein with many imperfections in the collagenous domain. *J Biol Chem* 281: 3494-3504.
- Vranka JA, Sakai LY & Bächinger HP (2004) Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. *J Biol Chem* 279: 23615-23621.
- Walker LC, Marini JC, Grange DK, Filie J & Yeowell HN (1999) A patient with Ehlers-Danlos syndrome type VI is homozygous for a premature termination codon in exon 14 of the lysyl hydroxylase 1 gene. *Mol Genet Metab* 67: 74-82.
- Walker LC, Overstreet MA, Siddiqui A, De Paepe A, Ceylaner G, Malfait F, Symoens S, Atsawasuwan P, Yamauchi M, Ceylaner S, Bank RA & Yeowell HN (2005b) A novel mutation in the lysyl hydroxylase 1 gene causes decreased lysyl hydroxylase activity in an Ehlers-Danlos VIA patient. *J Invest Dermatol* 124: 914-918.
- Walker LC, Overstreet MA, Willing MC, Marini JC, Cabral WA, Pals G, Bristow J, Atsawasuwan P, Yamauchi M & Yeowell HN (2004) Heterogeneous basis of the type VIB form of Ehlers-Danlos syndrome (EDS VIB) that is unrelated to decreased collagen lysyl hydroxylation. *Am J Med Genet A* 131: 155-162.
- Walker LC, Overstreet MA & Yeowell HN (2005b) Tissue-specific expression and regulation of the alternatively-spliced forms of lysyl hydroxylase 2 (LH2) in human kidney cells and skin fibroblasts. *Matrix Biol* 23: 515-523.
- 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, Luosujärvi H, Heikkinen J, Risteli M, Uitto L & Myllylä R (2002a) The third activity for lysyl hydroxylase 3: galactosylation of hydroxylysyl residues in collagens *in vitro*. *Matrix Biol* 21: 559-566.
- Wang C, Risteli M, Heikkinen J, Hussa AK, Uitto L & Myllylä R (2002b) Identification of amino acids important for the catalytic activity of the collagen glucosyltransferase associated with the multifunctional lysyl hydroxylase 3 (LH3). *J Biol Chem* 277: 18568-18573.
- Wang C, Valtavaara M & Myllylä R (2000) Lack of collagen type specificity for lysyl hydroxylase isoforms. *DNA Cell Biol* 19: 71-77.
- Wenstrup RJ, Florer JB, Davidson JM, Phillips CL, Pfeiffer BJ, Menezes DW, Chervoneva I & Birk DE (2006) Murine model of the Ehlers-Danlos syndrome. *col5a1* haploinsufficiency disrupts collagen fibril assembly at multiple stages. *J Biol Chem* 281: 12888-12895.
- Wilson R, Freddi S, Chan D, Cheah KS, Bateman JF. (2005) Misfolding of collagen X chains harboring Schmid metaphyseal chondrodysplasia mutations results in aberrant disulfide bond formation, intracellular retention, and activation of the unfolded protein response. *J Biol Chem* 280:15544-52.
- Wilson R, Lees JF & Bulleid NJ (1998) Protein disulfide isomerase acts as a molecular chaperone during the assembly of procollagen. *J Biol Chem* 273: 9637-9643.
- Wu J, Reinhardt DP, Batmunkh C, Lindenmaier W, Far RK, Notbohm H, Hunzelmann N & Brinckmann J (2006) Functional diversity of lysyl hydroxylase 2 in collagen synthesis of human dermal fibroblasts. *Exp Cell Res* 312: 3485-3494.
- Yeowell HN, Allen JD, Walker LC, Overstreet MA, Murad S & Thai SF (2000a) Deletion of cysteine 369 in lysyl hydroxylase 1 eliminates enzyme activity and causes Ehlers-Danlos syndrome type VI. *Matrix Biol* 19: 37-46.

- Yeowell HN, Ha V, Clark WL, Marshall MK & Pinnell SR (1994) Sequence analysis of a cDNA for lysyl hydroxylase isolated from human skin fibroblasts from a normal donor: differences from human placental lysyl hydroxylase cDNA. *J Invest Dermatol* 102: 382-384.
- Yeowell HN & Walker LC (1997) Ehlers-Danlos syndrome type VI results from a nonsense mutation and a splice site-mediated exon-skipping mutation in the lysyl hydroxylase gene. *Proc Assoc Am Physicians* 109: 383-396.
- Yeowell HN & Walker LC (1999) Tissue specificity of a new splice form of the human lysyl hydroxylase 2 gene. *Matrix Biol* 18: 179-187.
- Yeowell HN & Walker LC (2000) Mutations in the lysyl hydroxylase 1 gene that result in enzyme deficiency and the clinical phenotype of Ehlers-Danlos syndrome type VI. *Mol Genet Metab* 71: 212-224.
- Yeowell HN, Walker LC, Farmer B, Heikkinen J & Myllylä R (2000b) Mutational analysis of the lysyl hydroxylase 1 gene (*PLOD*) in six unrelated patients with Ehlers-Danlos syndrome type VI: prenatal exclusion of this disorder in one family. *Hum Mutat* 16: 90.
- Ylikärppä R, Eklund L, Sormunen R, Kontiola AI, Utriainen A, Määttä M, Fukai N, Olsen BR & Pihlajaniemi T (2003a) Lack of type XVIII collagen results in anterior ocular defects. *FASEB J* 17: 2257-2259.
- Ylikärppä R, Eklund L, Sormunen R, Muona A, Fukai N, Olsen BR & Pihlajaniemi T (2003b) Double knockout mice reveal a lack of major functional compensation between collagens XV and XVIII. *Matrix Biol* 22: 443-448.
- Yurchenco PD, Amenta PS & Patton BL (2004) Basement membrane assembly, stability and activities observed through a developmental lens. *Matrix Biol* 22: 521-538.
- Zhou Z, Wang J, Cao R, Morita H, Soininen R, Chan KM, Liu B, Cao Y, Tryggvason K. (2004) Impaired angiogenesis, delayed wound healing and retarded tumor growth in perlecan heparan sulfate-deficient mice. *Cancer Res* 64: 4699-702.

Original articles

- I Rautavuoma K, Takaluoma K, Passoja K, Pirskanen A, Kvist A-P, Kivirikko KI & Myllyharju J (2002) Characterization of three fragments that constitute the monomers of the human lysyl hydroxylase isoenzymes 1-3. *J Biol Chem* 277: 23084-23091.
- II Takaluoma K, Lantto J & Myllyharju J (2007) Lysyl hydroxylase 2 is a specific telopeptide hydroxylase, while all three isoenzymes hydroxylate collagenous sequences. Accepted for publication in *Matrix Biology*
- III Rautavuoma K, Takaluoma K, Sormunen R, Myllyharju J, Kivirikko KI & Soininen R (2004) Premature aggregation of type IV collagen and early lethality in lysyl hydroxylase 3 null mice. *Proc Natl Acad Sci USA* 101: 14120-14125.
- IV Takaluoma K, Hyry M, Lantto J, Sormunen R, Bank, RA, Kivirikko KI, Myllyharju J & Soininen R (2007) Tissue-specific changes in the hydroxylysine content and crosslinks of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice. *J Biol Chem* 282: 6588-6596.

The permission of the following copyright owners to reproduce the original articles has been granted.

(I and IV) Reprinted with permission of American Society for Biochemistry and Molecular Biology.

(II) Reprinted with permission from Elsevier.

(III) Copyright (2004) The National Academy of Sciences of the USA.

Original publications are not included in the electronic version of the dissertation.

904. Hinttala, Reetta (2006) Genetic causes of mitochondrial complex I deficiency in children
905. Niemi, Antti (2006) Röntgenhoitajien turvallisuuskulttuuri säteilyn lääketieteellisessä käytössä—Kulttuurinen näkökulma
906. Isokangas, Juha-Matti (2006) Endovascular treatment of 467 consecutive intracranial aneurysms in Oulu University Hospital. Angiographic and clinical results
907. Tähtinen, Tuula (2006) Insuliiniresistenssiin liittyvät kardiovaskulaariset riskitekijät suomalaisilla varusmiehillä. Tupakoinnin yhteys riskitekijöihin
908. Metsänen, Miia (2007) Thought disorder as a predictive sign of mental disorder. A study of high-risk and low-risk adoptees in the Finnish Adoptive Family Study of Schizophrenia
909. Herva, Anne (2007) Depression in association with birth weight, age at menarche, obesity and metabolic syndrome in young adults. The Northern Finland 1966 Birth Cohort Study
910. Lauronen, Erika (2007) Course of illness, outcome and their predictors in schizophrenia. The Northern Finland 1966 Birth Cohort study
911. Majava, Marja (2007) Molecular genetics of Stickler and Marshall syndromes, and the role of collagen II and other candidate proteins in high myopia and impaired hearing
912. Peltoniemi, Annu (2007) Terveystuotohenkilöstön valmiudet ohjata hemofiliaa sairastavia ja heidän perheitään
913. Daavittila, Iita (2007) Genetic risk factors for lumbar intervertebral disc disease characterized by sciatica
914. Pylkäs, Katri (2007) ATM, ATR and Mre11 complex genes in hereditary susceptibility to breast cancer
915. Kinnunen, Tuija (2007) Keuhkohtaumataudin sairaalahoito Suomessa: hoitoajan pituus ja sen yhteys ennusteeseen
916. Pursiainen, Ville (2007) Autonomic dysfunction in early and advanced Parkinson's disease
917. Paldanius, Mika (2007) Serological studies on *Chlamydia pneumoniae* infections
918. Parkkila, Timo (2007) Sutter metacarpophalangeal arthroplasty in rheumatoid patients

Book orders:
OULU UNIVERSITY PRESS
P.O. Box 8200, FI-90014
University of Oulu, Finland

Distributed by
OULU UNIVERSITY LIBRARY
P.O. Box 7500, FI-90014
University of Oulu, Finland

S E R I E S E D I T O R S

A
SCIENTIAE RERUM NATURALIUM
Professor Mikko Siponen

B
HUMANIORA
Professor Harri Mantila

C
TECHNICA
Professor Juha Kostamovaara

D
MEDICA
Professor Olli Vuolteenaho

E
SCIENTIAE RERUM SOCIALIUM
Senior Assistant Timo Latomaa

E
SCRIPTA ACADEMICA
Communications Officer Elna Stjerna

G
OECONOMICA
Senior Lecturer Seppo Eriksson

EDITOR IN CHIEF
Professor Olli Vuolteenaho

EDITORIAL SECRETARY
Publications Editor Kirsti Nurkkala

ISBN 978-951-42-8427-4 (Paperback)

ISBN 978-951-42-8428-1 (PDF)

ISSN 0355-3221 (Print)

ISSN 1796-2234 (Online)

